THE DYNAMICS AND CONTROL OF SOLAR-SAIL SPACECRAFT IN DISPLACED LUNAR ORBITS

A Dissertation

Submitted to the Faculty

of

Purdue University

by

Geoffrey George Wawrzyniak

In Partial Fulfillment of the

Requirements for the Degree

of

Doctor of Philosophy

August 2011

Purdue University

West Lafayette, Indiana

For my boy and his mama.

ACKNOWLEDGMENTS

I would like to thank my advisor, Professor Kathleen Howell, for her guidance of my research and many other efforts during the past six years. It has been a privilege to work with her. Her unwavering attention to detail and high standards have improved my work. I also thank her for her meticulous editorial direction in this dissertation and our other collaborations.

Professors James Longuski, Martin Corless, and Dan DeLaurentis are a few of the many great teachers at Purdue University and I am grateful to them for serving on my advisory committee. I have enjoyed the many conversations I have had with each of them. It has been a privilege to work with them as well.

Fellow research group member Diane Craig Davis has been a good friend and colleague. I also appreciate the help I have received from Tom Pavlak, Cody Short, and the many other graduate students in our research group.

I must also thank my family for their support. Among the many values that my parents instilled in me, the importance of an education was established at early age. My doctorate is a direct result of those values and their efforts as parents.

The person that deserves my greatest appreciation is my wife, Celeste. She has encouraged, loved, and patiently supported me through this process. She has also accomplished more in the past three years than I ever can in a lifetime by giving birth to our son, Paul, and raising him to be a great little kid. For his part, Paul has provided me with necessary breaks and motivation.

Finally, I would like to acknowledge the School of Aeronautics and Astronautics and the Mathematical Sciences Department for funding my program through various teaching assignments. Additional sponsorship was provided by the Purdue Forever Fellowship and the Space Shuttle Memorial Scholarship.

PROLOGUE

"The analysis of realistic-sail-shifted ellipses (or quasi-ellipses) should be another topic of frontier astrodynamics, to be addressed in M.S. theses or Ph.D. dissertations."

—Vulpetti, Johnson, and Matloff [1]

My motivation for pursuing a doctorate was my enthusiasm for spacecraft dynamics, navigation, and control. To me, the scientific or commercial purpose of a spacecraft serves to justify the fun of flying the vehicle. This investigation has allowed me apply that enthusiasm to the challenging and elegant problem of solar sails in a complex dynamical environment. Professor Howell is fond of talking about "working on the fringes" where interesting problems reside. Solar sailing is rich with areas for study, and I have enjoyed exploring this frontier of astrodynamics.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

ABSTRACT

Wawrzyniak, Geoffrey G. Ph.D., Purdue University, August 2011. The Dynamics and Control of Solar-Sail Spacecraft in Displaced Lunar Orbits. Major Professor: Kathleen C. Howell.

Trajectory generation for any spacecraft mission application typically involves either well-developed analytical approximations or a linearization with respect to a known solution. Such approximations are based on the well-understood dynamics of behavior in the system. However, when two or more large bodies (e.g., the Earth and the Moon or the Sun, the Earth and the Moon) are present, trajectories in the multi-body gravitational field can evolve chaotically. The problem is further complicated when an additional force from a solar sail is included. Solar sail trajectories are often developed in a Sun-centered reference frame in which the sunlight direction is fixed. New challenges arise when modeling a solar-sail trajectory in a reference frame attached to the Earth and the Moon (a frame that rotates in inertial space). Advantages accrue from geometry and symmetry properties that are available in this Earth–Moon reference frame, but the Sun location and the sunlight direction change with time.

Current trajectory design tools can reveal many solutions within these regimes. Recent work using numerical boundary value problem (BVP) solvers has demonstrated great promise for uncovering additional and, sometimes, "better" solutions to problems in spacecraft trajectory design involving solar sails. One such approach to solving BVPs is the finite-difference method. Derivatives that appear in the differential equations are replaced with their respective finite differences and evaluated at node points along the trajectory. The solution process is iterative. A candidate solution, such as an offset circle or a point, is discretized into nodes, and the equations that represent the relationships at the nodes are solved simultaneously. Finite-difference

methods (FDMs) exploit coarse initial approximations and, with the system constraints (such as the continuous visibility of the spacecraft from a point on the lunar surface), to develop orbital solutions in regions where the structure of the solution space is not well known. Because of their simplicity and speed, the FDM is used to populate a survey to assist in the understanding of the available design space. Trajectories generated by FDMs can also be used to initialize other nonlinear BVP solvers.

Any solution is only as accurate as the model used to generate it, especially when the trajectory is dynamically unstable, certainly the case when an orbit is purposefully offset from the Moon. Perturbations, such as unmodeled gravitational forces, variations in the solar flux, as well as mis-modeling of the sail and bus properties, all shift the spacecraft off the reference trajectory and, potentially, into a regime from which the vehicle is unrecoverable. Therefore, some type of flight-path control is required to maintain the vehicle near the reference path. Reference trajectories, supplied by FDMs, are used to develop guidance algorithms based on other, more accurate, numerical procedures, such as multiple shooting.

The primary motivation of this investigation is to determine what level of technology is required to displace a solar sail spacecraft sufficiently such that a vehicle equipped with a sail supplies a continuous relay between the Earth and an outpost at the lunar south pole. To accomplish this objective, numerical methods to generate reference orbits that meet mission constraints are examined, as well as flight-path control strategies to ensure that a sailcraft follows those reference trajectories. A survey of the design space is also performed to highlight vehicle-performance and ground-based metrics critical to a mission that monitors the lunar south pole at all times. Finally, observations about the underlying dynamical structure of solar sail motion in a multi-body system are summarized.

1. INTRODUCTION

New technologies enable new mission applications. Solar sails, only recently successfully operated in space, have long been proposed as a means to accomplish mission goals normally impossible for spacecraft in Keplerian orbits. In 1993, Robert L. Forward proposed and patented the Statite concept, a satellite equipped with a solar sail that uses solar radiation pressure to counteract gravity and hover statically above a planetary pole [2]. Concurrently, Colin McInnes identified equilibrium surfaces in the circular restricted Sun–Earth system that describe regions of space where a solar sail spacecraft (sailcraft) can be positioned indefinitely [3,4]. These original investigations were based on simple assumptions. More recent research efforts by McInnes, as well as other researchers, have improved the results from the original studies [5–19].

Sails supply a force to a spacecraft via light reflected from the sail surface. From special relativity's mass-energy equivalence, light, in the form of photons, possesses momentum. When light is reflected off a surface, incident and reflected forces are imparted to the sail. On an ideal solar sail, these forces are equal, and the resulting force is normal to the plane of the sail face and is directed away from the Sun. This force depends on the sail's attitude, the distance of the sail from the Sun, and sail properties such as reflectivity and thickness (or areal density). As a consequence of this sail force, trajectories and attitude profiles are tightly coupled for solar sail spacecraft.

A complex and rich dynamical environment exists in the Earth-Moon-Sun-sail regime; new analytical and numerical tools are required to understand, control, and exploit the sail force for mission design. Recent investigations were originally motivated by the 2004 Presidential Vision for Space Exploration, that is, a return of humans to the Moon and the establishment of an outpost at the lunar south pole by 2020 as a precursor to a human mission to Mars [20]. While this goal has since been

revised, the announcement of such a potential outpost motivated studies for establishing a communications architecture to support mission personnel [21–24]. Because line-of-sight transmission to the Earth (or to a relay in low-Earth orbit) is not guaranteed at the lunar south pole, multiple spacecraft in Keplerian orbits about the Moon are necessary to serve as communications relays [25–27]. Exotic solutions that exploit halo orbits about the cis- and trans-lunar Lagrange points have also been examined [28].

A spacecraft that hovers above a pole is often called a "polesitter." Using this concept, a sailcraft can be positioned as a relay between the lunar outpost and ground stations on Earth. Several researchers have examined this design problem. A team from the National Aeronautics and Space Administration's Jet Propulsion Laboratory (JPL) designed a non-periodic trajectory in the region below the Earth–Moon L_2 Lagrange point [29]. In this JPL study, a 235–255 kg spacecraft completes a conventional transfer from the Earth to a lunar orbit. Once in orbit, a solar sail with a maximum characteristic acceleration of 1.338 mm/s^2 ("a modest improvement in contemporary solar sail technology" [29]) is deployed and used to keep the spacecraft below the trans-lunar Lagrange point. A carrier vehicle is jettisoned, and the mass of the remaining vehicle is 195 kg. Sailcraft (and low-thrust propulsion) trajectories that are displaced below the Moon were developed by Ozimek, Grebow, and Howell at Purdue University using high-fidelity collocation techniques [30–33]. In the low-thrust investigations [31, 32], an NSTAR low-thrust engine is used to enable an expanding spiral departure that delivers a spacecraft of 500 kg mass from low-Earth orbit (LEO) to a propellant-optimal orbit below the trans-lunar Lagrange point. Simo and McInnes employ linear control theory for displacing halo orbits from the vicinity of the Earth–Moon Lagrange points [14–19]. Using an approach reminiscent of the collocation schemes exploited by Ozimek et al., Wawrzyniak and Howell have developed finite-difference methods (FDMs), augmented to include path constraints, to generate periodic sailcraft orbits displaced below the Moon [34]. This last approach is easily understood and implemented at the expense of reduced accuracy. Additionally,

schemes can be employed that use FDMs to compute a multitude of orbits in a short amount of time. These finite-difference methods are used successfully to examine the design space for displaced, lunar sailcraft orbits [35].

While each of the preceding techniques generates trajectories and nominal attitude profiles, all of the solutions are open loop: the prescribed attitude profile *is* the control profile that must be followed to remain on the baseline trajectory. Any deviation from the specified profile results in a sailcraft that diverges from the design trajectory. Additionally, the trajectories are merely numerical approximations to the solutions of the differential equations. Since a trajectory and its control profile are approximations, simply implementing the approximations will result in a trajectory that diverges from the desired path. Therefore, additional control schemes are required to correct the trajectory to track the desired path.

This investigation examines multiple aspects of mission design for solar sail applications in the Earth–Moon system. The design space is not well known. While the underlying dynamical structure of the flow through this region of space is partially revealed through the generation of the equilibrium surfaces, numerical methods that do not require a thorough understanding of the system dynamics (or a reasonable initial guess) are useful for generating orbits. Using these tools, a survey of the design space aids in developing insight into the solutions and offers metrics (such as attitude turn rates) for use in assessing the feasibility of each candidate solution. Finally, given a candidate solution that is feasible and desirable, control laws are developed to maintain the orbit.

1.1 History of solar sailing trajectory design and flight-path control

Although solar sailing has only recently been demonstrated in flight, space travel and solar sailing date back to Kepler. In his 1610 response to Galileo's *Sidereus Nuncius*, the first publication based on observations with a telescope, Kepler wrote

> [A]s soon as somebody demonstrates the art of flying, settlers from our species of man will not be lacking.... Given ships or *sails adapted to the breezes of heaven* (emphasis added), there will be those who will not shrink for even that vast expanse. Therefore, for the sake of those who, as it were, will presently be on hand to attempt this voyage, let us establish the astronomy, Galileo, you of Jupiter, and me of the Moon [36].

Kepler also correctly postulated that the rays of the Sun are directly responsible for the formation of comet tails, marking the first instance in which light is believed to exert a pressure on matter [37, 38]. In 1873, Maxwell deduced that electromagnetic radiation (including light) exerts a pressure, as a consequence of his theory of electromagnetism [39]. Bartoli independently determined that light must exert a pressure as a consequence of the second law of thermodynamics in 1876 [40]. The Maxwell-Bartoli theory of radiation pressure was experimentally confirmed by Lebedev in 1900 [41] as well as Nichols and Hull in 1901 [42].

Practical concepts for solar sailing have existed for approximately 100 years, beginning with Tsiolkovsky and Tsander in the 1920s. A team at JPL completed the first serious mission study in the late 1970s for a rendezvous with Halley's Comet. A watershed effort by McInnes in the 1990s and the publication of his Ph.D. dissertation as a textbook on solar sailing re-invigorated interest in solar sailing as a research topic [6]. Solar sail concepts have been funded in competitions sponsored by the National Aeronautics and Space Administration's (NASA) New Millennium Program, created in 1995 to improve technological readiness levels for new, "leading-edge" technologies.

Early trajectory design work by McInnes focuses, in part, on non-Keplerian orbits. Keplerian orbits are characterized in terms of conic sections (e.g., ellipses, parabolas, hyperbolas) where a primary body is located at the focus of the section. Spiral trajectories, common in low-thrust space missions, including solar sailcraft, can be considered to be perturbed Keplerian ellipses. Non-Keplerian orbits frequently occur within the context of the restricted three-body problem as orbits associated with the

Lagrange points. However, the continuous force supplied by a high-performance solar sail might allow hovering. Such spacecraft orbits can be described as "displaced" or "offset." Examples of such offset orbits with respect to the Sun, as well as a planet in orbit about the Sun, are illustrated in Fig. 1.1 [5, 6, 43]. Orbit (a) appearing in the left plot of Fig. 1.1 is in a plane parallel to the plane containing the $\hat{\boldsymbol{\imath}}$ and $\hat{\boldsymbol{\jmath}}$ unit vectors, coaxial with the $\hat{\boldsymbol{k}}$ unit vector originating at Sun. The off-axis orbit on the right, Orbit (b), in Fig. 1.1 is centered on the $\hat{\boldsymbol{h}}$ axis which is perpendicular to $\hat{\boldsymbol{\imath}}$ in the plane containing $\hat{\boldsymbol{\jmath}}$ and $\hat{\boldsymbol{k}}$. The final example near the planet, Orbit (c), is coaxial with $\hat{\boldsymbol{\jmath}}$ and its center is in opposition.

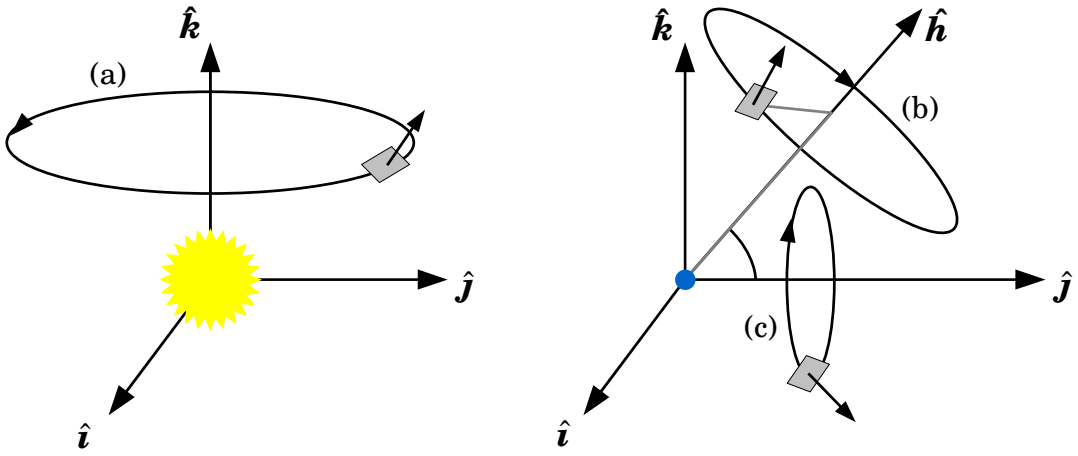
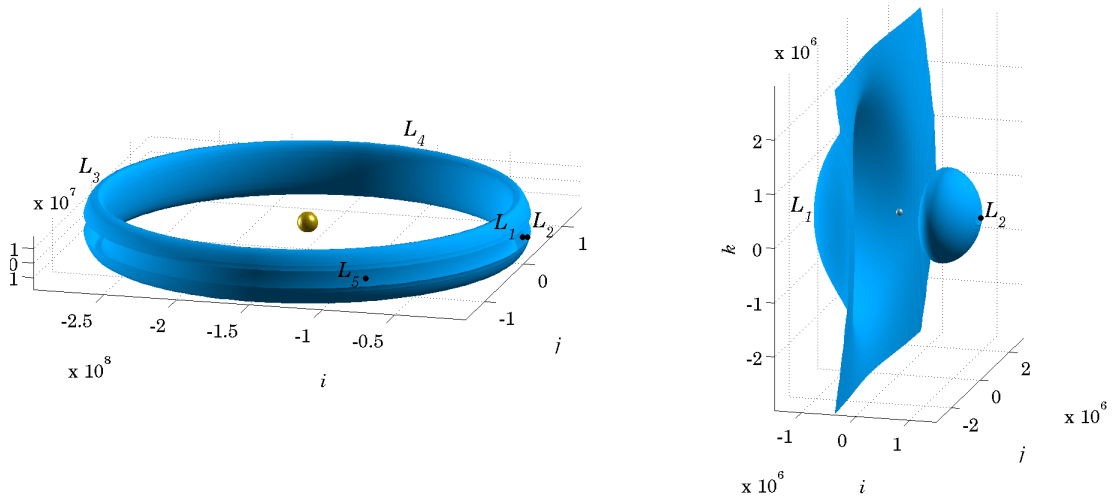


Fig. 1.1: Three examples of offset orbits. Each gray square represents a solar sail; the arrow attached to each sail represents the direction of the sail-face normal.

McInnes also examined artificial equilibrium points in circular restricted three-body (CR3B) systems. Five equilibrium solutions exist in a CR3B formulation; these are the well-known Lagrange points. However, with the addition of a sail, two distinct equilibrium surfaces emerge in the Sun–planet CR3B system, as seen in Fig. 1.2 and defined by an infinite number of artificial equilibrium points. Except for a dimple in the vicinity of the planet, the first surface is essentially a torus with an outer radius equal to the Sun–planet distance (the circumferential gap about the exterior



(a) Equilibrium torus about the Sun

(b) Dimple in torus section and secondary lobe

Fig. 1.2: Two views of the equilibrium surface for a sail with a characteristic acceleration of $a_c = 1.70$ mm/s^2 in the Sun–Earth rotating frame. The Sun and the Earth are magnified by a factor of 10 in their respective figures. Positions of the Lagrange points are denoted with black dots. Note that some points are obscured by the surface in these views.

of the torus is an artifact of rendering). Four of the Lagrange points lie on the outer diameter of this torus. The second surface is a lobe on the far side of the planet. The remaining Lagrange point, L_2, is located on the far side of the lobe. The size of these surfaces depends on the sail's characteristic acceleration, that is, a measure of the acceleration that a sail can supply given its reflectivity properties and the mass of the entire sailcraft (sail, bus, payload, structure, etc.). If a sailcraft is positioned on an equilibrium surface with an appropriate attitude, it will remain there indefinitely unless perturbed. Perhaps the most common application of these equilibrium surfaces is the generation of an artificial equilibrium point closer to the Sun than L_1. The Geostorm (a.k.a., Heliostorm) Warning Mission would position a sailcraft at this sub-

L_1 point to provide an early warning of incoming coronal mass ejections that cause geomagnetic storms [44–46].

To control the spacecraft flight path in the Geostorm class of missions, both Yen [44] and Sauer [45] control to the sub-L_1 point as opposed to an orbit about that point as suggested by Lisano et al. [46]. Lawrence and Piggott compare a linear-quadratic control scheme and a Gramian controller to maintain the trajectory about the sub-L_1 point [47]. Farrés and Jorba exploit dynamical systems theory to design station-keeping maneuvers with a solar sail to maintain orbits about a sub-L_1 point and move to other points on the equilibrium surface [11–13]. Waters and McInnes linearize relative to fixed-points that define a solar-sail libration-point orbit near the sub-L_1 point, and then produce an optimal control scheme that delivers the trajectory to those fixed points [48].

Another class of mission applications often examined for solar sail trajectory design comprises optimal (minimum-time) transfers from the Earth to other solar system bodies. Sauer develops continuous control laws to orient the sail-face normal for mission trajectories destined to reach the inner planets and the asteroid Eros [49]. Otten and McInnes employ a fixed-attitude steering law for a two-dimensional Earth-to-Mars transfer whereby the sail is restricted to a fixed orientation during each of N evenly spaced segments along the trajectory [50]. Mengali and Quarta further restrict the discrete attitude problem by the providing a finite set of possible attitudes for the sail during a two-dimensional transfer [51]. Mengali and Quarta do allow two of the five admissible controls to be determined as optimization parameters. However, all of these optimal steering laws are open loop and do not accommodate errors in the system. Other investigations into optimal solar-sail steering laws for interplanetary transfer are available in Refs. [52–55].

Recently, NASA's Constellation Program generated renewed interest in missions to the Moon. A single solar sail spacecraft has been proposed as a possible solution to the lunar communications coverage problem. West outlines a mission architecture whereby a sailcraft hovers in the vicinity below the Earth–Moon L_2 point [29]. Oz-

imek et al. applies low- and high-degree Gauss-Lobatto rules to a collocation scheme to generate solar sail trajectories in the Earth–Moon system [30, 33]. The local error along the trajectory path is proportional to Δt^5 and Δt^{12}, respectively, due to truncation in the approximating polynomials [56]. The problem is not formulated as an optimization problem. Rather, the authors employ this approach to solve the equations of motion iteratively. The converged solution rarely resembles the initial guess; however, the solution does depend on the initial guess. Grebow et al. exploit these same collocation techniques to generate low-thrust trajectories that remain in view of a point on the lunar surface at all times [31, 57].

McInnes also examines equilibrium surfaces in the Earth–Moon system [3]. These surfaces are instantaneous because the Sun moves with respect to the fixed Earth and Moon. The sail attitude required for equilibrium depends on both the pseudo-gravity gradient defined in the rotating Earth–Moon system as well as the position of the Sun. Short-term missions to points on the instantaneous equilibrium surface are postulated in McInnes [3]. This work has been expanded in Simo and McInnes [14–19] to include linear control laws and stability analysis for the motion of solar sails in orbits displaced from the Earth–Moon rotating plane in the vicinity of the lunar Lagrange points.

1.2 Solar sail missions

Mission designers typically incorporate solar radiation pressure (SRP) as a perturbation. For a solar sail, SRP is exploited as a primary means of propulsion and, thus, may significantly affect the trajectory, enabling new concepts in trajectory design. Solar sailing dates back to Tsiolkovsky, Tsander, and Oberth in the 1920s [6]. If the definition of a solar sail includes any spacecraft that exploits SRP, an early "sail" mission involved the Mariner 10 spacecraft that flew by Mercury three times in 1974–75 and exploited SRP for attitude control [58]. More recently, MESSENGER employed SRP for trajectory control during the 2008–09 Mercury flybys and for angu-

lar momentum management [59]. These two missions "sailed" using their solar panels and not with a highly reflective, lightweight sheet of large dimensions conventionally defined as a solar sail. However, small sails attached to traditional spacecraft have been proposed for attitude and trajectory control [60–66] over the last 50 years.

Only recently has a spacecraft flown with a solar sail as its only means of propulsion. In May of 2010, only 2.5 years after initial development, the Japanese Space Agency, JAXA, launched a solar sail spacecraft named IKAROS (Interplanetary Kitecraft Accelerated by Radiation Of the Sun) as a secondary payload to another spacecraft (*Akatsuki*, "Dawn" in Japanese, also known as the Venus Climate Orbiter). The sailcraft is the first in-flight demonstration of solar sailing [67]. After separating from the main launch payload, the solar sail on IKAROS fully deployed on June 10, 2010. An image from one of the separation cameras (DCAM2) appears in Fig. 1.3. The

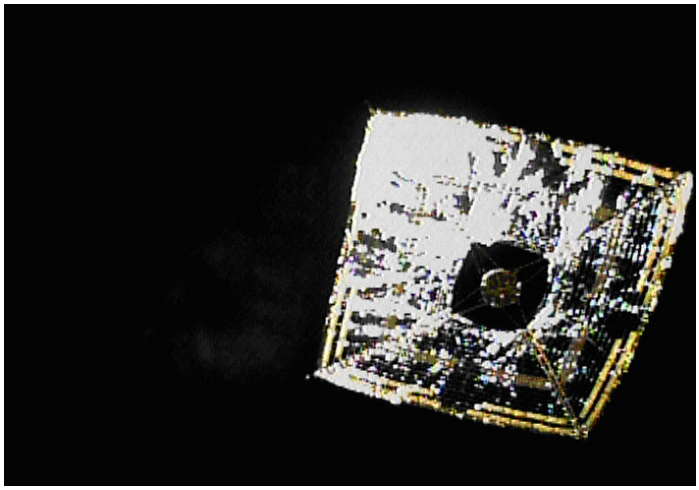


Fig. 1.3: The IKAROS spacecraft shortly after sail deployment. Credit: JAXA.

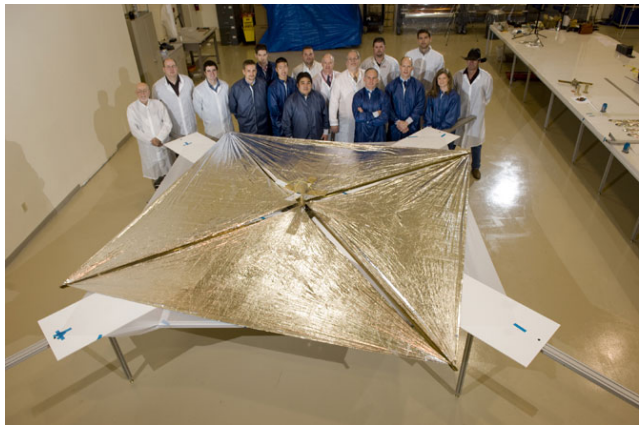
sail is 14-by-14 meters and the vehicle spins at 2 rotations per minute (RPM). Liquid crystal display (LCD) panels are arranged circumferentially around the sail. The vehicle's attitude is controlled by turning the LCD panels on-and-off on opposite sides of the sail at a rate of 2 RPM, creating a net torque on the sail and allowing the vehicle to target the Venus B-plane [67, 68]. The IKAROS spacecraft flew by Venus

on December 8, 2010 [69]. Its mission has been extended to March 2012, when it is expected to complete one orbit about the Sun [70].

On November 19, 2010, NASA launched the NanoSail-D2 ("D" for "deploy, deorbit, demonstration, and drag" and "2" because it is the ground spare of the NanoSail-D spacecraft that was lost in a launch vehicle failure) inside the FASTSAT satellite, which was one of eight spacecraft onboard a Minotaur IV launch vehicle [71]. NanoSail-D2 is a three-element CubeSat; two elements contained the stowed 10-by-10 foot sail and one element is the spacecraft bus [72]. Because of an unknown malfunction, NanoSail-D2 ejected from FASTSAT nearly two months later than scheduled on January 17, 2011, into a 72° inclination, 640-km altitude orbit [71]. Three days later, the sail deployed, and approximately 11 hours after that, the batteries lost power and radio contact ended, all as expected. The primary purpose of the project is to demonstrate the stowage and deployment of a small solar sail in LEO and to use high-atmospheric drag to deorbit the vehicle. As of June 13, 2011, the perigee and apogee altitudes are 575 km and 609.2 km, respectively [73]. While the vehicle is no longer transmitting to Earth, it is visible from the ground, and amateur photographers are encouraged to submit photos to spaceweather.com of NanoSail-D2 streaking across the night sky. One such image, taken on January 27, 2011, from Buenos Aires with a one-second exposure, appears in Fig. 1.4(a). The sail appears as a faint diagonal streak under the text. Another image taken after the ground deployment test of NanoSail-D appears in Fig. 1.4(b) and illustrates the sail size. Drag has caused the sailcraft to reorient so that its sail-face normal is perpendicular to the direction of motion [71]. The result is reduced atmospheric drag and a longer orbital lifetime, extending into 2012.

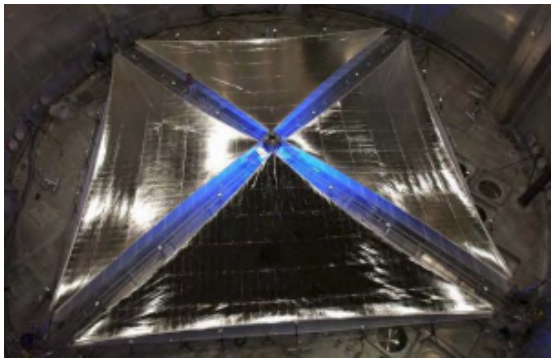
Prior to the NanoSail effort, sailcraft technologies were regularly entered the New Millennium Program's Space Technology (ST) competitions. The most recent competition, ST9, included the ground deployment of two 20-by-20 meter sails in the 100 foot vacuum chamber at NASA's Plum Brook Space Power Facility in Sandusky, Ohio. The two prototypes appear deployed in the vacuum chamber in Fig. 1.5. If

(a) Credit: Enzo De Bernardini, used with permission.



(b) Credit: NASA/MSFC/D. Higginbotham.

Fig. 1.4: NanoSail-D2 in space and NanoSail-D after a ground deployment test.



(a) ATK sail

(b) L'Garde sail

Fig. 1.5: Two ST9 candidate sails. Note the people for scale in the image on the right. Credit: NASA/MSFC.

solar sails had been selected to proceed to the next level in the ST9 competition, the L'Garde sail would have been increased in size to 40 meters on each side and been deployed in space. Regardless, the New Millennium Program was defunded in 2009, and these sails represent the most advanced stage of large solar sail development at NASA [74]. However, because of these efforts, the most recent technology readiness

level (TRL) cited for solar sails is TRL6 [74], which indicates that sails have been demonstrated in a *relevant* environment, such as ground or space (e.g., a vacuum chamber). Demonstrating a system prototype "near or at the scale of the planned operational system" in a *space* environment is necessary for a technology to be considered at TRL7 [75]. The next TRLs reflect a concept that is "flight-qualified" and "flight-proven" for mission operations. After a draw-down in solar sail technology following the ST9 effort, a renewal of solar sail development within NASA is listed in the announcement of Crosscutting Capability Demonstrations as part of the focus of NASA's new Office of the Chief Technologist [76].

In addition to government efforts, the Planetary Society has developed a solar sail ready for flight called *LightSail-1* [77]. This mission is planned as the first of three solar sail missions by the Planetary Society. Similar to NanoSail-D2, *LightSail-1* consists of a three-element CubeSat. Launch is anticipated for November 2011.

Researchers at the University of Strathclyde's Advanced Space Concepts Laboratory have developed a catalog for solar sail missions [78]. Rendezvous missions to the inner and outer solar system are better accomplished using traditional technologies, such as electric or chemical propulsion. Highly non-Keplerian orbits (such as pole sitters), solar observation, and trajectories beyond the outer planets are significantly enabled by solar sails. However, these categories of missions are significantly more advanced than the ground deployments, LEO mission, and Venus flyby already demonstrated with solar sailing technology. Sails must become both lighter and larger; the total vehicle mass (including the sail itself) per sail area must decrease and the sail area must increase to enable these advanced missions.

Sailing technology is still in its infancy, and advancing the TRL requires in-space demonstrations, which can be costly and risky. Project scientists and program managers must be convinced that solar sails are low-risk investments for their applications. The continuing advancement of solar sail technology might again occur in conjunction with other development programs, such as a hybrid propulsion system. For example, a solar kite might be employed to assist in attitude control, as is the case with

MESSENGER. Alternatively, a solar sail can be combined with a low-thrust device to develop a hybrid system. By combining propulsion systems, programmatic risk is reduced, and the degree of difficulty in advancing through TRLs is similarly reduced. The Strathclyde researchers assert that in-space demonstrations such as these will increase the confidence in solar sailing as a technology and eventually lead to a sail that serves as the primary propulsion system on larger-scale missions [78].

1.3 The physics of solar radiation pressure

A brief description of the source of solar radiation pressure is necessary to develop subsequent solar sail force models. A more thorough explanation is available in McInnes [6]. It is known that the irradiance, or the incident power on a surface, from solar radiation, I_0, at 1 AU is approximately 1368 W/m^2, that is, through every square meter at the Sun–Earth distance, 1368 Joules of energy from the Sun pass per second. This value varies by 3–4% per year as a result of the Earth's eccentric orbit about the Sun and because the Sun does not emit a constant amount of energy. Similar to gravity, irradiance from a point source also scales as the square of the inverse of the distance from the Sun, that is,

$$I = I_0 \left(\frac{\mathrm{AU}}{R} \right)^2 \tag{1.1}$$

where R is the distance from the Sun to the region of interest. The term *insolation* is used when describing the irradiance from the Sun.

Momentum transfer from photons imparts a force on any illuminated body. Usually, this force is imperceptible. However, in a space environment composed of field forces and essentially free of contact forces, photon momentum will appreciably accelerate the body. The pressure force from electromagnetic radiation can be derived from the principles of either relativity and quantum mechanics or electromagnetism [6]. For pressure due to solar radiation, the insolation is divided by the speed of light, that is,

$$P = \frac{I}{c} \tag{1.2}$$

where P is the pressure from solar radiation and c is the speed of light in a vacuum. Therefore, the magnitude of the solar radiation pressure (SRP) force on any body that is 1 AU from the Sun is approximately 4.56×10^{-6} N/m^2. For a flat body of area A, the resulting force is

$$f_\ell \hat{\boldsymbol{\ell}} = PA(\hat{\boldsymbol{\ell}} \cdot \hat{\mathbf{u}})\hat{\boldsymbol{\ell}} \tag{1.3}$$

where $\hat{\boldsymbol{\ell}}$ is the direction of incident radiation, or the Sun-light vector, and $\hat{\mathbf{u}}$ is the direction of the sail-face normal.[1] Solar sailing is enabled by employing a large, flat, low-mass surface to maximize the resulting acceleration on the space vehicle. Sails also reflect much (and ideally all) of the incident energy in such a way that additional acceleration is generated opposite to the direction of the reflected photons. For an ideal solar sail, that reflected force is

$$f_r \hat{\mathbf{u}}_r = -PA(\hat{\mathbf{u}}_r \cdot \hat{\mathbf{u}})\hat{\mathbf{u}}_r \tag{1.4}$$

where $\hat{\mathbf{u}}_r$ is the direction of reflected radiation. These forces appear in Fig. 1.6. The model for an ideal solar sail is based on the assumption that all reflection is specular, and, for specular reflection, the angle of incidence and the angle of reflection are identical. The angle of incidence is also known as the pitch angle, α, and the dot products in Eqs. (1.3) and (1.4) are, therefore, both $\cos\alpha$. Summing the forces in Eqs. (1.3) and (1.4) leads to a net force in the direction of the sail-face normal, that is,

$$f\hat{\mathbf{u}} = 2PA\cos^2\alpha\,\hat{\mathbf{u}} \tag{1.5}$$

In reality, however, this force is attenuated by some energy that is either absorbed by the sail (and re-emitted as heat) or reflected diffusely, that is, some photons scatter at an angle other than α from the sail-face normal. Assumptions of an ideal model verse a non-ideal, higher-fidelity model are addressed in Appendix A.

Finally, the characteristic acceleration, a_c, is introduced to incorporate the pressure, area, mass, and reflective efficiency of the sail at 1 AU. It is formulated as

$$a_c = \frac{2(4.56)\nu}{\sigma\ [\text{g m}^{-2}]} \quad [\text{mm s}^{-2}] \tag{1.6}$$

[1]Vectors are denoted with boldface. The symbol " ˆ " indicates a unit vector.

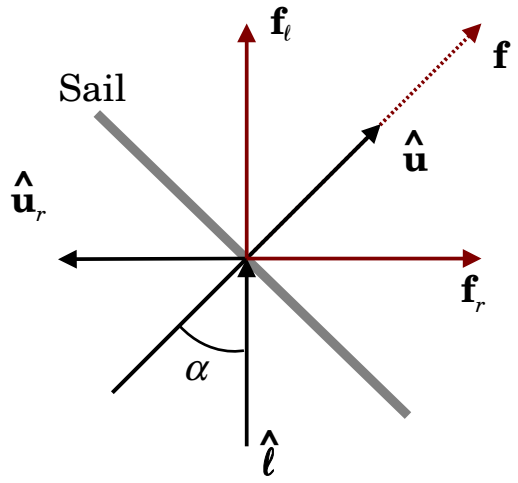


Fig. 1.6: Forces from an ideal reflection on a solar sail [6].

where ν is the reflective efficiency, typically between 0.85 and 0.90 for a solar sail, and σ is the areal density of the sailcraft (also known as the total solar sail loading), or

$$\sigma = \frac{m}{A} \tag{1.7}$$

where m is the total mass of the vehicle and A is the area of the sail. An alternate form of Eq. (1.6) separates the mass of the sail and its associated support structure, m_s, from the mass of the spacecraft bus and payload, m_p. In this form,

$$a_c = \frac{2\nu P}{\sigma_s + (m_p/A)}, \quad \sigma_s = \frac{m_s}{A} \tag{1.8}$$

where $P = 4.56 \times 10^{-6}$ N/m^2 and σ_s is termed the sail assembly loading. The acceleration contribution from an ideal solar sail spacecraft is then

$$\mathbf{a}_s = a_c \left(\frac{\text{AU}}{R}\right)^2 \cos^2\alpha\, \hat{\mathbf{u}} \tag{1.9}$$

The analyses in this investigation occur in the Earth–Moon regime, where $R \approx$ AU, and the squared fraction in Eq. (1.9) reduces to 1.

One goal for solar sail hardware technology development is to maximize the surface reflectivity and minimize the total solar sail loading. A recent sailcraft design

for NASA's Space Technology competition (ST9) that was built by L'Garde possesses an overall characteristic acceleration of $a_c = 0.58$ mm/s^2; however, the characteristic acceleration of the sail and its support structure alone is closer to 1.70 mm/s^2 [79]. The total sail loading parameter for ST9 is then $\sigma = 14.1$ g/m^2 and the sail assembly loading is $\sigma_s = 4.8$ g/m^2. The equations of motion do not separate the sail and its associated structure from the overall spacecraft. In a survey of recently designed solar sails, a characteristic acceleration of 1.70 mm/s^2 is at the limit of existing sailcraft technology. In fact, the characteristic accelerations of the two solar sail spacecraft demonstrated in space are approximately 0.004 mm/s^2 for IKAROS [80] and 0.02 mm/s^2 for NanoSail-D2 [81]. Many future applications, including the polesitter highlighted in this investigation, require characteristic accelerations on the order of 0.5 to 2.0 mm/s^2, and the sizes of these future sails are expected to be 100 to 300 meters on each side [78].

1.4 Present scope

Fundamental to the mission design of a sailcraft in an orbit displaced below the Moon is an understanding of the underlying dynamical structure, an ability to generate reference orbits, and an ability to control the spacecraft to fly near those reference orbits; all are of equal importance to the understanding and development of this mission application. Because of the forces required from the solar sail to maintain a periodic orbit, the regime in which it operates is highly sensitive to modeling accuracy and not well suited to analytical methods. A numerical tool is developed that generates sailcraft orbits in idealized gravitational regimes, such as the Earth–Moon system, as well as high-fidelity, ephemeris-based gravitational models. Numerically generated orbits are simpler to produce and may be more accurate than analytically derived reference trajectories. The resulting orbits can be understood within the context of the underlying dynamical structure. The orbits essentially "hover" near a point on an instantaneous equilibrium surface, a surface that depends on the frame

in which the system equations are formulated. These surfaces are composed of sets of artificial Lagrange points, attached to theoretical Lagrange points in a circular restricted three-body system. Finally, since any reference solution is inexact, adjustments to the control profile are necessary. Control laws based on numerical techniques to solve boundary value problems are developed. This investigation is organized as follows:

- Chapter 2: The dynamical models used in this investigation are introduced. Differential equations for the motion of a solar sail spacecraft within the context of the Earth–Moon CR3B system are presented. Vector formulations for thrust control (i.e., the sail-face normal direction) are derived. Roy and Ovenden's Mirror Theorem (1955) is adapted for solar sail trajectories, and models for linear dynamics are introduced.

- Chapter 3: Various numerical tools for solving boundary value problems (BVPs) are presented. Shooting methods, collocation techniques, and finite-difference methods (FDMs) are discussed and relevant examples are furnished. The FDM approach receives greater emphasis as it is not currently as popular in the literature for solving BVPs as the other methods. An error analysis for the finite-difference algorithm is also supplied, as well as a formulation that incorporates mission constraints.

- Chapter 4: The FDM is employed as the foundation for a survey of the design space for solar sail trajectories that solve the lunar south pole coverage problem. Various initial guesses are exploited to enable the generation of tens-of-millions of simulations in a search for viable trajectories. Critical mission metrics are used to evaluate the plethora of solutions.

- Chapter 5 : Solutions from the numerical tools are linked to equilibrium properties of the Earth–Moon dynamical system. Solar gravity is minor compared to the gravity of the Earth and the Moon. Five sample reference trajectories are demonstrated to track instantaneous equilibrium surfaces that arise from

the balance of solar radiation pressure, gravity of the primary bodies, and the effects of a rotating reference frame.

- Chapter 6: Solutions developed in the idealized Earth–Moon rotating system with the FDM algorithm are applied to initialize a similar FDM process in a high-fidelity, ephemeris-based model. The procedure to transition trajectories designed in the simpler model to the higher-fidelity model is presented. The same five sample trajectories from Chapter 5 are used to illustrate the transition.

- Chapter 7: Three techniques for solving BVPs are employed to control the flight path of a solar sail within the context of the Earth–Moon circular restricted three-body model to reference trajectories developed in that model. The basis for the controller design is a turn-and-hold scheme, whereby the attitude of the sail is held constant in inertial space and then instantaneously turned one or more times to target the reference flight path at some future time.

- Chapter 8: Concluding remarks are offered along with recommendations for future research.

2. SOLAR SAILING DYNAMICS

The history of science and mathematics is largely the history of celestial mechanics. Ancient societies made those who could predict solar and lunar eclipses their high priests. Aristotle and Ptolemy influenced the western world for over a thousand years with their ideas on heavenly motion. Giants coupling mathematics and astronomy such as Copernicus, Kepler, and Galileo transformed humanity's understanding of its place in the universe. Newton described motion by incorporating mechanics, inventing new fields of mathematics, and describing the effect of gravity. Euler, Lagrange, Gauss, Bessel, Poincaré, and many other visionary thinkers of their respective times made significant contributions to the study of dynamics, especially as applied to celestial mechanics.

In the past one hundred years, a new branch of science, stemming from celestial mechanics, has emerged in the form of astrodynamics, which is the motion of artificial satellites, or spacecraft. Tsiolkovsky, Goddard, and Von Braun are a few of the significant contributors that established this field and are responsible for enabling spacecraft that reach low-Earth orbit, the Moon, Mars, and the outer planets. In recent years, new pathways through the solar system have been discovered by applying a mixture of classical and modern mathematical techniques to describe celestial and spacecraft motion.

The dynamical behavior of particles or bodies within any system is frequently described by a set of ordinary differential equations. The solutions to these equations, whether analytical or the result of numerical simulations, are employed to chart the motion of a particle in the presence of gravity and other forces affecting its motion. The equations of motion for a sailcraft in a rotating multi-body system are initially introduced. By assuming the system dynamics are governed by the gravity of the Earth and the Moon, as well as the effect of solar radiation pressure, interesting sym-

metries arise. Additionally, the equations of motion are nonlinear, and linearization about a reference path results in tools that are useful for generating solutions and controlling motion in a nonlinear regime.

2.1 Circular restricted three-body motion and an idealized solar sail

It is advantageous to formulate the lunar south pole coverage problem within the context of a circular restricted three-body (CR3B) system. Motion is modeled in a frame, R, that is fixed to the Earth and the Moon. The R frame rotates with respect to an inertially fixed frame, I, about a common $\hat{\mathbf{Z}}$ axis with an angular rate of ω. The angular velocity vector is then

$$\boldsymbol{\omega} = \left\{ \begin{array}{ccc} 0 & 0 & \omega \end{array} \right\}^{\mathrm{T}} \tag{2.1}$$

where superscript " $^{\mathrm{T}}$ " indicates a vector (or matrix) transpose. Both frames originate at a common barycenter. An illustration of these two frames appears in Fig. 2.1. A position vector in the inertial frame, I, is described by the variables $\left\{ \begin{array}{ccc} X & Y & Z \end{array} \right\}^{\mathrm{T}}$, and

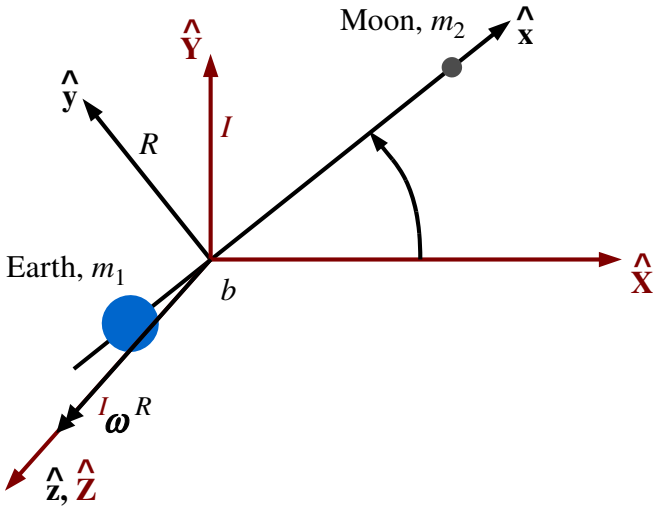


Fig. 2.1: The R frame rotating with respect to the I frame.

a position vector in the rotation frame, R, is described by the variables $\left\{ \begin{array}{ccc} x & y & z \end{array} \right\}^{\mathrm{T}}$. The velocity and acceleration corresponding to the first and second derivatives of a position element in the appropriate frame, where time is the independent variable, are denoted with a single dot and double dots, respectively (e.g., $\dot{x}$ and $\ddot{x}$).

Modeling behavior in the CR3B system typically involves characteristic quantities and nondimensionalization based on distances, masses, and time. To begin, the first characteristic quantity is GM^*, which is the sum of the gravitational parameters of the two primaries, that is,

$$GM^* = GM_1 + GM_2 \tag{2.2}$$

where GM_1 and GM_2 are the gravitational parameters of the Earth (398600.436 $\mathrm{km}^3/\mathrm{s}^2$) and the Moon (4902.799 $\mathrm{km}^3/\mathrm{s}^2$), respectively. The average distance between the Earth and the Moon is the characteristic quantity L^* and is equal to 384,400 km in this investigation. A final characteristic quantity is based on time and originates with Kepler's equation,

$$t^* = \sqrt{\frac{L^{*3}}{GM_1 + GM_2}} \tag{2.3}$$

Because a solar sail delivers an acceleration to the spacecraft, another quantity, the *system* characteristic acceleration, is also useful. This quantity is easily derived as the ratio of the characteristic length, L^*, to the square of the characteristic time, t^*, that is,

$$a^* = \frac{L^*}{t^{*2}} = 2.7307 \text{ mm/s}^2 \tag{2.4}$$

Nondimensionalization based on characteristic quantities is often used in astrodynamics to simplify the scale of the numbers involved in computation [82]. The mass ratio of the Moon to the sum of the masses of the two primaries is

$$\mu = \frac{GM_2}{GM_1 + GM_2} \tag{2.5}$$

The nondimensional gravitational parameter of the Earth is consequently $1 - \mu$. The nondimensional distance between the Earth and the Moon is 1. The period for one revolution of the Earth and the Moon about a common barycenter is one sidereal

month (27.321 days), and after dividing by t^*, the nondimensional period is 2π in nondimensional time units. As a result, the nondimensional rotational rate of the Earth–Moon system with respect to an inertial frame is 1, whereas the dimensional rotation rate is 2.67×10^{-6} rad/s^2.

After nondimensionalizing distance, mass, and time quantities, the development of the equations of motion in the R frame, rotating relative to an inertial frame, I, employs a kinematic expansion [82–85]. Consistent with McInnes [6], the nondimensional vector equation of motion for a spacecraft at a location $\mathbf{r}$ relative to the barycenter (center of mass of the primaries) is

$$ {}^{R}\mathbf{a} + 2\left({}^{I}\boldsymbol{\omega}^{R} \times {}^{R}\mathbf{v}\right) + \boldsymbol{\nabla} U(\mathbf{r}) = \mathbf{a}_s(t) \tag{2.6} $$

where the first term is the acceleration relative to the rotating frame (more precisely expressed as $\frac{{}^{R}d^2\mathbf{r}}{dt^2}$, where the left superscript R indicates a derivative in the rotating frame) and the second term is the corresponding Coriolis acceleration, which requires the velocity relative to the rotating frame, ${}^{R}\mathbf{v}$ (more precisely $\frac{{}^{R}d\mathbf{r}}{dt}$).[1] The angular-velocity vector, ${}^{I}\boldsymbol{\omega}^{R}$ (or $\boldsymbol{\omega}$), relates the rate of change of the rotating frame with respect to the inertial frame. In this nondimensional system, $\boldsymbol{\omega} = \left\{ \begin{array}{ccc} 0 & 0 & 1 \end{array} \right\}^{\mathrm{T}}$. The applied acceleration, from a solar sail in this case, is indicated on the right side by $\mathbf{a}_s(t)$. The centripetal and gravitational accelerations are combined in a pseudo-gravity gradient, $\boldsymbol{\nabla} U(\mathbf{r})$, that is,

$$ \boldsymbol{\nabla} U(\mathbf{r}) = \left(\boldsymbol{\omega} \times (\boldsymbol{\omega} \times \mathbf{r}) \right) + \left(\frac{(1-\mu)}{r_1^3}\mathbf{r_1} + \frac{\mu}{r_2^3}\mathbf{r_2} \right) \tag{2.7} $$

where μ represents the mass fraction of the smaller body, or $m_2/(m_1 + m_2)$, and r_1 and r_2 are the distances from the larger and smaller bodies, respectively, that is,

$$ r_1 = \sqrt{(\mu + x)^2 + y^2 + z^2} $$
$$ r_2 = \sqrt{(\mu + x - 1)^2 + y^2 + z^2} $$

[1] Derivatives of the position vector (${}^{R}\mathbf{v}$ and ${}^{R}\mathbf{a}$) are assumed to be relative to the rotating frame and, consequently, R is dropped.

Motion in the vicinity of the Moon is affected by both lunar and terrestrial gravity. To a far lesser degree, motion is also influenced by the Sun and other solar system bodies. These other gravitational influences are not generally incorporated in the present formulation. The Earth's gravitational effect is on the same order of magnitude as lunar gravity in the proximity of the Moon. Additionally, modeling the vehicle's motion in the CR3B system supplies a solution in a frame in which the two primary bodies are fixed. This feature greatly simplifies the geometric interpretation of the solution. Furthermore, the simplifications resulting from the assumptions inherent in the CR3B facilitate the generation of periodic trajectories.

The preceding discussion centers on "natural" motion that is derived from a pseudo-potential. An applied acceleration term appears on the right side of Eq. (2.6). At a distance of 1 AU, an appropriate distance to assume for a sailcraft in the Earth–Moon system, the nondimensional applied acceleration from a solar sail is modeled as

$$\mathbf{a}_s(t) = \frac{a_c}{a^*}(\hat{\boldsymbol{\ell}}(t) \cdot \hat{\mathbf{u}})^2 \hat{\mathbf{u}} \tag{2.8}$$

where $\hat{\mathbf{u}}$ is the sail-face normal and $\hat{\boldsymbol{\ell}}(t)$ is a unit vector in the Sun-to-spacecraft direction. These vectors appear in Fig. 2.2. Observed from the rotating frame, R, the Sun moves in a clockwise direction about the fixed primaries. The sail mass, m_3, is negligible compared to the masses of the Earth and the Moon, which are m_1 and m_2,

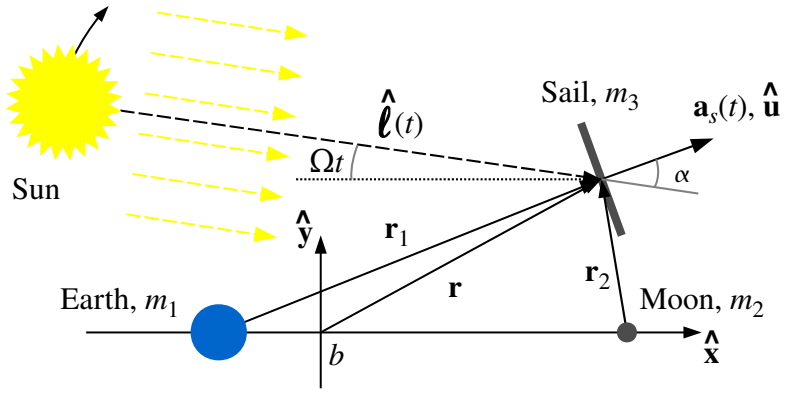


Fig. 2.2: Earth–Moon system model

respectively. The origin of the system is located at the barycenter, b, of the Earth and the Moon. The term $(\hat{\boldsymbol{\ell}}(t) \cdot \hat{\mathbf{u}})$ is also expressed as $\cos\alpha$, where α is the sail pitch angle, or the angle between the solar incidence direction and the sail-face normal. In a Sun-fixed frame where the orthonormal triad is $\left\{ \begin{array}{ccc} \hat{\boldsymbol{\ell}} & (\hat{\boldsymbol{\ell}} \times \hat{\mathbf{z}}) & \hat{\mathbf{z}} \end{array} \right\}^{\mathrm{T}}$, the sail-face normal, $\hat{\mathbf{u}}$, is expressed in terms of α and δ, that is,

$$\hat{\mathbf{u}} = \left\{ \begin{array}{c} \cos\alpha \\ \sin\delta\sin\alpha \\ \cos\delta\sin\alpha \end{array} \right\} \tag{2.9}$$

where δ is also known as the clock angle and is measured about the Sun-line, $\hat{\boldsymbol{\ell}}$, from the z axis in the R frame, as illustrated in Fig. 2.3 (and described in Appendix B).

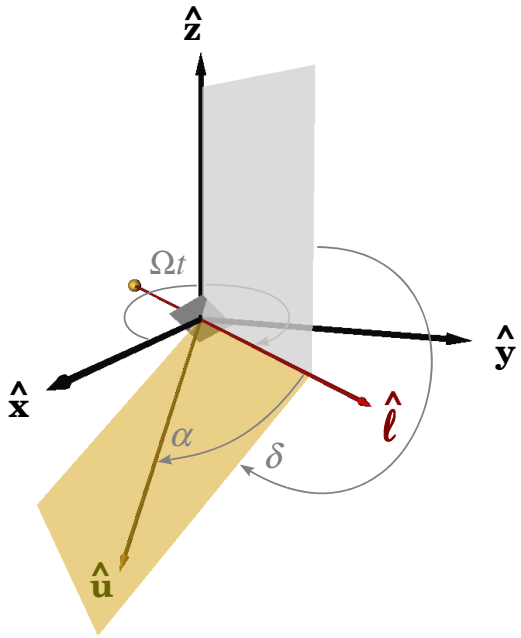


Fig. 2.3: Pitch, α, and clock, δ, angles for the sail-face normal with respect to the Sun-line. The axes are fixed in the rotating frame and the Sun moves about the Earth–Moon system at a rate of Ω.

When $\hat{\mathbf{u}}$ is rotated to the working frame, R, it is

$$\hat{\mathbf{u}} = \left\{ \begin{array}{c} \cos\Omega t \cos\alpha + \sin\Omega t \sin\alpha \sin\delta \\ -\sin\Omega t \cos\alpha + \cos\Omega t \sin\alpha \sin\delta \\ \cos\delta \sin\alpha \end{array} \right\} \tag{2.10}$$

Formulations of the sail-face normal in other reference frames are available and are introduced as necessary.

As demonstrated in Eq. (2.8), the direction of the sunlight vector with respect to the sail-face normal is a critical and mitigating factor in the acceleration supplied by the sail. The sunlight direction is expressed relative to the rotating frame and is a function of time, that is,

$$\hat{\boldsymbol{\ell}}(t) = \cos(\Omega t)\,\hat{\mathbf{x}} - \sin(\Omega t)\,\hat{\mathbf{y}} + 0\,\hat{\mathbf{z}} \tag{2.11}$$

where Ω is the ratio of the synodic rate of the Sun as it moves along its path to the system rate, approximately 0.9192. One physical constraint is imposed on the attitude of the spacecraft: the sail normal, $\hat{\mathbf{u}}$, which is coincident with the direction of the resultant force in an ideal model, is always directed away from the Sun. This is written mathematically as

$$\hat{\boldsymbol{\ell}}(t) \cdot \hat{\mathbf{u}} \geq \cos\alpha_{\max} \tag{2.12}$$

where $\alpha_{\max}$ is 90°. At $\alpha = 90°$, the sail is edge-on to the Sun. If $\alpha > 90°$, the reflecting side is directed away from the Sun and the "back" side now faces the Sun. An incident force is still generated, but the back side of a solar sail is typically not designed for high reflectivity, rather it is coated to emit excess heat (countering the force from reflected sunlight). In reality, because of imperfections in the sail, $\alpha_{\max}$ is typically less than 90°. The differences between an ideal model and other, more realistic, models are discussed in Appendix A.

2.2 Symmetry in solar sail trajectories

Construction of some solar sail trajectories is aided by observations that are fundamental to the dynamical model. It is well known that certain symmetries exist in

the circular restricted three-body problem (CR3BP). One such symmetry arises from the time-invariance, or reversibility, of the equations of motion. The scalar equations of motion for the CR3BP are

$$\ddot{x} - 2\dot{y} - x + \frac{(1-\mu)(x+\mu)}{r_1^3} + \frac{\mu(x-1+\mu)}{r_2^3} = 0 \tag{2.13}$$

$$\ddot{y} + 2\dot{x} - y + \frac{(1-\mu)y}{r_1^3} + \frac{\mu y}{r_2^3} = 0 \tag{2.14}$$

$$\ddot{z} + \frac{(1-\mu)z}{r_1^3} + \frac{\mu z}{r_2^3} = 0 \tag{2.15}$$

If trajectories are reversed across the xz plane, then

$$x \to x, \quad y \to -y, \quad z \to z, \quad t \to -t$$

The subsequent path derivatives are

$$\frac{d(x)}{d(-t)} = -\dot{x}, \qquad \frac{d(-y)}{d(-t)} = \dot{y}, \qquad \frac{d(z)}{d(-t)} = -\dot{z}$$
$$\frac{d(-\dot{x})}{d(-t)} = \ddot{x}, \qquad \frac{d(\dot{y})}{d(-t)} = -\ddot{y}, \qquad \frac{d(-\dot{z})}{d(-t)} = \ddot{z}$$

If these states are substituted into Eqs. (2.13) through (2.15), the same equations of motion as those appearing in Eqs. (2.13) through (2.15) are recovered. The advantage of this phenomenon is that a *mirror configuration* occurs at the crossing of the xz plane or the y axis [86]. Two types of mirror configurations are possible: (P) when all point masses (i.e., the primaries and the spacecraft) lie in a plane and all velocity vectors are at right angles to that plane and (A) when all point masses lie along an axis and all velocity vectors are at right angles to that axis, but not necessarily parallel. Orbits are periodic if, at two separate epochs, a mirror configuration occurs [86]. An example of a periodic orbit with two mirror configurations in an arbitrary CR3B system appears in Fig. 2.4. Because of the mirror configurations, certain states are assumed to be zero at the perpendicular crossings of the symmetry plane (or axis). One implication from these assumptions is that a trajectory only requires numerical integration from one mirror configuration to the subsequent mirror configuration, or one half-period, to determine a symmetric periodic orbit.

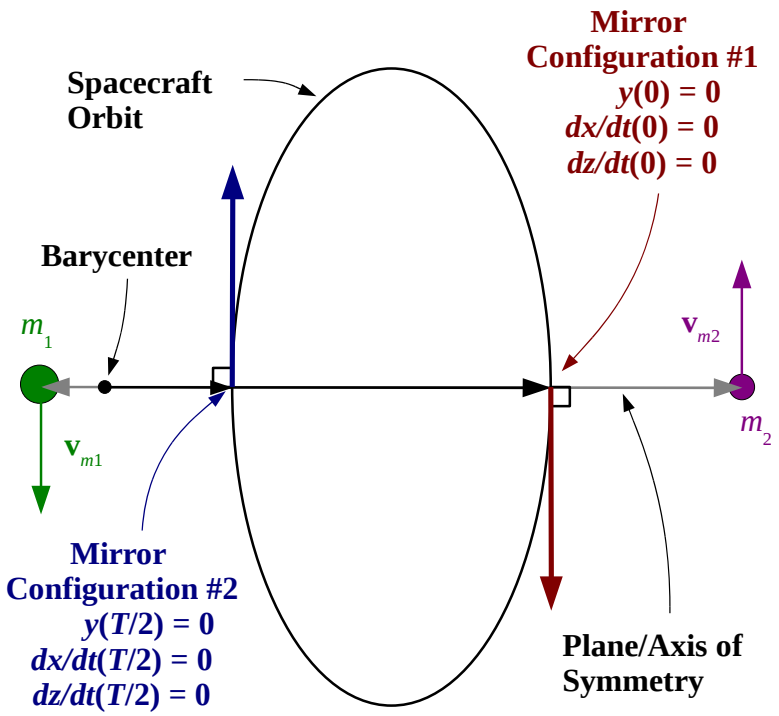


Fig. 2.4: An example of a mirror configuration (either P-type or A-type).

For a solar sail moving in the Earth-Moon CR3B system, a symmetry similar to a mirror configuration occurs for time reversal in the equations of motion. When expressed in terms of the Earth–Moon rotating frame, the equations of motion for a solar sail spacecraft are

$$\ddot{x} - 2\dot{y} - x + \frac{(1-\mu)(x+\mu)}{r_1^3} + \frac{\mu(x-1+\mu)}{r_2^3} = \beta(\mathrm{c}_{\Omega t}\mathrm{c}_\alpha^3 + \mathrm{s}_{\Omega t}\mathrm{s}_\alpha\mathrm{c}_\alpha^2\mathrm{s}_\delta) \tag{2.16}$$

$$\ddot{y} + 2\dot{x} - y + \frac{(1-\mu)y}{r_1^3} + \frac{\mu y}{r_2^3} = \beta(-\mathrm{s}_{\Omega t}\mathrm{c}_\alpha^3 + \mathrm{c}_{\Omega t}\mathrm{s}_\alpha\mathrm{c}_\alpha^2\mathrm{s}_\delta) \tag{2.17}$$

$$\ddot{z} + \frac{(1-\mu)z}{r_1^3} + \frac{\mu z}{r_2^3} = \beta(\mathrm{s}_\alpha\mathrm{c}_\alpha^2\mathrm{c}_\delta) \tag{2.18}$$

where c $\equiv$ cos, s $\equiv$ sin, and $\beta = a_c/a^*$. When these equations are propagated in negative time, the following equations result

$$\ddot{x} - 2\dot{y} - x + \frac{(1-\mu)(x+\mu)}{r_1^3} + \frac{\mu(x-1+\mu)}{r_2^3} = \beta(\mathrm{c}_{\Omega t}\mathrm{c}_\alpha^3 - \mathrm{s}_{\Omega t}\mathrm{s}_\alpha\mathrm{c}_\alpha^2\mathrm{s}_\delta) \tag{2.19}$$

$$\ddot{y} + 2\dot{x} - y + \frac{(1-\mu)y}{r_1^3} + \frac{\mu y}{r_2^3} = \beta(-\mathrm{s}_{\Omega t}\mathrm{c}_\alpha^3 - \mathrm{c}_{\Omega t}\mathrm{s}_\alpha\mathrm{c}_\alpha^2\mathrm{s}_\delta) \tag{2.20}$$

$$\ddot{z} + \frac{(1-\mu)z}{r_1^3} + \frac{\mu z}{r_2^3} = \beta(\mathrm{s}_\alpha\mathrm{c}_\alpha^2\mathrm{c}_\delta) \tag{2.21}$$

Note the sign differences on the right side when comparing Eqs. (2.16) and (2.17) to Eqs. (2.19) and (2.20), respectively. Time invariance is only possible for solar sails in the Earth–Moon system if the clock angle $\delta = 0°$ or $180°$ at the time of the mirror configuration. This property is also observed and employed by Farrés and Jorba [87]. While Roy and Ovenden [86] specify that the "n-point masses are acted upon by their mutual gravitational forces only," the Mirror Theorem still holds when incorporating a solar sail in the Earth–Moon system because the Sun is assumed to be in a plane containing the Earth, Moon, and sailcraft at the times of the sailcraft crossing of the symmetry plane. Although solar gravity is not modeled, solar radiation pressure acts in a manner similar to gravity, albeit repulsive, in this system.

2.3 Linear dynamics

One convenient and straightforward approach for the design and control of trajectories in the CR3B system is the use of linear variational equations relative to some reference arc. Consider a system of equations such that

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t)) \tag{2.22}$$

where $\mathbf{x}(t)$ is a vector containing the position and velocity states, or

$$\mathbf{x}(t) = \left\{ \begin{array}{cc} \mathbf{r}(t)^{\mathrm{T}} & \mathbf{v}(t)^{\mathrm{T}} \end{array} \right\}^{\mathrm{T}} \tag{2.23}$$

at an arbitrary time t. Now let $\mathbf{x}^r(t)$ be a reference trajectory that satisfies the nonlinear equations of motion under the conditions in which they are formulated. A

nearby path also satisfies the equations of motion and is denoted $\mathbf{x}^n(t)$. This path is generally unknown a priori. It may be considered a desired path or a path that is employed simply to track the reference trajectory, depending on the application. At time t, the difference between the nearby and reference trajectories is

$$\delta\mathbf{x}(t) = \mathbf{x}^n(t) - \mathbf{x}^r(t) \tag{2.24}$$

Equation (2.24) can be evaluated at some initial time, t_0, such that $\mathbf{x}^n(t_0)$ is perturbed from $\mathbf{x}^r(t_0)$ by $\delta\mathbf{x}(t_0)$, or $\mathbf{x}_0^n$, $\mathbf{x}_0^r$, and $\delta\mathbf{x}_0$, respectively, for convenience. If Eq. (2.22) is formulated as an initial value problem,

$$\mathbf{x}^n(t) = \mathbf{x}(t, \mathbf{x}_0^r + \delta\mathbf{x}_0) \tag{2.25}$$

$$\mathbf{x}^r(t) = \mathbf{x}(t, \mathbf{x}_0^r) \tag{2.26}$$

An illustration of these neighboring paths, along with $\delta\mathbf{x}_0$, appears in Fig. 2.5.

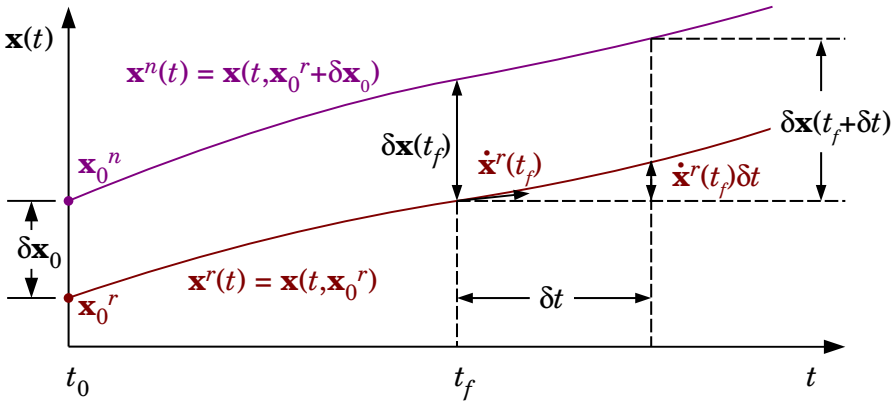


Fig. 2.5: Variations between neighboring trajectories [88].

The small step $\delta\mathbf{x}(t)$ relative to $\mathbf{x}^r(t)$ is approximated straightforwardly via a set of linear differential equations. Differentiating Eq. (2.24) with respect to an inertial observer results in

$$\delta\dot{\mathbf{x}}(t) = \dot{\mathbf{x}}^n(t) - \dot{\mathbf{x}}^r(t) \tag{2.27}$$

Continuing with the assumption that $\mathbf{x}^r(t)$ and $\mathbf{x}^n(t)$ are close (and, therefore, their derivatives are also close), a Taylor series expansion about $\mathbf{x}^r(t)$ yields a first-order approximation for $\delta\dot{\mathbf{x}}(t)$,

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}(t)\delta\mathbf{x}(t) \tag{2.28}$$

where $\mathbf{A}(t)$ is a 6×6 Jacobian matrix

$$\mathbf{A}(t) = \left.\frac{\partial\dot{\mathbf{x}}(t)}{\partial\mathbf{x}(t)}\right|_{x^r} = \left.\frac{\partial\mathbf{f}(\mathbf{x}(t))}{\partial\mathbf{x}(t)}\right|_{x^r} \tag{2.29}$$

For the autonomous CR3B system without an applied acceleration (such as the acceleration supplied by a solar sail),

$$\mathbf{f}(\mathbf{x}(t)) = \left\{ \begin{array}{c} \mathbf{v}(t) \\ -2\left(\boldsymbol{\omega} \times \mathbf{v}(t)\right) - \boldsymbol{\nabla} U(\mathbf{r}(t)) \end{array} \right\} \tag{2.30}$$

For this system, the elements of $\mathbf{A}(t)$ are time-varying and of the form

$$\mathbf{A}(t) = \left[\begin{array}{cc} \mathbf{0} & \mathbf{I} \\ -\mathbf{M} & -\mathbf{S} \end{array} \right] \tag{2.31}$$

where $\mathbf{M}$ is the set of partial derivatives of $\boldsymbol{\nabla} U$ from Eq. (2.7) with respect to $\mathbf{r}$, $\mathbf{S}$ is the set of partial derivatives of the Coriolis terms in Eq. (2.6) with respect to $\mathbf{v}$, and $\mathbf{I}$ is the 3×3 identity matrix.

The solution to Eq. (2.28) is significant to serve as the basis of the design and control algorithms. If $\delta\mathbf{x}(t)$ is available at some future time, t_f, it can be added to the known reference at t_f, that is, $\mathbf{x}^r(t_f)$, to approximate the state along the nearby trajectory, at the specified time, $\mathbf{x}^n(t_f)$. This variation appears as $\delta\mathbf{x}(t_f)$ in Fig. 2.5 and is denoted *contemporaneous* because the two trajectories are compared at the same time. To solve Eq. (2.28), the variables are separated and both sides are integrated from t_0 to t_f, resulting in

$$\delta\mathbf{x}(t_f) = \boldsymbol{\Phi}(t_f, t_0)\delta\mathbf{x}(t_0) \tag{2.32}$$

where $\boldsymbol{\Phi}(t_f, t_0)$ is the state transition matrix (STM). The STM supplies a linear mapping of the perturbation $\delta\mathbf{x}(t_0)$ at t_0 to the deviation from the reference path,

$\delta\mathbf{x}(t_f)$, at a later time. These sensitivities are essentially linear and are expressed in the form of partial derivatives; in matrix form, they are represented

$$\mathbf{\Phi}(t_f, t_0) = \frac{\partial \mathbf{x}(t_f)}{\partial \mathbf{x}(t_0)} \tag{2.33}$$

Unfortunately, calculating these partial derivatives is not always straightforward.

Generating the STM can be accomplished in various ways depending on the characteristics of the system of equations and the computational capabilities. In the first approach, the linear system of equations comprising $\mathbf{A}(t)$ is assumed to be constant over an interval from t_i to t_{i+1}. In this case, the STM is represented as

$$\mathbf{\Phi}(t_f, t_0) = e^{\mathbf{A}\Delta t} \tag{2.34}$$

where $\Delta t = t_{i+1} - t_i$. (Of course, if $\mathbf{A}$ is invariant, this solution applies for all time.) A second approach involves a Taylor series expansion of the exponential function in Eq. (2.34), resulting in

$$e^{\mathbf{A}\Delta t} = \sum_{k=0}^{\infty} \frac{1}{k!}(\Delta t \mathbf{A})^k \tag{2.35a}$$

$$= \mathbf{I} + \Delta t\mathbf{A} + \frac{1}{2!}(\Delta t\mathbf{A})^2 + \frac{1}{3!}(\Delta t\mathbf{A})^3 + \ldots \tag{2.35b}$$

where $\Delta t = t_f - t_0$. In practice, higher-order approximations are no typically employed.

Often, the Jacobian is time-varying. An alternative when $\mathbf{A}(t)$ is a continuous function of time is to integrate the STM. With a slight change of notation, the derivative of the STM in Eq. (2.33) is

$$\dot{\mathbf{\Phi}}(t, t_0) = \frac{\partial \dot{\mathbf{x}}(t)}{\partial \mathbf{x}(t)} \frac{\partial \mathbf{x}(t)}{\partial \mathbf{x}(t_0)} \tag{2.36}$$

$$= \mathbf{A}(t)\mathbf{\Phi}(t, t_0) \tag{2.37}$$

The same result appears when Eq. (2.32) is substituted into Eq. (2.28). Equation (2.37) comprises a set of 36 scalar differential equations. To determine the elements in $\mathbf{\Phi}(t, t_0)$, the STM is integrated along with the equations of motion for

the system for a set of 42 scalar differential equations or, assuming they are available, directly uses the states from the reference path to obtain $\mathbf{A}(t)$.

The preceding strategy exploits the linear variational equations. Because the STM is a matrix of the sensitivities to a perturbation in each element of $\mathbf{x}(t_0)$, another approach for generating the STM is direct numerical integration of the nonlinear equations of motion, once for each element of the state vector. For each integration, the corresponding state is perturbed by some small value, ϵ, at the initial epoch. The state at the end of the integration forms the column in $\mathbf{\Phi}(t_f, t_0)$ associated with the particular state variable. No partial derivatives are required; however, this approach requires additional computational processing and requires more time to generate the STM compared to the numerical integration of Eq. (2.37). Also, selection of an appropriate value of ϵ across various types of states can be nontrivial.

Three useful properties of the STM are noteworthy. The first arises from the fact that if $\delta\mathbf{x}(t_1) = \mathbf{\Phi}(t_1, t_0)\delta\mathbf{x}(t_0)$ and $\delta\mathbf{x}(t_2) = \mathbf{\Phi}(t_2, t_1)\delta\mathbf{x}(t_1)$, then

$$\mathbf{\Phi}(t_n, t_0) = \mathbf{\Phi}(t_n, t_{n-1}) \dots \mathbf{\Phi}(t_2, t_1)\mathbf{\Phi}(t_1, t_0) \tag{2.38}$$

where n is a finite number of intervals. If the interval from t_i to t_{i+1} is small, $\mathbf{A}$ may be assumed to be time-invariant over that interval. Then, the STMs generated in either Eq. (2.34) or Eq. (2.35) that correspond to each small interval may be stacked, per Eq. (2.38), to produce a STM from t_0 to t_n. Another useful property of the STM is that

$$\mathbf{\Phi}(t_0, t_f) = \mathbf{\Phi}^{-1}(t_f, t_0) \tag{2.39}$$

If the STM is generated in forward time, its inverse is the associated mapping for the deviation at t_f, $\delta\mathbf{x}(t_f)$, to the perturbation at t_0, $\delta\mathbf{x}(t_0)$. Finally, from the definition in Eq. (2.32), when $t = t_0$,

$$\mathbf{\Phi}(t_0, t_0) = \mathbf{I} \tag{2.40}$$

where $\mathbf{I}$ is the 6×6 identity matrix. This identity property indicates that a perturbation at t_0 is unchanged when it is mapped to itself.

The preceding discussion is based on a contemporaneous variation. Note that a time variation between the reference and neighboring trajectories is also possible. Given that all perturbations and deviations are assumed to be small, the difference between two states on the neighboring and reference trajectories is expressed as the sum of the deviation at t_f between the two paths and the derivative of the reference path at t_f multiplied by a noncontemporaneous variation, δt, or

$$\delta\mathbf{x}(t_f + \delta t) = \delta\mathbf{x}(t_f) + \dot{\mathbf{x}}^r(t_f)\delta t \tag{2.41}$$

These variations appear in Fig. 2.5. Substituting Eq. (2.32) into Eq. (2.41) results in the full form of the first-order variational equations of motion,

$$\delta\mathbf{x}(t_f + \delta t) = \mathbf{\Phi}(t_f, t_0)\delta\mathbf{x}(t_0) + \dot{\mathbf{x}}^r(t_f)\delta t \tag{2.42}$$

This equation also emerges from a Taylor expansion. Note that the noncontemporaneous variations are mapped via the derivative of the position and velocity state elements (i.e., velocity and acceleration, respectively) at the final time, t_f, along $\mathbf{x}^r(t)$. Equation (2.42) is useful for correcting $\mathbf{x}^r(t_0)$ (discussed in further detail in Chapter 3) to solve for $\delta\mathbf{x}(t_0)$, assuming all other quantities are available. These corrections are classified as *final-time free* problems. Often, the final time is specified and a noncontemporaneous variation is not desired; then, $\delta t = 0$ and Eq. (2.42) reduces to Eq. (2.32). The solar sail trajectories in this investigation are generally *final-time fixed* problems, unless otherwise noted.

If the problem includes a set of controls and other parameters that may be adjusted, the variational equations are augmented. The system of differential equations is then formulated such that a dependency on the controls and a set of parameters also exists, that is,

$$\dot{\mathbf{x}}(t) = \mathbf{f}(\mathbf{x}(t), t) + \mathbf{g}(\mathbf{x}(t), \mathbf{y}(t), t) \tag{2.43}$$

where $\mathbf{y}$ represents a generic set of controls and parameters and $\mathbf{g}(\mathbf{x}(t), \mathbf{y}(t), t)$ is a function of the state, controls, and parameters, as well as time, t. For the nonautonomous system defined by Eqs. (2.6) and (2.8), $\mathbf{g}(\mathbf{y}(t), t) = \mathbf{a}_s(t)$. Note, however,

for this system, that $\mathbf{f}(\mathbf{x}(t))$ from Eq. (2.30) is autonomous and time invariant, while $\mathbf{g}(\mathbf{y}(t), t)$ is non-autonomous and time variant, but not a function of $\mathbf{x}(t)$. The linear vector equation of motion, i.e., the first-order variational equations, is written in the familiar state-space form

$$\delta\dot{\mathbf{x}}(t) = \mathbf{A}(t)\delta\mathbf{x}(t) + \mathbf{B}(t)\delta\mathbf{y}(t) \tag{2.44}$$

where

$$\mathbf{B} = \left.\frac{\partial(\mathbf{f}+\mathbf{g})}{\partial\mathbf{y}}\right|_{\mathbf{y}^r} \tag{2.45}$$

and $\mathbf{y}^r = \mathbf{y}^r(t)$, which refers to a reference control profile as well as a parameter set that appear in the equations of motion. The matrix form of Eq. (2.44) is

$$\left\{ \begin{array}{c} \delta\dot{\mathbf{x}}(t) \\ \delta\dot{\mathbf{y}}(t) \end{array} \right\} = \left[\begin{array}{cc} \mathbf{A}(t) & \mathbf{B}(t) \\ \mathbf{A}_y(t) & \mathbf{B}_y(t) \end{array} \right] \left\{ \begin{array}{c} \delta\mathbf{x}(t) \\ \delta\mathbf{y}(t) \end{array} \right\} \tag{2.46}$$

where

$$\mathbf{A}_y(t) = \left.\frac{\partial\dot{\mathbf{y}}}{\partial\mathbf{x}}\right|_{\mathbf{x}^r} \text{ and } \mathbf{B}_y(t) = \left.\frac{\partial\dot{\mathbf{y}}}{\partial\mathbf{y}}\right|_{\mathbf{y}^r} \tag{2.47}$$

The matrices $\mathbf{A}_y(t)$ and $\mathbf{B}_y(t)$ are typically zero if $\delta\dot{\mathbf{y}} = \mathbf{0}$ or if $\mathbf{y}$ is not a function of time.

Similar to the uncontrolled problem discussed previously, generating variational relationships for the set of controls and parameters is not necessarily straightforward. Consistent with the procedure to generate Eq. (2.37), the matrix in Eq. (2.46) is incorporated into the differential equations to develop a state transition matrix, $\mathbf{\Phi}$, and a state-control (or state-parameter) transition matrix, $\mathbf{\Psi}$,

$$\frac{d}{dt}\left[\begin{array}{cc} \mathbf{\Phi} & \mathbf{\Psi} \\ \mathbf{\Phi}_y & \mathbf{\Psi}_y \end{array} \right] = \left[\begin{array}{cc} \mathbf{A}(t) & \mathbf{B}(t) \\ \mathbf{A}_y(t) & \mathbf{B}_y(t) \end{array} \right] \left[\begin{array}{cc} \mathbf{\Phi} & \mathbf{\Psi} \\ \mathbf{\Phi}_y & \mathbf{\Psi}_y \end{array} \right] \tag{2.48}$$

Initially, the matrix on the far right of Eq. (2.48) is an identity matrix of appropriate size. The matrices that map deviations in the controls and the parameters at t_0 to deviations in the state and control elements at t are included for completeness; if $\mathbf{A}_y$ and $\mathbf{B}_y$ are zero, then $\mathbf{\Phi}_y = \mathbf{0}$ and $\mathbf{\Psi}_y = \mathbf{I}$, respectively. The state-control transition

matrix, $\mathbf{\Psi}$, maps the deviation from the reference control and parameter set at t_0, or $\delta\mathbf{y}(t_0)$, to a deviation in position and velocity at some future time, $\delta\mathbf{x}(t_f)$, that is,

$$\mathbf{\Psi}(t_f, t_0) = \frac{\partial \mathbf{x}_f(t)}{\partial \mathbf{y}(t_0)} \tag{2.49}$$

The overall mapping of the errors in the state, controls (or parameters), and final time is

$$\left\{ \begin{array}{c} \delta\mathbf{x}(t_f + \delta t) \\ \delta\mathbf{y}(t_f + \delta t) \end{array} \right\} = \left[\begin{array}{cc} \mathbf{\Phi}(t_f, t_0) & \mathbf{\Psi}(t_f, t_0) \\ \mathbf{\Phi}_y(t_f, t_0) & \mathbf{\Psi}_y(t_f, t_0) \end{array} \right] \left\{ \begin{array}{c} \delta\mathbf{x}(t_0) \\ \delta\mathbf{y}(t_0) \end{array} \right\} + \left\{ \begin{array}{c} \dot{\mathbf{x}}(t_f, \mathbf{x}(t_0), \mathbf{y}(t_0)) \\ \dot{\mathbf{y}}(t_f, \mathbf{x}(t_0), \mathbf{y}(t_0)) \end{array} \right\} \delta t \tag{2.50}$$

when noncontemporaneous variations are included (i.e., $\delta t \neq 0$). For a given controller, variations in the state at time t_0 may not be allowed and $\delta\mathbf{x}(t_0) = 0$. Additionally, if $\delta t = 0$, then

$$\delta\mathbf{x}(t) = \mathbf{\Psi}(t, t_0)\delta\mathbf{y}(t_0) \tag{2.51}$$

which is an important relationship for developing trajectory targeting schemes. Both $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are useful for developing targeters for shooting schemes.

3. NUMERICAL TOOLS FOR SOLVING TWO-POINT BOUNDARY VALUE PROBLEMS

Trajectory design for any spacecraft mission application typically involves either well-developed analytical approximations or linearization relative to a known solution. These approximations are typically based on well-understood dynamics. However, when two or more large bodies (e.g., the Earth and the Moon, or the Sun, the Earth, and the Moon) are present, trajectories in a multi-body gravitational environment can become chaotic, and the formulation of solutions, whether analytical or numerical, in these multi-body systems possesses new challenges.

Ion propulsion and solar sails supply an additional force on the spacecraft, one that can be directed, often expanding the design options. However, envisioning the necessary control history and the resulting trajectory prior to any analysis can be difficult, especially when the advanced-propulsion device is employed in a complex dynamical regime, such as a CR3B system. Analytical techniques that incorporate these control forces in such a regime are not sufficiently developed to reveal the extent of the design possibilities quickly.

Numerical methods are critical elements in uncovering new trajectory options and furnishing mission designers with a better sense of the design space, particularly in complex regimes. Spacecraft trajectories are, fundamentally, solutions to nonlinear equations of motion (EOMs), which themselves are ordinary differential equations (ODEs). Some ODEs are posed as initial value problems (IVPs) and are solved numerically using explicit or implicit integration, incorporating single or multiple steps. In an IVP, the state is defined at some point in the time domain, and all subsequent states along the path evolve as a result of the differential equations and the given initial condition. For explicit integration, the state at the current time is

calculated based on the state at the previous time (or, in a multi-step method, at a number of preceding times); it is the method of integration most analysts use regularly. Implicit integration may also be single or multi-step and also relies on the initial conditions and the differential equations, but the state at the current time is expressed implicitly. An iterative method, such as a Newton-Raphson algorithm, is required to solve the differential equations at the current point in the time domain. Obtaining solutions using implicit methods is generally more difficult, but the accuracy of the solution is improved for stiff ODEs [89, 90].

Trajectories can also be represented as solutions to boundary value problems (BVPs) in terms of ordinary differential equations (ODEs). For a BVP, conditions are specified at the beginning and, also, at the end of the time domain. The specific values of the states at the extremes may or may not be fixed in a BVP with a periodicity constraint; periodicity only requires that the values at the extremes are equal. Differential-corrections schemes, including shooting methods [91], collocation approaches, and finite-difference methods [89,92–95] are common numerical processes to solve BVPs and are similar in that they all employ linear corrections to a nonlinear system. Variational methods are also used to solve BVPs when the problem possesses certain minimality properties [90]. Classical perturbation techniques have long supplied analytical approximations for trajectories [63, 96–98]. Finally, recent efforts using evolutionary algorithms to solve optimization problems also offer an approach to solving boundary value problems [99, 100].

Any differential-corrections scheme can be formulated simply as an initial value problem with a constrained initial position and a free initial velocity. Using a shooting method, for example, the difference between the end state and the target state at the terminal point on the trajectory arc is used to adjust the initial velocity [89, 91, 101, 102]. Shooting is often used in the optimization of spacecraft trajectories [103–105]. Trajectory arcs involving low-thrust ion engines are typically discretized such that the thrust arcs resemble a series of finite burns. Unfortunately, continuous control,

such as the control supplied by a solar sail, cannot be easily incorporated into typical shooting algorithms.

Collocation techniques have arisen in the past few decades as an alternative strategy to generate spacecraft trajectories, coincident with improvements in computer processor capabilities [30, 106–109]. Collocation routines are also common tools for solving trajectory optimization problems [104, 105]. The advantage of a collocation scheme in solving a BVP is a formulation such that the control can be discretized; the algorithm then delivers a solution for the control and trajectory states simultaneously [110]. Collocation involves minimization of the difference between the derivative of a continuous polynomial approximation along an arc of the trajectory and the derivative from the equations of motion at an intermediate, or *defect*, point between nodes along the path. Mathematically, collocation methods require the selection of a representative solution (typically a polynomial) that is consistent with the vector field corresponding to the differential equations [111]. In the 1970's, collocation methods were applied to boundary value problems by Russell and Shampine [110]. Doedel also approached the solution to nonlinear two-point boundary problems using finite differences in a collocation strategy [112].

In the mid-1980's, Hargraves and Paris applied a collocation method to a direct trajectory optimization problem [106]. A functional (indirect) optimization problem is transcribed to a parameter (direct) optimization problem by using discretized interior states as optimization parameters. Nonlinear programming techniques are then employed to solve for the set of interior states that optimizes some cost function, usually propellant or time-of-flight [106–108]. Herman and Conway improved the accuracy of the technique by incorporating high-degree Gauss-Lobatto quadrature rules into a collocation and nonlinear-programming algorithm [109]. Tang and Conway applied these techniques to low-thrust interplanetary trajectories [113], and Melton successfully adapted them to the interplanetary solar sail problem [114]. Other authors have also explored trajectory optimization as a parameter optimization problem [104,115]. The focus of the preceding studies is direct optimization methods, which are gener-

ally less accurate than indirect methods [103], but can be used as starting points for indirect optimization schemes.

Most recently, Ozimek, Grebow, and Howell applied low- and high-degree Gauss-Lobatto rules to generate solar sail trajectories in the Earth–Moon system [30]. The local error along the trajectory path is proportional to Δt^5 and Δt^{12}, respectively, due to truncation in the approximating polynomials [56]. The problem is not formulated as an optimization problem. Rather, the authors employ this approach to iteratively solve the equations of motion. The converged solution is not expected to resemble the initial guess and rarely does so; however, the solution does depend on the initial guess.

The preceding numerical methods possess high degrees of accuracy. Any type of shooting strategy requires some knowledge of the design space to initiate the procedure. Collocation schemes do not require a priori knowledge of the solution corresponding to the BVP, but the implementation is not intuitive. Furthermore, the partial derivatives of the constraints with respect to the states are nontrivial in collocation schemes and are often approximated numerically [30].

A more rudimentary approach to solve BVPs is a finite-difference method (FDM), in which the derivatives that appear in the differential equations are replaced with their respective finite differences and evaluated at node points along the trajectory [34, 89]. The solution process is iterative. The trajectory is discretized, and the equations that represent the relationships at the nodes are solved simultaneously. Using a central-difference approximation, the best accuracy available from an FDM is proportional to Δt^2. However, model fidelity and uncertainty may obviate any accuracy gains from a highly accurate scheme; solutions from such methods may not persevere without additional control in actual flight. If a solution based on a high-precision approach (e.g., shooting or high-degree collocation) is migrated from a model with limited accuracy to one of higher accuracy (e.g., CR3B model vs. ephemeris model or actual flight), it may be no more reliable in that higher-fidelity model than a solution from a low-precision approach (e.g., finite-differencing). More-

over, the FDM is simple to formulate and offers an improved understanding of the design space. Partial derivatives are easily obtained via standard analytical techniques, without resorting to numerical or automatic differentiation (both of which are often computationally expensive). Once the analyst is familiar with the feasible options, higher-fidelity methods for solving the BVP can be used to refine trajectories or combined with optimization techniques to produce optimal trajectories. Another option is to employ the FDM technique to generate trajectories that fit path constraints to explore a broad design space quickly. Using the MATLAB® computing environment, Wawrzyniak and Howell [35, 116] survey over 10 million initial guess combinations for the path and control profile, as well as a range of characteristic accelerations and path constraints appropriate to the lunar south pole coverage problem. The time to generate the millions of trajectories for the survey is approximately one week when the algorithms run on eight cores over five platforms.

3.1 Shooting methods

Shooting methods are common tools for generating trajectories in astrodynamics. In general, some sort of approximation is developed, perhaps via analytical techniques that incorporate knowledge of the system dynamics. Often, a trajectory is approximated by employing linearization relative to a known reference solution to the nonlinear differential equations, such as an equilibrium point. States from this approximation are used to initialize the numerical integration of the nonlinear equations of motion. A state transition matrix (STM) is propagated along with a six-state (position and velocity) trajectory and is used to linearly (or differentially) correct the states and, thus, select the initial values of the elements. Dynamical characteristics extracted from the STM are effective at describing the consequences of small perturbations and, as a result, differential corrections implemented with a shooting scheme is a popular tool for generating, or updating, numerical representations of orbits. Shooting methods are often described in terms of two types: (1) single shooting and

(2) multiple shooting. In single shooting, a single arc is defined from an initial point to a final target. Multiple shooting [90] exploits short arc segments, usually along non-periodic orbits, for a more stable numerical process; path constraints (e.g., elevation) and continuity between arc segments are enforced at the endpoints along each segment. The sail attitude is generally fixed along each arc segment (or the entire orbit if single shooting is employed).

Nuss [117] and A. McInnes [118] both use single-shooting methods to generate halo orbits in the SE-CR3B regime. Nuss begins with a third-order approximation for a halo orbit in the vicinity of a collinear Lagrange point (L_1). Nuss then uses shooting to deliver a nonlinear solution and then applies continuation based on a_c to extend the nonlinear, solar sail halo orbits toward the Sun [117]. A. McInnes expands the approximation to include effects from a solar sail and generates families near artificial Lagrange points for various a_c [118]. Waters and C. McInnes [119] exploit a single-shooting corrections process to correct Lindstedt-Poincaré approximations of periodic orbits above the SE ecliptic plane. Likewise, Farrés and Jorba [87] employ multiple shooting as a tool to compute periodic orbits near the Sun–Earth–Sail L_1 point. In each of the preceding examples, sail attitude is assumed to be fixed relative to the modeled frame.

3.1.1 Single shooting example: Liapunov orbits

A periodic orbit about the L_1 libration point is generated per Szebehely [120, pp. 242–249], whereby the equations of motion are linearized in the neighborhood of L_1.[1] The resulting elliptical path in the xy plane is represented in black in Fig. 3.1. If a state is extracted (the black dot in Fig. 3.1) from the linear path and propagated in the nonlinear equations of motion for the CR3B system until the next crossing of the x axis, the green path in Fig. 3.1 results. Clearly, the linear approximation does not immediately translate into a nonlinear solution.

[1]Note that the nomenclature for L_1 and L_2 is reversed in Ref. [120].

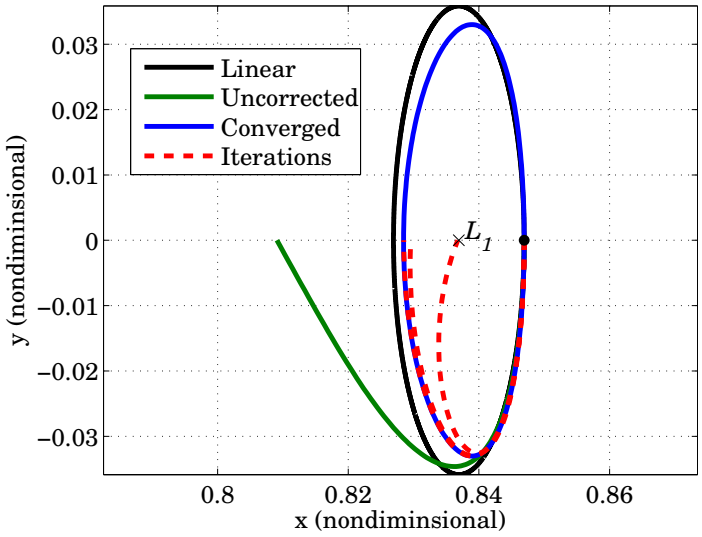


Fig. 3.1: An example of single shooting for a Liapunov orbit.

The goal is to produce a periodic trajectory. Because the Mirror Theorem states that a trajectory is periodic if, at two different epochs, a mirror configuration occurs, two perpendicular crossings of the x axis are employed for generating a periodic orbit [86]. A perpendicular crossing occurs when $\dot{x} = 0$ and $y = 0$. The first perpendicular crossing is specified as the initial state, that is, at the black dot on the right side of the ellipse in Fig. 3.1. For this two-dimensional example, y_0 and $\dot{x}_0$ are equal to zero. The initial value x_0 is fixed. The initial value $\dot{y}_0$ is allowed to vary throughout the corrections process so that the targeter ultimately determines an initial velocity that produces a perpendicular state, or a *mirror configuration*, at the next x axis crossing.

To target a perpendicular crossing, the path is propagated and the state transition matrix is generated per Eq. (2.37) until the next crossing of the x axis. The time of this crossing is denoted as t_f. The variational equations are

$$\begin{Bmatrix} \delta x_f \\ \delta y_f \\ \delta \dot{x}_f \\ \delta \dot{y}_f \end{Bmatrix} = \mathbf{\Phi} \begin{Bmatrix} \cancelto{0}{\delta x_0} \\ \cancelto{0}{\delta y_0} \\ \cancelto{0}{\delta \dot{x}_0} \\ \delta \dot{y}_0 \end{Bmatrix} + \begin{Bmatrix} \delta \dot{x}_f \\ \delta \dot{y}_f \\ \delta \ddot{x}_f \\ \delta \ddot{y}_f \end{Bmatrix} \delta t \tag{3.1}$$

Because δx_f and $\delta \dot{y}_f$ are unconstrained and only $\delta \dot{y}_0$ may be adjusted, Eq. (3.1) is simplified to

$$\left\{ \begin{array}{c} \delta y_f \\ \delta \dot{x}_f \end{array} \right\} = \left[\begin{array}{cc} \frac{\partial y_f}{\partial \dot{y}_0} & \dot{y}_f \\ \frac{\partial \dot{x}_f}{\partial \dot{y}_0} & \ddot{x}_f \end{array} \right] \left\{ \begin{array}{c} \delta \dot{y}_0 \\ \delta t \end{array} \right\} \tag{3.2}$$

where the partial derivatives in the above equation are elements from the state transition matrix, $\mathbf{\Phi}$, in Eq. (3.1). Note that, since the integration is terminated at the x axis crossing, $\delta y_f = 0$. The difference between the desired value of $\dot{x}_f$ (i.e., zero) and the actual value from the integration is $\delta \dot{x}_f = -\dot{x}_f$. With these substitutions and some algebra, the system of equations in Eq. (3.2) is rewritten as a single equation to solve for $\delta \dot{y}_0$, that is,

$$\delta \dot{y}_0 = \frac{\ddot{x}_f \left(\frac{1}{\dot{y}_f} \frac{\partial y_f}{\partial y_0} \delta \dot{y}_0 \right) - \dot{x}_f}{\frac{\partial \dot{x}_f}{\partial \dot{y}_0}} \tag{3.3}$$

A linear correction to $\delta \dot{y}_0$ results from Eq. (3.2) or (3.3) and is added to the initial guess for $\dot{y}_0$ from the reference trajectory. A new reference is propagated (the dashed red line going to L_1 in Fig. 3.1). The process is repeated until $\delta \dot{x}_f \rightarrow 0$, to within some acceptable tolerance, and the dashed red lines converge onto the blue path in Fig. 3.1. The converged value of $\dot{y}_0$ is employed to generate the blue trajectory, which is a periodic solution to the nonlinear equations of motion.

This particular example is a *final-time-free* targeter where the time of the next crossing of the x axis is allowed to vary. If the period is fixed at a particular value T_0, then δt is pre-specified as zero (for a *final-time-fixed* targeter) and, consequently, δx_0 is free to vary so that Eq. (3.1) is not over-determined.

3.1.2 Single shooting example: solar sail orbits

As an example of the computation of a solar sail trajectory using shooting in the EM problem, a linear approximation for elliptical orbits that are shifted, and remain, below the Lagrange points and parallel to the EM orbit plane is detailed in Simo and McInnes [16]. In these orbits, the pitch angle, α, is fixed to supply a constant out-of-plane force; based on the linear equations of motion, the pitch angle that supplies

the maximum out-of-plane force is 35.264°. The orbits in Simo and McInnes [16] are plotted directly from the approximation. To compute orbits that are periodic solutions to the *nonlinear* equations of motion, the Simo-McInnes approximation is used to seed a differential corrections scheme. The process is developed to determine a set of periodic orbits in the nonlinear model. The initial state, pre-specified to be located on the xz plane, from a linear approximation corresponding to an out-of-plane orbit in Simo and McInnes [16] supplies an initial guess for a nonlinear trajectory that is corrected using a differential corrections technique.

A state transition matrix (STM) is constructed based on the variations in the full state vector $\mathbf{x}$, as well as a_c, such that deviations at the beginning of the arc at t_0 are linearly mapped to some future time t_f. Higher-order state transition formulations are available [121, 122]; however, STMs based on first-order variations and linear mappings are typically used for most applications and are sufficient for this analysis. A deviation in $\mathbf{z}$, where $\mathbf{z}$ consists of $\mathbf{x}$ as well as a_c, is computed as $\delta\mathbf{z} = \mathbf{z}^d - \mathbf{z}^a$, where $\mathbf{z}^d$ is the desired value of $\mathbf{z}$ and $\mathbf{z}^a$ is the actual value.

Given the characteristics of this dynamical system, a periodic orbit is assumed to be symmetric across the xz plane,[2] thus, (1) the plane crossings are perpendicular, (2) the value for t_f can be set equal to the half-period, or the second crossing of the xz plane, and (3) a reduced set of variables is required for the differential corrections process. The objective is to reduce $\left\{ \begin{array}{ccc} \delta y_f & \delta\dot{x}_f & \delta\dot{z}_f \end{array} \right\}^{\mathrm{T}}$ to zero by adjusting either $\left\{ \begin{array}{ccc} \delta x_0 & \delta\dot{y}_0 & \delta a_c \end{array} \right\}^{\mathrm{T}}$ or $\left\{ \begin{array}{ccc} \delta z_0 & \delta\dot{y}_0 & \delta a_c \end{array} \right\}^{\mathrm{T}}$, depending on the advantages of fixing the initial values of z_0 or x_0, respectively. The other elements at t_0 are zero because t_0 occurs at the preceding perpendicular crossing. Using either of the linear mappings,

$$\left\{ \begin{array}{c} \delta y_f \\ \delta\dot{x}_f \\ \delta\dot{z}_f \end{array} \right\} = \left[\begin{array}{ccc} \frac{\partial y}{\partial x} & \frac{\partial y}{\partial \dot{y}} & \frac{\partial y}{\partial a_c} \\ \frac{\partial \dot{x}}{\partial x} & \frac{\partial \dot{x}}{\partial \dot{y}} & \frac{\partial \dot{x}}{\partial a_c} \\ \frac{\partial \dot{z}}{\partial x} & \frac{\partial \dot{z}}{\partial \dot{y}} & \frac{\partial \dot{z}}{\partial a_c} \end{array} \right] \left\{ \begin{array}{c} \delta x_0 \\ \delta\dot{y}_0 \\ \delta a_c \end{array} \right\} \tag{3.4}$$

[2]The justification for this symmetry is described in Section 2.2 and is an extension of the Mirror Theorem [86].

or

$$\left\{ \begin{array}{c} \delta y_f \\ \delta \dot{x}_f \\ \delta \dot{z}_f \end{array} \right\} = \left[\begin{array}{ccc} \frac{\partial y}{\partial z} & \frac{\partial y}{\partial \dot{y}} & \frac{\partial y}{\partial a_c} \\ \frac{\partial \dot{x}}{\partial z} & \frac{\partial \dot{x}}{\partial \dot{y}} & \frac{\partial \dot{x}}{\partial a_c} \\ \frac{\partial \dot{z}}{\partial z} & \frac{\partial \dot{z}}{\partial \dot{y}} & \frac{\partial \dot{z}}{\partial a_c} \end{array} \right] \left\{ \begin{array}{c} \delta z_0 \\ \delta \dot{y}_0 \\ \delta a_c \end{array} \right\} \tag{3.5}$$

where the variations in the matrices are from the state transition matrices in Eq. (2.33) and (2.49), the vectors on the right are used to update the reference path, generating new deviations in the final state. The new characteristic acceleration, a_c, remains constant along the orbit. The process is repeated until a convergence tolerance of 1×10^{-12} is met on the vector norm, that is, either $\left\{ \begin{array}{ccc} \delta x_0 & \delta \dot{y}_0 & \delta a_c \end{array} \right\}^{\mathrm{T}}$ or $\left\{ \begin{array}{ccc} \delta z_0 & \delta \dot{y}_0 & \delta a_c \end{array} \right\}^{\mathrm{T}}$, depending on whether Eq. (3.4) or (3.5) are employed, respectively. This particular formulation is a final-time-fixed targeter. Because the solar sail orientation and trajectory are coupled, and because the time-dependent direction of the sunlight vector and orientation determine the available acceleration, time is not free to vary when targeting future states (a *final-time-fixed* targeter). The final time is set to one half the synodic period such that the vehicle and the Sun cross the xz plane at the same time.

Once the shooting routine converges on a solution, its state at t_0, $\mathbf{x}_k(t_0)$, along with the states at t_0 of up to three preceding solutions, $\mathbf{x}_{k-1}(t_0)$ to $\mathbf{x}_{k-3}(t_0)$, is employed in a cubic polynomial that is used to extrapolate the predicted initial state along the next orbit, $\mathbf{x}^*_{k+1}(t_0)$, in the set. If the adjustment in the z components from $\mathbf{x}_k(t_0)$ to $\mathbf{x}^*_{k+1}(t_0)$ exceeds the modification in the x components, then the z component is fixed (a.k.a., continuation in z) and Eq. (3.4) is used for the iterative process. If the change in x exceeds the change in z, then the x component is fixed (continuation in x) and Eq. (3.5) is employed. Therefore, either x_0 or z_0 is used as the continuation variable. Note that α and δ are fixed at 35.264° and 0°, respectively, for each orbit, and y_0, $\dot{x}_0$, and $\dot{z}_0$ are all zero for each orbit in the set.

A set of orbits that is initialized by the linear approximation in Simo and McInnes [16] appears in Figs. 3.2 through 3.5. An "x" indicates the initial position in the xz plane of each orbit. For all of the orbits in these figures, $\dot{y}_0 < 0$. Recall that the choice

of continuation variable, x_0 or z_0, depends on the predicted values of these variables. For the orbits appearing in Fig. 3.2, the continuation variable (i.e., the variable that is fixed while other variables are adjusted to generate a periodic orbit) is initially z_0 and then switches to x_0 for the blue orbits. Note that the thick orbits in Fig. 3.2 through 3.5 coincide with bifurcations in the eigenstructure, and possibly indicate a branching to other sets of orbits. The color associated with each orbit in the figures represents the characteristic acceleration, a_c, that is required to produce the corresponding path. The final orbit that is plotted in Fig. 3.5 (i.e., the periodic orbit with the largest amplitude in the z direction) is defined for a characteristic acceleration of zero, indicating that it is a natural orbit. Two hodographs appear in Fig. 3.6. The first hodograph displays the x and z coordinates corresponding to the initial state along the trajectory in the xz plane for each orbit in the set. The second displays the maximum z value of each orbit in the set and the corresponding x and a_c values. Orbits corresponding to bifurcations as identified via the eigenstructure are indicated with a large, black dot (and note that black is not aligned with the color scales in Fig. 3.6). Recall that a_c is not time-dependent and is constant along an individual orbit. The approximation from Simo and McInnes [16] that is employed to initialize the differential corrections process is successful in producing periodic orbits for initial out-of-plane distances from zero to about 300 km below the EM orbit plane (an ellipse 100 km in x and 1900 km in y). Thereafter, the continuation scheme is employed, using the initial states from the previous orbits to predict the initial state for the current orbit. At the second bifurcation (represented by the middle thick orbit in Fig. 3.3), the continuation procedure yields the orbit that requires a maximum characteristic acceleration level equal to 1.6 mm/s^2. Other bifurcations along this set are also observed, indicating potential intersections with various families of solutions. Furthermore, bifurcations to various families of orbits are also possible between the orbits plotted in the figures; it is possible that the prediction algorithm skips over initial conditions that would result in a bifurcation. Further investigation of potential families of orbits might yield additional insight. The nonlinear orbits represented in

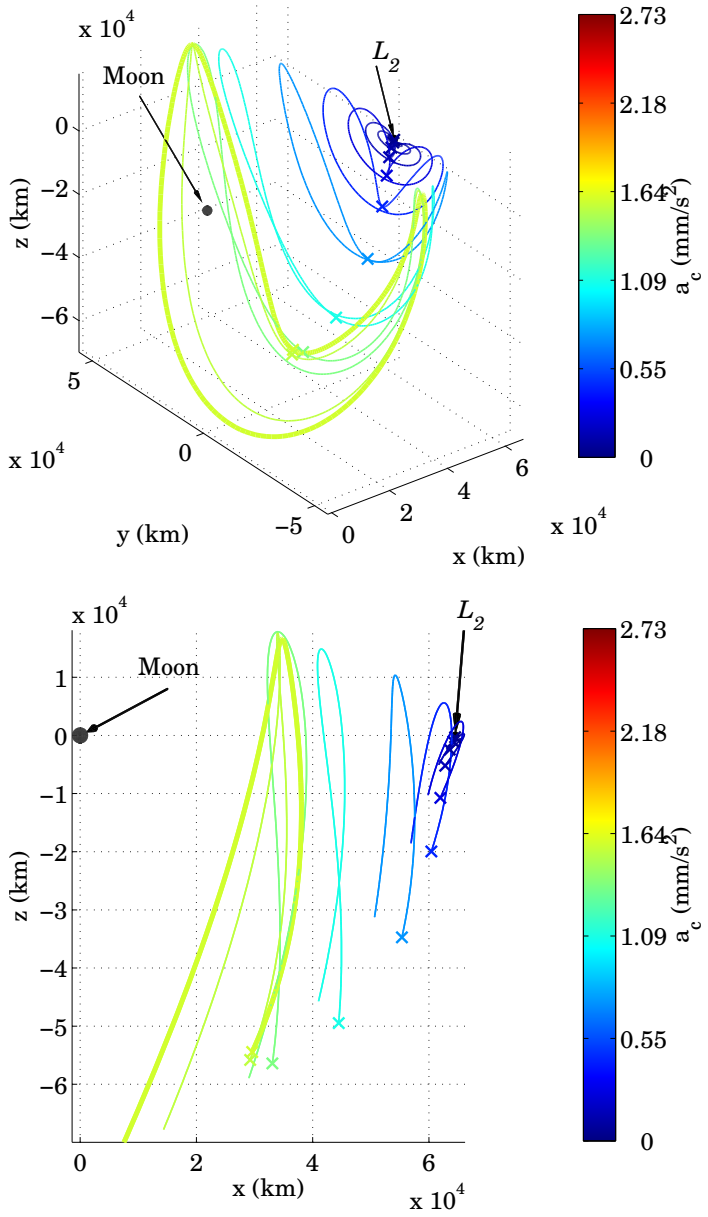


Fig. 3.2: A set of offset orbits in the vicinity of L_2. The Moon is plotted for scale.

Fig. 3.2 clearly do not supply solutions that achieve continuous communications coverage at the LSP. However, the do demonstrate the differential corrections process. The complete set of orbits also highlights the infinite number and some of the variety of periodic orbits that continue to exist even with the inclusion of a solar sail.

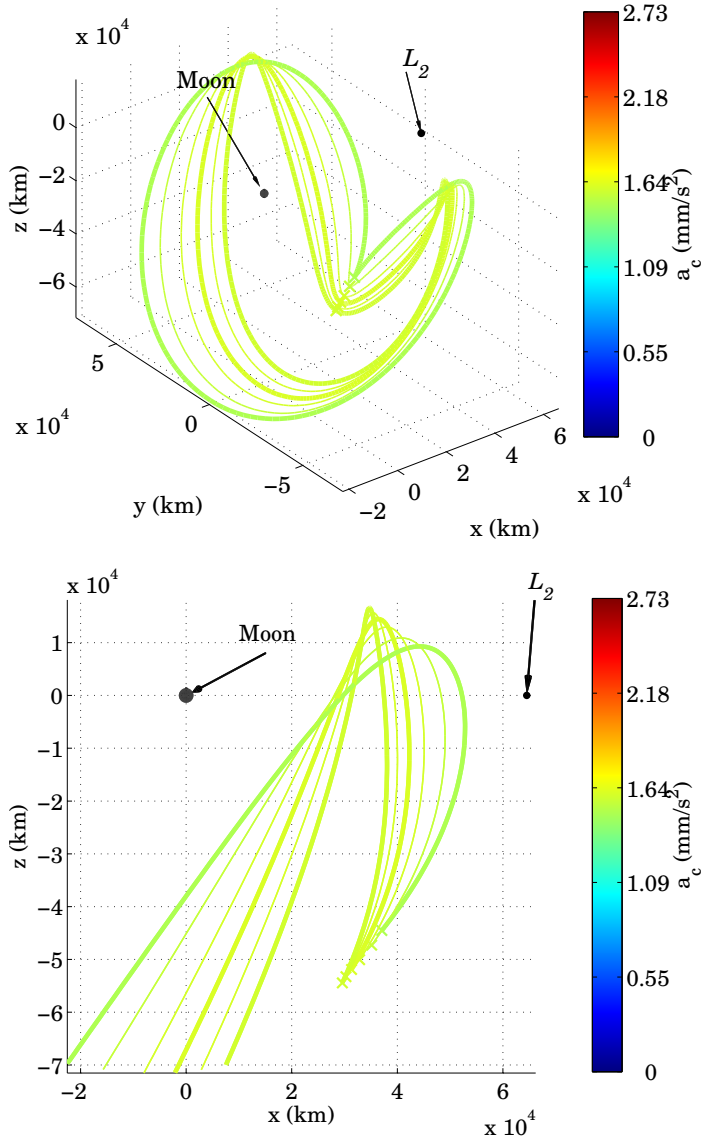


Fig. 3.3: A continuation of the set of orbits from Fig. 3.2.

The accuracy of a shooting method depends on the iteration and integration tolerances, as well as the model itself. The deviations at the end of a propagation sequence are compared to iteration tolerances to determine the convergence characteristics. The preceding example employed MATLAB®'s ODE113 explicit integration routine to propagate the orbits [123]. Similar to other MATLAB® propagators, the theoretical error associated with each propagated state at a given step is compared

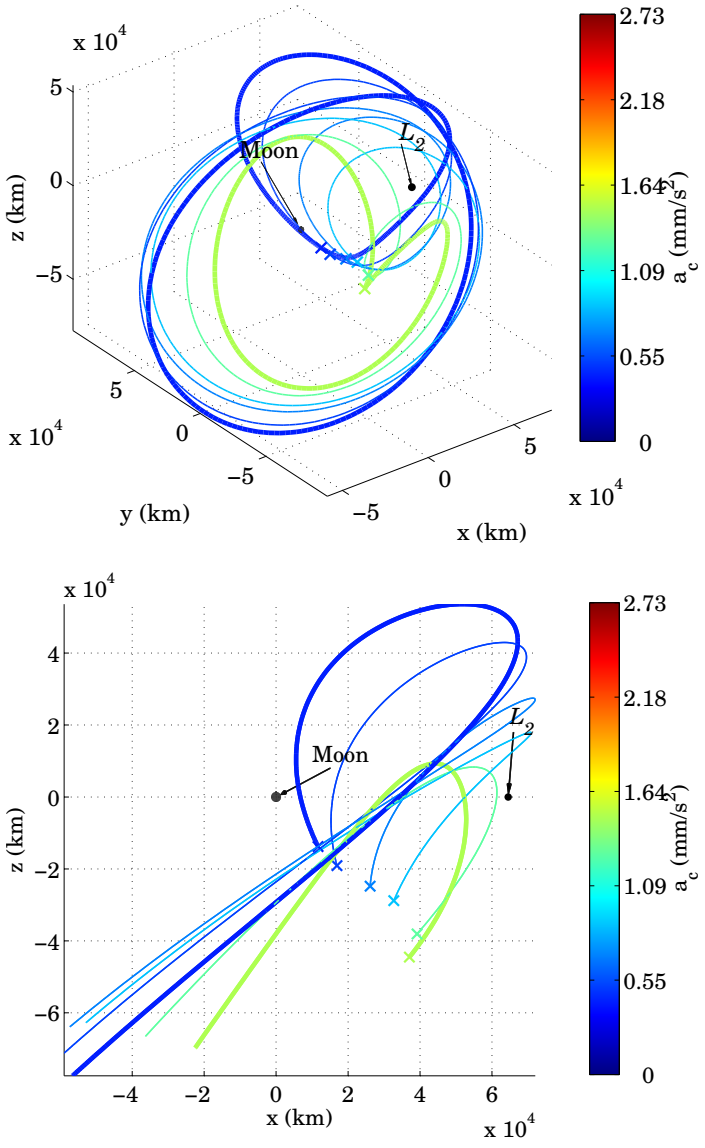


Fig. 3.4: A continuation of the set of orbits from Fig. 3.3.

to relative and absolute integration tolerances to control the step size (tolerances for both are equal to 1×10^{-13} in this simulation). In reality, a propagated trajectory diverges from a true path, but the magnitude of the divergence is reduced with tighter tolerances on the integration. Consequently, propagated trajectories are often regarded as "truth" models in mission design.

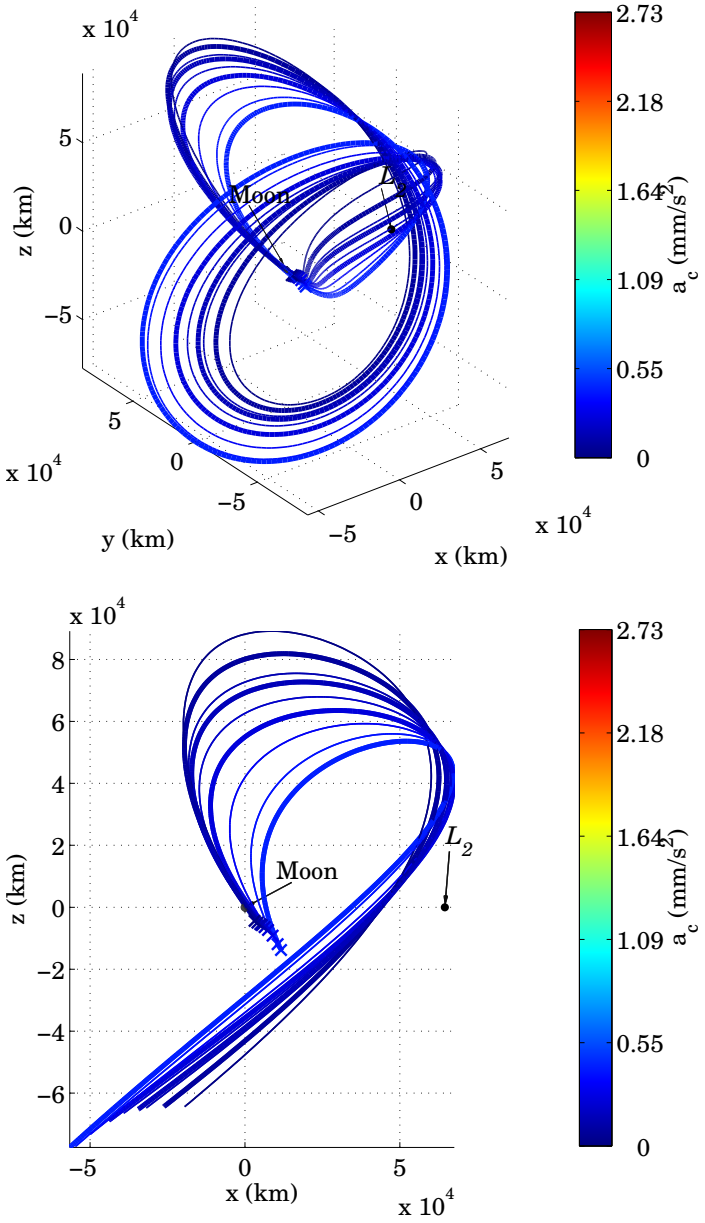
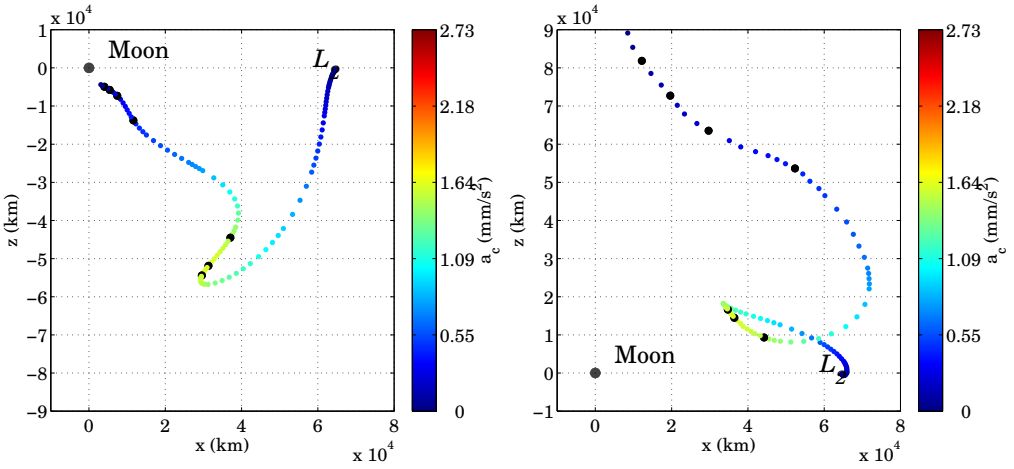


Fig. 3.5: A continuation of the set of orbits from Fig. 3.4.

3.2 Collocation

A popular family of algorithms to solve boundary value problems is based on *collocation*. In these approaches, both the trajectory and control are discretized and this set of states along the trajectory is determined via nonlinear programming



(a) x_0 and z_0 coordinates (b) maximum z and corresponding x

Fig. 3.6: Hodographs of the orbits appearing Figs. 3.2 through 3.5.

techniques. Collocation involves minimization of the difference between the derivative of a continuous approximating polynomial and the derivative from the differential equations at an intermediate, or defect, point (or points) along an arc between the nodes. The accuracy of the solution arising from a collocation strategy is related to the degree of the approximating polynomials; however, the complexity of the approach increases as well [109].

Suppose that, for a one-dimensional case, $\dot{x}(t) = f(x(t))$, where $f(x(t))$ is an ordinary differential equation, or an equation of motion, and $x(t)$ solves the equation of motion. Also, suppose that $S_k(t)$ is an N^{th}-degree polynomial that approximates $x(t)$ in the interval between the mesh points at τ_k and τ_{k+1}. If, at a time $\tau_a \in [\tau_k, \tau_{k+1}]$, $S_k(\tau_a)$ is equivalent to $x(\tau_a)$, then the approximating polynomial is said to be collocated to the solution of the differential equation at τ_a.

As an example of a collocation scheme, a third-degree polynomial (in red) based on state and derivative information at the mesh points (black) at τ_k and τ_{k+1} appears in Fig. 3.7 [109]. For the example in the figure, only states and derivative information from their associated differential equations are required at the boundaries to generate a third-degree polynomial. However, internal points between the mesh points may

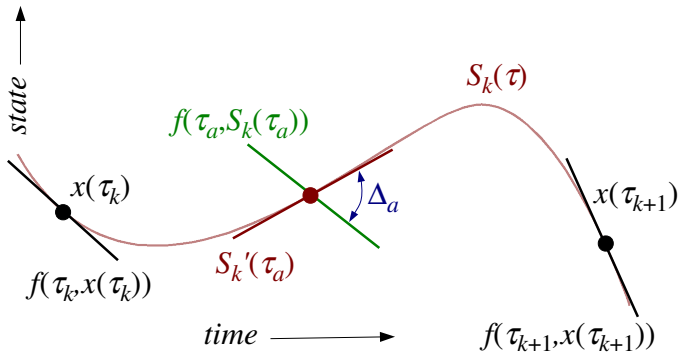


Fig. 3.7: Third-degree collocation example.

be required, depending on the degree of the collocating polynomial. At one or more epochs (τ_a in the figure) between τ_k and τ_{k+1}, the derivative of the polynomial (red in the figure) is compared to the derivative from the ordinary differential equation (green). The residual between these two values is a defect (Δ_a in the figure). Generally, $x(\tau_k)$ and $x(\tau_{k+1})$ are not precisely known initially, and any defect will be non-zero. Therefore, values for $x(\tau_k)$ and $x(\tau_{k+1})$ are updated through an iterative process, thereby changing the interpolating polynomial and derivative information from the equations of motion, until the defects are zero.

Because of their relative insensitivity to a poor initial guess and relative accuracy when compared to other approaches, collocation methods are increasingly common for solving BVPs, including optimal control problems. Optimization software packages such as DIDO [124] and GPOPS [125,126], as well as MATLAB®'s BVP*C functions (* = 4,5,6) [123,127], rely on collocation. The type of polynomial and the integration rules vary for different accuracy levels. Nassiri et al. [128] employ a technique that relies on a Hermite interpolating polynomial and Simpson quadrature rules (a.k.a., Hermite-Simpson collocation), which possess a local accuracy of $\mathscr{O}(\Delta t^5)$, to minimize the time-of-flight along a solar sail interplanetary transfer. Ozimek et al. [30] demonstrate: (1) a Hermite-Simpson collocation scheme for a solar sail in the EM system, as well as (2) a highly accurate method for application in the same system

that is based on a seventh-degree polynomial subject to Gauss-Lobatto integration constraints with a local accuracy of $\mathcal{O}(\Delta t^{12})$; both methods in Ozimek et al. [30] employ equally spaced nodes. In Ozimek et al. [129], mesh refinement is employed in the seventh-degree Gauss-Lobatto process, strategically redistributing and reducing the number of nodes required for a solution of equal accuracy across all nodes.

3.2.1 Collocation example

Collocation algorithms are successfully applied to the lunar south pole coverage problem by Ozimek [130], Grebow [131], and Ozimek et al. [30, 129]. Four sample orbits generated by employing the Hermite-Simpson collocation scheme[3] as described in Ozimek et al. [30] appear in Fig. 3.8. These orbits are generated with a minimum elevation-angle constraint of 15° as viewed from the LSP. The Moon is plotted to

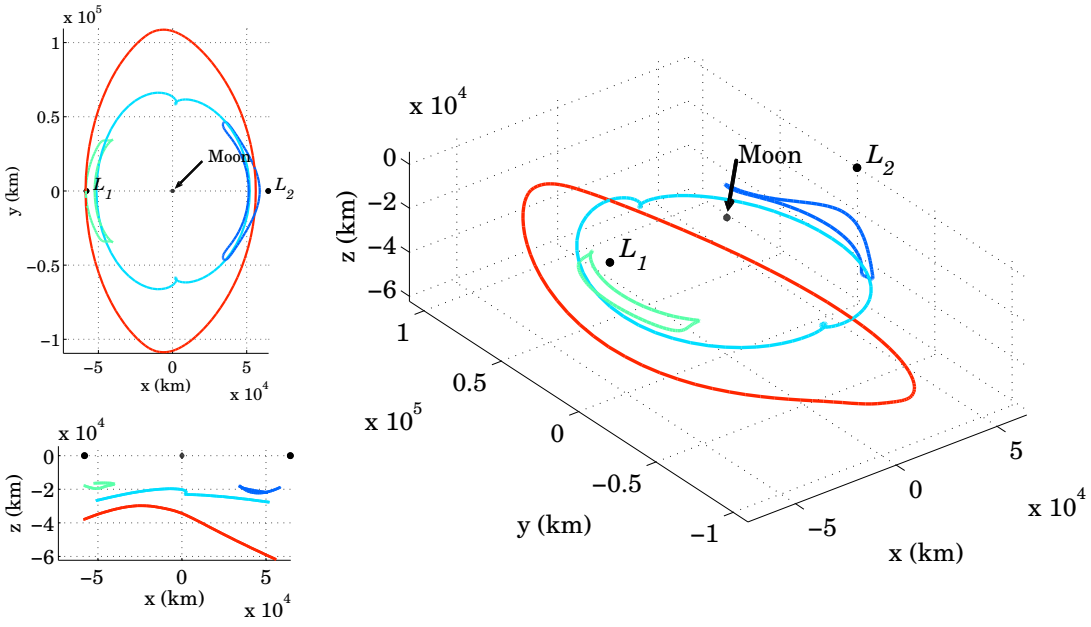


Fig. 3.8: Three views of four sample orbits generated by a Hermite-Simpson collocation method.

[3]With the exception of partial derivatives derived via central-difference approximations.

scale; $L_{1,2}$ appear as black dots. In this view in the rotating Earth–Moon frame, the Sun moves in a clockwise fashion about the z axis. The initial time is defined when the Sun is located along the x axis on the L_1 side of the Moon (in the direction of the Earth). At the initial time, the spacecraft are each positioned on the L_2 side of the Moon in their respective orbits, as well as in the xz plane. The initial guesses for the paths in these periodic orbits are concentric circles of arbitrary radii located below the Moon. The initial guess for the control profile is $\alpha = 35.26°$ and $\delta = 180°$ (refer to Fig. 2.3 for definitions of these angles) as the Sun revolves about the Earth–Moon system.

If a trajectory is developed using a higher-accuracy collocation scheme, a direct numerical integration process initiated by extracting a set of these states, along with the specified control history, often results in a simulated arc that resembles the original collocation solution. However, this observation in no way implies that solutions from one numerical approach are superior to those emerging from another method. Each periodic orbit generated here can serve as a reference trajectory for mission design. Naturally, navigation and flight-path control must be considered for an actual flight, since model uncertainties and mis-modeling result in an actual trajectory that diverges from any reference.

3.3 Augmented finite difference methods

A third category of numerical boundary value problem solvers are finite-difference methods (FDM). These approaches are not as popular as collocation or shooting methods; therefore, a thorough development is warranted. The derivation for a finite-difference method applied to the solar sail problem begins with a generic, nonlinear, second-order, two-point BVP that can be represented as

$$\ddot{\mathbf{r}} = \mathbf{f}(t, \mathbf{r}, \dot{\mathbf{r}}) \tag{3.6}$$

where $\mathbf{r}(t_1)$ and $\mathbf{r}(t_n)$ are fixed at the extremes of the time span, $[t_1, t_n]$ in interval notation [89, 90, 92, 93]. The solution of Eq. (3.6) is a trajectory, $\mathbf{r}(t)$. To solve the

BVP, the classical FDM discretizes the domain into nodes, or epochs, $t_1, t_2, \ldots, t_n$ and replaces the derivatives in Eq. (3.6) by their respective finite differences. Central-difference approximations (CDAs) are commonly used because they are more accurate than forward or backward differences. The first and second time derivatives of $\mathbf{r}(t)$ are, of course, the velocity and acceleration vectors, and are approximated by CDAs as

$$\widetilde{\mathbf{v}}_i = \frac{\mathbf{r}_{i+1} - \mathbf{r}_{i-1}}{t_{i+1} - t_{i-1}} \tag{3.7}$$

$$\widetilde{\mathbf{a}}_i = 2\frac{(\mathbf{r}_{i+1} - \mathbf{r}_i)\,(t_i - t_{i-1}) - (\mathbf{r}_i - \mathbf{r}_{i-1})\,(t_{i+1} - t_i)}{(t_i - t_{i-1})(t_{i+1} - t_{i-1})(t_{i+1} - t_i)} \tag{3.8}$$

where $\mathbf{r}(t_i)$ is defined as $\mathbf{r}_i$, $t_i \in (t_{i-1}, t_{i+1})$, and the symbol "$\widetilde{\ }$" indicates a numerical approximation. Equation (3.8) arises from three central differences: one for the velocity at $t_{i-1/2}$, which is midway between t_{i-1} and t_i, another at $t_{i+1/2}$, midway between t_i and t_{i+1}, and a third using the first two intermediate velocities resulting in the acceleration at t_i. The FDM does not require uniform node placement;[4] however, when the time steps between t_{i-1}, t_i, and t_{i+1} are fixed at a value of Δt, then the midpoint of $[t_{i-1}, t_{i+1}]$ is t_i and Eqs. (3.7) and (3.8) simplify to

$$\widetilde{\mathbf{v}}_i = \frac{\mathbf{r}_{i+1} - \mathbf{r}_{i-1}}{2\Delta t} \tag{3.9}$$

$$\widetilde{\mathbf{a}}_i = \frac{\mathbf{r}_{i+1} - 2\mathbf{r}_i + \mathbf{r}_{i-1}}{\Delta t^2} \tag{3.10}$$

Note that $\mathbf{r}_{n-1}$ is used for $\mathbf{r}_0$ when calculating $\widetilde{\mathbf{v}}_1$ for a periodic solution. Similarly, $\mathbf{r}_1$ substitutes for $\mathbf{r}_n$ in $\widetilde{\mathbf{v}}_{n-1}$, so the problem does not include redundant constraints. The calculations for $\mathbf{a}_1$ and $\mathbf{a}_{n-1}$ are similar. The differences between the actual velocity and acceleration and their respective CDAs at t_i are

$$\mathbf{v}_i - \widetilde{\mathbf{v}}_i = -\frac{\Delta t^2}{6}\mathbf{r}'''(\rho_v) \tag{3.11}$$

$$\mathbf{a}_i - \widetilde{\mathbf{a}}_i = -\frac{\Delta t^2}{12}\mathbf{r}^{\mathrm{iv}}(\rho_a) \tag{3.12}$$

[4]Mesh refinement and non-uniform node spacing may improve the accuracy of the FDM.

where $\rho_{v,a} \in (t_{i-1}, t_{i+1})$ and $\mathbf{r}'''(t)$ and $\mathbf{r}^{\text{iv}}(t)$ are continuous on $[t_{i-1}, t_{i+1}]$. Derivations of these truncation errors are available in the literature [89, 92, 93]. The approximations in Eqs. (3.9) and (3.10) replace the first and second time derivatives in Eq. (3.6), and the result is an equation at each epoch of the form

$$\frac{\mathbf{r}_{i-1} + 2\mathbf{r}_i + \mathbf{r}_{i+1}}{\Delta t^2} = \mathbf{f}\left(t, \mathbf{r}, \frac{\mathbf{r}_{i+1} - \mathbf{r}_{i-1}}{2\Delta t}\right) \tag{3.13}$$

Equation (3.13) may still be nonlinear, but it can be linearized into a system of equations, one for each t_i, excluding the extremes, and solved iteratively. Due to the approximations in Eqs. (3.9) and (3.10), the classic FDM results in a solution that will approximate each component of the true solution to the BVP by an error proportional to Δt^2 at each node [92, 93].

For the current problem, the FDM is augmented to incorporate path and system constraints, as well as a control history. Additionally, although BVPs usually require specific values for some or all of the boundary conditions, the FDM requires only that the first and the last state be equal for periodicity. Two variations of this FDM strategy are developed. In the first, position is the only explicitly discretized trajectory state (FDM-R). The second FDM is formulated to solve directly for position *and* velocity along the discretized trajectory (FDM-RV). Algebraically, these FDMs are equivalent. In practice, the two formulations yield slightly different results. Given that the objective is a simple and quick method to approximate the design space, neither strategy should be considered superior to the other; both achieve the stated goal.

As mentioned the FDM-R scheme directly calculates the three position elements at each node. The FDM-R formulation is based on approximating Eq. (2.6) via the CDAs for velocity and acceleration as represented in Eqs. (3.9) and (3.10), respectively, that is,

$$\widetilde{\mathbf{a}}_i + 2\left(\boldsymbol{\omega} \times \widetilde{\mathbf{v}}_i\right) + \boldsymbol{\nabla} U(\mathbf{r}_i) - \mathbf{a}_s(t_i) = \mathbf{0} \tag{3.14}$$

Note that the velocity term in Eq. (3.14) is replaced with its CDA, $\widetilde{\mathbf{v}}_i$. To solve the BVP, the trajectory is discretized at $t_1, t_2 \ldots, t_n$. Position is expressed in terms of Cartesian coordinates at each node,

$$\mathbf{r}_i = \left\{ \begin{array}{ccc} x_i & y_i & z_i \end{array} \right\}^{\mathrm{T}} \tag{3.15}$$

The control, $\mathbf{u}_i$, that is, the sail pointing vector in this application, as well as a set of m slack variables, η_i, are also included at each node. Normally, $\|\mathbf{u}\| = 1$, but for purposes of the FDM, the magnitude of each component of $\mathbf{u}$ is one when the process converges on a solution, but not necessarily unit magnitude at any other time when searching for a solution. (Throughout this document, $\hat{\mathbf{u}}$ will indicate a sail pointing vector known to be unit length and $\mathbf{u}$ will indicate a sail pointing vector not necessarily of unit length.) The positions, control, and slack variables at node i for the FDM-R are collected into the subvector,

$$\mathbf{q}_i = \left\{ \begin{array}{ccc} \mathbf{r}_i^{\mathrm{T}} & \mathbf{u}_i^{\mathrm{T}} & \boldsymbol{\eta}_i^{\mathrm{T}} \end{array} \right\}^{\mathrm{T}}_{(6+m)} \tag{3.16}$$

where the length of $\mathbf{q}_i$ is $(6 + m)$. The subvectors corresponding to each node are collected into a column vector,

$$\mathbf{X} = \left\{ \begin{array}{cccc} \mathbf{q}_1^{\mathrm{T}} & \mathbf{q}_2^{\mathrm{T}} & \cdots & \mathbf{q}_{n-1}^{\mathrm{T}} \quad \mathbf{q}_n^{\mathrm{T}} \end{array} \right\}^{\mathrm{T}}_{(6+m)n} \tag{3.17}$$

which represents the discretized trajectory, as well as the associated control history and slack variables. The velocities associated with the trajectories generated by the FDM-R can be reconstructed with CDAs of $\mathbf{r}(t)$ during post-processing.

The FDM-RV process is similarly formulated to the FDM-R. The most significant difference between the two approaches is that there are two distinct sets of ODEs to solve such that velocity is directly calculated by the FDM-RV, i.e.,

$$\widetilde{\mathbf{v}}_i - \mathbf{v}_i = \mathbf{0} \tag{3.18}$$

$$\widetilde{\mathbf{a}}_i + 2\left(\boldsymbol{\omega} \times \mathbf{v}_i\right) + \boldsymbol{\nabla} U(\mathbf{r}_i) - \mathbf{a}_s(t_i) = \mathbf{0} \tag{3.19}$$

Note the difference in the velocity terms in Eqs. (3.14) and (3.19). With this FDM-RV alternative, the subvector is

$$\mathbf{q}_i = \left\{ \begin{array}{cccc} \mathbf{r}_i^{\mathrm{T}} & \mathbf{v}_i^{\mathrm{T}} & \mathbf{u}_i^{\mathrm{T}} & \boldsymbol{\eta}_i^{\mathrm{T}} \end{array} \right\}_{(9+m)}^{\mathrm{T}} \tag{3.20}$$

since $\mathbf{v}_i$ is now explicitly part of the solution. This formulation results in an $\mathbf{X}$ vector with length $(9+m)n$.

Both the FDM-R and the FDM-RV algorithms are employed to update the initial guess of the elements in the state vector, $\mathbf{X}$. To accomplish this correction, the difference between the evaluated EOMs at node i and the associated numerically derived approximations is constrained to be zero. Typically, the initial guess for the elements in $\mathbf{X}$ will not result in a satisfied constraint. Additional constraints are required for periodicity, control-direction magnitude, and path characteristics. All of these constraints are dependent on $\mathbf{X}$ and are collected in the column vector $\mathbf{F}(\mathbf{X})$. An expansion of $\mathbf{F}(\mathbf{X})$ about $\mathbf{X}^j$ yields a linear approximation for $\mathbf{F}(\mathbf{X}^{j+1})$,[5] that is,

$$\mathbf{F}(\mathbf{X}^{j+1}) = \mathbf{F}(\mathbf{X}^j) + D\mathbf{F}(\mathbf{X}^j)\left(\mathbf{X}^{j+1} - \mathbf{X}^j\right) + \mathscr{O}(\Delta\mathbf{X}^2) \tag{3.21}$$

where the $D\mathbf{F}(\mathbf{X}^j)$ is the Jacobian matrix, $\frac{\partial \mathbf{F}(\mathbf{X}^j)}{\partial \mathbf{X}^j}$.[6] By assuming that $\mathbf{X}^j$ is in the neighborhood where $\mathbf{F}(\mathbf{X}^{j+1}) = \mathbf{0}$ (or $\|\mathbf{X}^{j+1} - \mathbf{X}^j\| \ll 1$), Eq. 3.21 is rearranged into a minimum-norm update equation, that is,

$$\mathbf{X}^{j+1} = \mathbf{X}^j - D\mathbf{F}(\mathbf{X}^j)^{\mathrm{T}}\left[D\mathbf{F}(\mathbf{X}^j)\cdot D\mathbf{F}(\mathbf{X}^j)^{\mathrm{T}}\right]^{-1}\mathbf{F}(\mathbf{X}^j) \tag{3.22}$$

Equations (3.21) and (3.22) reflect the Newton-Raphson method of root finding for several variables. For efficient computation, $D\mathbf{F}(\mathbf{X})$ is pre-allocated and stored in memory as a sparse matrix [123]. Of course, $\mathbf{F}(\mathbf{X}^{j+1})$ is not necessarily equal to zero after the first update of $\mathbf{X}$. Therefore, new values for $\mathbf{F}(\mathbf{X})$ are generated, a new Jacobian matrix is evaluated, and Eq. (3.22) is recomputed in subsequent iterations

[5]Superscripts on vectors refer to iteration number. The initial guess is denoted $j = 0$; a converged solution by $j = f$.

[6]Partial derivatives in the Jacobian matrix are derived analytically, but are not specifically developed here.

Eq. (3.22) until some convergence tolerance depending on $\mathbf{F}(\mathbf{X})$ or $\mathbf{X}$ is met. For this investigation, if $\mathbf{X}$ from Eq. (3.17) is written in terms of scalar elements, where

$$\mathbf{X} = \left\{ \begin{array}{cccc} \xi_1 & \xi_2 & \dots & \xi_{(6+m)n} \end{array} \right\}^{\mathrm{T}} \tag{3.23}$$

then a useful criterion for convergence is one such that the summed magnitude of the relative changes of each element in $\mathbf{X}$ is below the convergence tolerance, or

$$\sqrt{\sum_{k=1}^{(6+m)n} \left(\frac{\xi_k^{j+1} - \xi_k^j}{\xi_k^j} \right)^2} \leq tol \tag{3.24}$$

When an initial guess is in the neighborhood of a solution, convergence is quadratic. Specifying $tol \approx 10^{-7}$ is sufficient, since any further iteration approaches double precision. If the initial guess is not in the neighborhood of a solution, $\mathbf{X}^{j+1}$ bears little resemblance to $\mathbf{X}^j$; the step from $\mathbf{X}^j$ to $\mathbf{X}^{j+1}$ may even be chaotic. However, particularly in the current application, the new $\mathbf{X}^{j+1}$ might be near an alternate solution and subsequently lead to convergence. The converged discretized solution, $\mathbf{X}^f$, that satisfies these constraints is the trajectory that solves the EOMs to within a theoretical error level. Note that, because all elements are not the same type of quantities (i.e., position, velocity, control direction, slack variables), no reason exists that the different types will scale similarly, and it is possible that not all quantities are actually determined to the same accuracy.

3.3.1 Error analysis for the FDM

At each epoch, the converged subvector $\mathbf{q}_i$ differs from the true solution, $\overline{\mathbf{q}}_i$, due to the truncation error associated with the FDMs and the limited machine precision. The proceeding error analysis is described by Kincaid and Cheney [93, pp. 589–592]. For a step size Δt that is greater than some value, the error due to truncation in Eqs. (3.7)–(3.12) dominates the machine error. Assuming that there exists a true set

of position states, $\bar{\mathbf{r}}$, that solve Eqs. (2.6) and (3.11) as formulated for the FDM-R approach, Eq. (3.6) is written as

$$\frac{\bar{\mathbf{r}}_{i-1}+2\bar{\mathbf{r}}_i+\bar{\mathbf{r}}_{i+1}}{\Delta t^2}-\frac{\Delta t^2}{12}\bar{\mathbf{r}}^{\mathrm{iv}}(\rho_a)=-\frac{\boldsymbol{\omega}\times(\bar{\mathbf{r}}_{i+1}-\bar{\mathbf{r}}_{i-1})}{\Delta t}+\frac{\Delta t^2}{6}\boldsymbol{\omega}\times\bar{\mathbf{r}}'''(\rho_v)-\boldsymbol{\nabla}U(\bar{\mathbf{r}}_i)+\mathbf{a}_s(t_i) \tag{3.25}$$

The pseudo-gravity gradient associated with the true solution, $\bar{\mathbf{r}}_i$, is expanded as a function of an approximate solution, $\mathbf{r}_i$, that is,

$$\boldsymbol{\nabla}U(\bar{\mathbf{r}}_i)=\boldsymbol{\nabla}U(\mathbf{r}_i)+\mathbf{M}(\bar{\mathbf{r}}_i)\mathbf{e}_i+\mathscr{O}(\mathbf{e}_i^{\mathrm{T}}\mathbf{e}_i) \tag{3.26}$$

where $\mathbf{e}_i=\bar{\mathbf{r}}_i-\mathbf{r}_i$ and $\mathbf{M}(\bar{\mathbf{r}}_i)$ is the Hessian of U. Using the solution from the FDM-R algorithm, Eq. (3.13) is subtracted from Eq. (3.25), resulting in

$$\Delta t^{-2}\left(\mathbf{e}_{i-1}+2\mathbf{e}_i+\mathbf{e}_{i+1}\right)=-\Delta t^{-1}\left(\boldsymbol{\omega}\times(\mathbf{e}_{i+1}-\mathbf{e}_{i-1})\right)-\mathbf{M}_i\mathbf{e}_i+\Delta t^2\mathbf{h}_i \tag{3.27}$$

where $\mathbf{h}_i=\frac{1}{12}\bar{\mathbf{r}}^{\mathrm{iv}}(\rho_a)+\frac{1}{6}\bar{\mathbf{r}}'''(\rho_v)$. It is assumed that $\bar{\mathbf{r}}(t)$ is continuously differentiable to fourth order. Rearranging and multiplying by Δt^2 yields

$$(-\mathbf{I}+\Delta t\boldsymbol{\omega}\times)\mathbf{e}_{i-1}+(2\mathbf{I}-\Delta t^2\mathbf{M}_i)\mathbf{e}_i+(-\mathbf{I}-\Delta t\boldsymbol{\omega}\times)\mathbf{e}_{i+1}=-\Delta t^4\mathbf{h}_i \tag{3.28}$$

where $\mathbf{I}$ is the identity matrix and $\boldsymbol{\omega}\times$ is the skew-symmetric gyroscopic matrix. Some terms in Eq. (3.28) can be simplified to a form that will be useful later, that is

$$(-\mathbf{I}+\Delta t\boldsymbol{\omega}\times)+(2\mathbf{I}-\Delta t^2\mathbf{M}_i)+(-\mathbf{I}-\Delta t\boldsymbol{\omega}\times)=-\Delta t^2\mathbf{M}_i \tag{3.29}$$

Now, let $\boldsymbol{\lambda}$ correspond to the difference element $\mathbf{e}_i$ that possesses the largest magnitude. Equation (3.28) reduces to an upper bound on $\boldsymbol{\lambda}$,

$$(-\mathbf{I}+\Delta t\boldsymbol{\omega}\times)\boldsymbol{\lambda}+(2\mathbf{I}-\Delta t^2\mathbf{M}_i)\boldsymbol{\lambda}+(-\mathbf{I}-\Delta t\boldsymbol{\omega}\times)\boldsymbol{\lambda}\geq-\Delta t^4\mathbf{h}_i \tag{3.30}$$

Using Eq. (3.29), Eq. (3.30) simplifies to

$$\begin{aligned}-\Delta t^2\mathbf{M}_i\boldsymbol{\lambda}&\geq-\Delta t^4\mathbf{h}_i\\ \mathbf{M}_i\boldsymbol{\lambda}&\leq\Delta t^2\mathbf{h}_i\\ \boldsymbol{\lambda}&\leq\Delta t^2\mathbf{M}_i^{-1}\mathbf{h}_i\end{aligned} \tag{3.31}$$

The inequality in Eq. (3.31) represents three scalar equations. Thus, the position error corresponding to the FDM-R formulation is $\mathscr{O}(\Delta t^2)$ as $\Delta t \to 0$. To calculate the velocity from the FDM-R algorithm, the CDA from Eq. (3.9) is used, also with an error $\mathscr{O}(\Delta t^2)$. The error analysis for the alternate FDM-RV algorithm incorporates the fact that the velocity, $\mathbf{v} = \mathbf{f}(t, \mathbf{r})$, possesses errors that can also be expressed by Eq. (3.11); therefore, the derivation continues with the same steps as those for the FDM-R formulation beginning with Eq. (3.25).

The augmented finite-difference methods sacrifice precision for simplicity. At each node, the error in the trajectory is proportional to Δt^2, significantly less precise than common collocation algorithms [56, 106]. To improve the precision of the FDMs, smaller step sizes can be employed. In this analysis, $\Delta t = 0.067$, or $n = 101$, which translates to an error equal to 0.452% of the Earth–Moon distance, i.e., approximately 1740 km, in each direction. To improve the precision of the FDM, the step size between nodes, Δt, must decrease. The Jacobian matrix $D\mathbf{F}(\mathbf{X})$ becomes quite large, imposing a significant computational cost. Also, partial derivatives of $\widetilde{\mathbf{a}}_i$ from Eqs. (3.8) and (3.10) with respect to $\mathbf{r}_{i-1,i,i+1}$ are on the order of Δt^{-2}, whereas $\frac{\partial \mathbf{a}_i}{\partial \mathbf{r}_i}$ is independent of step size; a small Δt may lead to a Jacobian matrix, $D\mathbf{F}(\mathbf{X})$, in which the sensitivities of the numerical terms greatly exceeds those of the analytical terms. Furthermore, the smaller the step size, the larger the machine error [89]. Alternatively, Richardson extrapolation, or extrapolation to the limit, could be employed to further minimize the error in the solutions [89]. Incorporating a process to determine the specific step size that minimizes some combination of truncation and machine errors, or incorporating an additional process like Richardson extrapolation, adds complexity to a method that is formulated to be simple to understand and implement, as well as quick to yield results. The design space in this problem is not well known, so a trajectory that is precise to within Δt^2 relative to an actual solution is meaningful. If the goal is a general understanding of the design space, accomplished in a relatively quick analysis, then these results can be very insightful.

3.3.2 Algebraic constraint vector, F(X)

At the core of the augmented finite-difference approach, as developed here, is the algebraic constraint vector, $\mathbf{F}(\mathbf{X})$, that contains the algebraic approximations to the equations of motion and other constraints necessary for the desired periodic solar sail trajectory to emerge. In this formulation, each element in $\mathbf{F}(\mathbf{X})$ should be zero to satisfy simultaneously the ODEs and path constraints. For a periodic solar sail trajectory, subject to path constraints $\mathbf{g}(\mathbf{X})$,

$$\mathbf{F}(\mathbf{X}) = \left\{ \begin{array}{ccccccc} \Delta\mathbf{a}(\mathbf{X})^{\mathrm{T}} & [\Delta\mathbf{v}(\mathbf{X})^{\mathrm{T}}] & \mathbf{T}(\mathbf{X})^{\mathrm{T}} & y_1 & \mathbf{N}(\mathbf{X})^{\mathrm{T}} & \mathbf{g}(\mathbf{X})^{\mathrm{T}} \end{array} \right\}^{\mathrm{T}} \tag{3.32}$$

The presence of $\Delta\mathbf{v}(\mathbf{X})$ depends on whether the formulation is the FDM-R or FDM-RV, and the length of the $\mathbf{F}(\mathbf{X})$ vector is either $(2 + (4 + m)n)$ or $(2 + (7 + m)n)$, accordingly. The order of these elements within the vector is arbitrary, and no difference in performance is apparent when $\mathbf{F}(\mathbf{X})$ and $\mathbf{X}$ are rearranged such that $D\mathbf{F}(\mathbf{X})$ is a sparse, banded-diagonal matrix. Using the configuration in Eq. (3.32), each element set appears as block diagonal in the corresponding Jacobian matrix, $D\mathbf{F}(\mathbf{X})$. Each element set is subsequently discussed, and the corresponding size of each element set is indicated in a subscript outside of the braces.

The first element set in Eq. (3.32) is the difference in acceleration. The acceleration at a given epoch, t_i, as evaluated from the equations of motion, depends only on the position, velocity, and control, $\mathbf{a}_i = \mathbf{f}(\mathbf{r}_i, \mathbf{v}_i, \mathbf{u}_i)$. The numerically derived acceleration, $\widetilde{\mathbf{a}}_i$ from Eq. (3.8) or (3.10), depends on the state at multiple epochs. A valid trajectory possesses the same acceleration whether computed from the EOMs or from the CDAs, within the truncation error. Therefore, the difference between accelerations resulting from the EOMs and those from numerical calculation should nominally be zero for a converged solution; thus, the difference forms the first set of equality constraints in the composite constraint vector,

$$\Delta\mathbf{a}(\mathbf{X}) = \left\{ \begin{array}{c} \mathbf{a}_1(\mathbf{r}_1, \mathbf{v}_1, \mathbf{u}_1) - \widetilde{\mathbf{a}}_1 \\ \vdots \\ \mathbf{a}_{n-1}\left((\mathbf{r}, \mathbf{v}, \mathbf{u})|_{n-1}\right) - \widetilde{\mathbf{a}}_{n-1} \end{array} \right\}_{3(n-1)} \tag{3.33}$$

Likewise, the difference between the velocity from the EOMs and the numerically derived velocity is zero for a valid trajectory. Velocity from the EOMs at each epoch, $\mathbf{v}_i$, is available from $\mathbf{X}$ (see Eqs. (3.17) and (3.20)). This difference forms a second set of equality constraints (only in the FDM-RV algorithm) in $\mathbf{F}(\mathbf{X})$,

$$\Delta\mathbf{v}(\mathbf{X}) = \left\{ \begin{array}{c} \mathbf{v}_1 - \widetilde{\mathbf{v}}_1 \\ \vdots \\ \mathbf{v}_{n-1} - \widetilde{\mathbf{v}}_{n-1} \end{array} \right\}_{3(n-1)} \tag{3.34}$$

where $\mathbf{v}_i$ may be approximated as $\widetilde{\mathbf{v}}_i$. These differences in acceleration and velocity are also denoted as *defects* and are located at the node corresponding to each t_i. In an initial guess for the trajectory state vector, the defects $\Delta\mathbf{a}_i$ and $\Delta\mathbf{v}_i$ at each node will most likely not be zero. The Newton-Raphson iteration process adjusts the path of the trajectory to resolve these differences.

The next set of constraints enforces periodicity and unit length of the control vector. To enforce periodicity, the originating and final state vector constrained such that they are equal. Periodicity is represented as a set of constraints in $\mathbf{F}(\mathbf{X})$

$$\mathbf{T}(\mathbf{X}) = \left\{ \begin{array}{c} \mathbf{r}_n - \mathbf{r}_1 \\ [\mathbf{v}_n - \mathbf{v}_1] \\ \mathbf{u}_n - \mathbf{u}_1 \\ \boldsymbol{\eta}_n - \boldsymbol{\eta}_1 \end{array} \right\}_{(6|9+m)} \tag{3.35}$$

where m is the number of path constraints, expressed as slack variables. The length of $\mathbf{T}(\mathbf{X})$ depends on the application of an FDM-R or an FDM-RV approach. Although the only boundary condition is periodicity, for consistency the position at the boundary is constrained to be in the xz plane. Therefore, the y-coordinate at the first point along the trajectory is constrained, $y_1 = 0$. In practice, without this constraint, y_1 remains very close to zero, approximately 40 km from the xz plane once the solution is converged. With this constraint, y_1 is reduced to 0.004 mm from the

plane. Next, the magnitude of the control vector must remain of unit length in the iteration process

$$\mathbf{N}(\mathbf{X}) = \left\{ \begin{array}{c} \mathbf{u}_1^{\mathrm{T}}\mathbf{u}_1 - 1 \\ \vdots \\ \mathbf{u}_{n-1}^{\mathrm{T}}\mathbf{u}_{n-1} - 1 \end{array} \right\}_{(n-1)} \tag{3.36}$$

The formulation in Eq. (3.36) leads to a simpler partial derivative than if the magnitude is incorporated directly. This constraint is incorporated into the implementation of the FDM since the control history in this analysis is described in terms of a unit vector. Alternatively, the control history may be described in terms of angles, obviating this constraint.

Path constraints are included as inequality constraints, which are converted to equality constraints by the use of slack variables, a successful numerical adaptation from nonlinear programming [30, 132]. The slack variables are incorporated into $\mathbf{X}$ and are associated with the other state elements at their particular epoch, t_i. In this analysis, the elevation angle constraint, $E_{\min}$, maintains the visibility of the spacecraft from an outpost located near the south pole of the Moon, and the spacecraft altitude constraint, $A_{\max}$, is imposed for radio power restrictions. Altitude is defined as the distance from the lunar south pole, that is,

$$A_i = \sqrt{(x_i - 1 + \mu)^2 + y_i^2 + (z_i + R_m)^2} \tag{3.37}$$

The third path constraint requires that the sail-face normal, or control, $\mathbf{u}_i$, is always directed away from the Sun (the sunlight vector is $\hat{\boldsymbol{\ell}}_i$). Alternatively, $\alpha_{\max} = 90^\circ$ in Eq. (2.12). Of the inequality constraints, only this attitude requirement is mandated. Together, these three inequality constraints are written as

$$E_{\min} \leq E_i \equiv \arcsin\left(-\frac{z_i + R_m}{A_i}\right) \tag{3.38}$$

$$A_{\max} \geq A_i \tag{3.39}$$

$$\cos\alpha_{\max} \leq \hat{\boldsymbol{\ell}}_i \cdot \mathbf{u}_i \tag{3.40}$$

For the given problem and model, adding a path constraint to avoid the penumbra and umbra of the Earth's or the Moon's shadow is unnecessary because of the el-

evation constraint. A shadow constraint could be added for another application or shadowing effects could be incorporated into the dynamical model directly [133]. Additional inequality path constraints could include limits on the body turn rates and accelerations governed by the attitude control system of the spacecraft. The associated slack variables are squared and added to the inequality constraints, resulting in the following equality constraints

$$\mathbf{g}_i(\mathbf{r}_i, \boldsymbol{\eta}_i) = \left\{ \begin{array}{l} \sin E_{\min} + \frac{z_i + R_m}{A_i} + \eta_{E,i}^2 \\ A_i - A_{\max} + \eta_{A,i}^2 \\ \cos \alpha_{\max} - (\hat{\boldsymbol{\ell}}_i \cdot \mathbf{u}_i) + \eta_{\alpha,i}^2 \end{array} \right\}_m \tag{3.41}$$

Combining the path constraints at all epochs results in $m(n-1)$ total elements in $\mathbf{g}(\mathbf{X})$. When using the path constraints in Eq. (3.41), $m = 3$. The first two path constraints in Eq. (3.41), as well as the periodicity constraint in Eq. (3.35), are illustrated in Fig. 3.9. The $\cos \alpha_{\max}$ constraint appears in Fig. 2.2. The sailcraft in

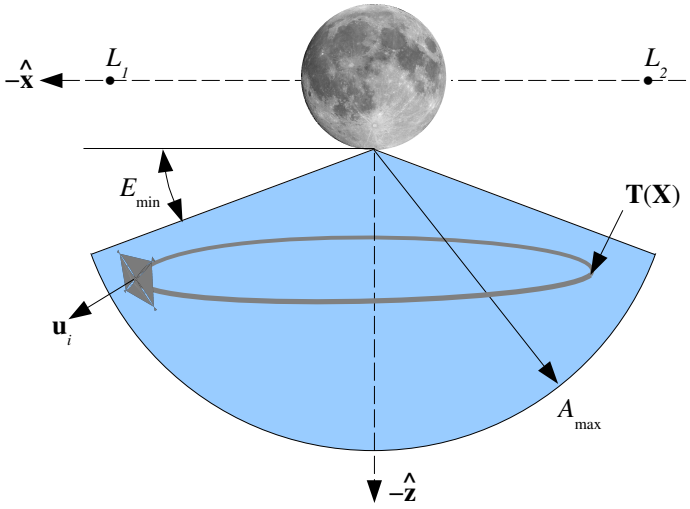


Fig. 3.9: Path constraints for an orbit below the Moon (Moon image from nasa.gov).

Fig. 3.9 orbits below the Moon. Feasible solutions also exist that do not orbit below the Moon, rather orbit below either the L_1 or the L_2 point.

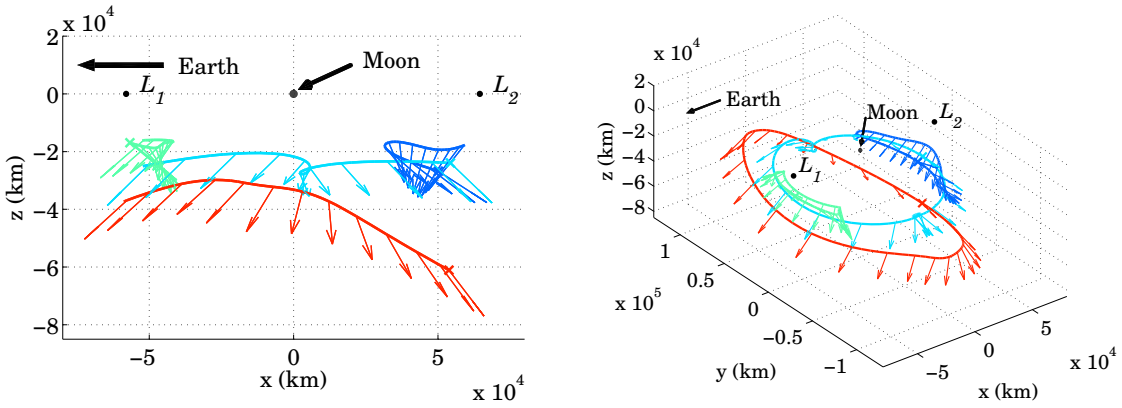
Only the first $(n-1)$ epochs are required in formulating $\Delta\mathbf{a}$, $\Delta\mathbf{v}$, $\mathbf{N}$, and $\mathbf{g}(\mathbf{X})$ in the constraint vector $\mathbf{F}(\mathbf{X})$. Due to periodicity, the first and n^{th} epochs are identical.

Activating the n^{th} epoch yields a problem that is over-constrained and the Jacobian matrix, $D\mathbf{F}(\mathbf{X})$, is not full rank.

3.3.3 Sample results from the FDM

Another four sample orbits, generated from an augmented finite-difference approach, appear in Fig. 3.10. The arrows directed away from points along the trajectory indicate the direction of the sail-face normal, $\hat{\mathbf{u}}$, and are separated in time by approximately one day. The Moon is plotted to scale. The path in each of these four solutions originates on the L_2 side of the Moon, in a location indicated by an "x." Motion along these paths is clockwise (retrograde), except for the dark-blue orbit. This dark-blue orbit is an example of a prograde orbit, which originates under L_1 and moves in a counterclockwise direction under the Moon. In this Earth–Moon frame, the Sun is initially aligned with the $-x$ axis, but moves clockwise about the Earth–Moon system with a period of one sidereal month (29.5 days); not coincidentally, the sailcraft trajectories possess the same period. Note that each of these orbits meets the constraints established previously and each is based on a characteristic acceleration of $a_c = 1.70$ mm/s^2. The color scheme for all orbits in this document is based on the maximum specific transverse torque for a spinning spacecraft, a property described in Sections 4.2.1 and 4.4.2. The associated control profiles for the four sample orbits in Fig. 3.10 appear in Fig. 3.11. Of these four trajectories, the cyan path under the Moon requires the smallest maximum pitch angle, α, and the red-orange path under the Moon is associated with the highest maximum pitch angle. These maximum angles occur on the right side ($+x$ side) of the cyan orbit in Fig. 3.10(a), when the Sun is initially positioned along the $-x$ axis (at day 0 in Fig. 3.11). The maximum pitch angle associated with the red-orange orbit occurs when the spacecraft is near the extremities in the $\pm y$ directions of that trajectory (at days 9 and 20 in Fig. 3.11).

Employing the FDM and incorporating the same initial guesses that are used to generate the orbits via the Hermite-Simpson collocation methods discussed previously,



(a) Side view, xz (b) 3-D view

Fig. 3.10: Four sample orbits generated from an augmented finite-difference method.

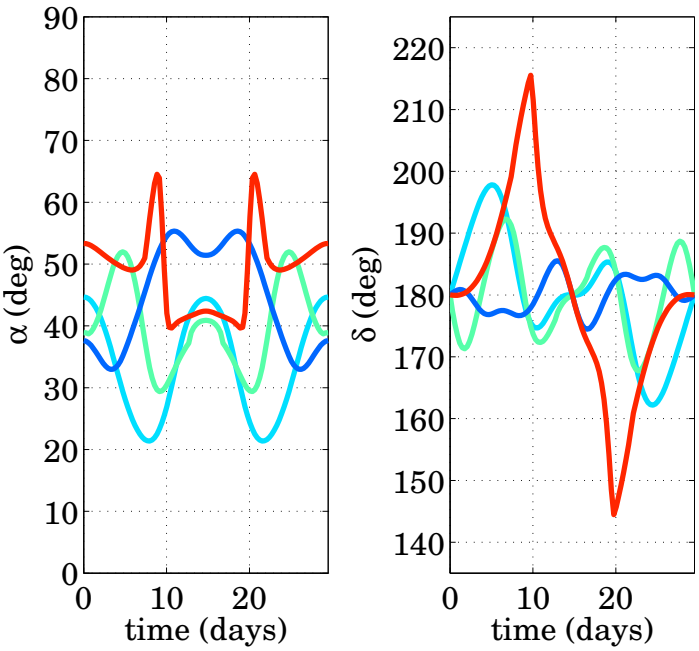


Fig. 3.11: Associated control profiles corresponding to the orbits in Fig. 3.10.

trajectories similar to those appearing in Fig. 3.8 emerge. The orbits generated by the Hermite-Simpson collocation scheme appear in Fig. 3.12. The only visible difference in the four periodic orbits is the shape of the "twists" along the y axis in the cyan trajectory. Near these twists, the two cyan orbits in Figs. 3.8 and 3.12 differ by

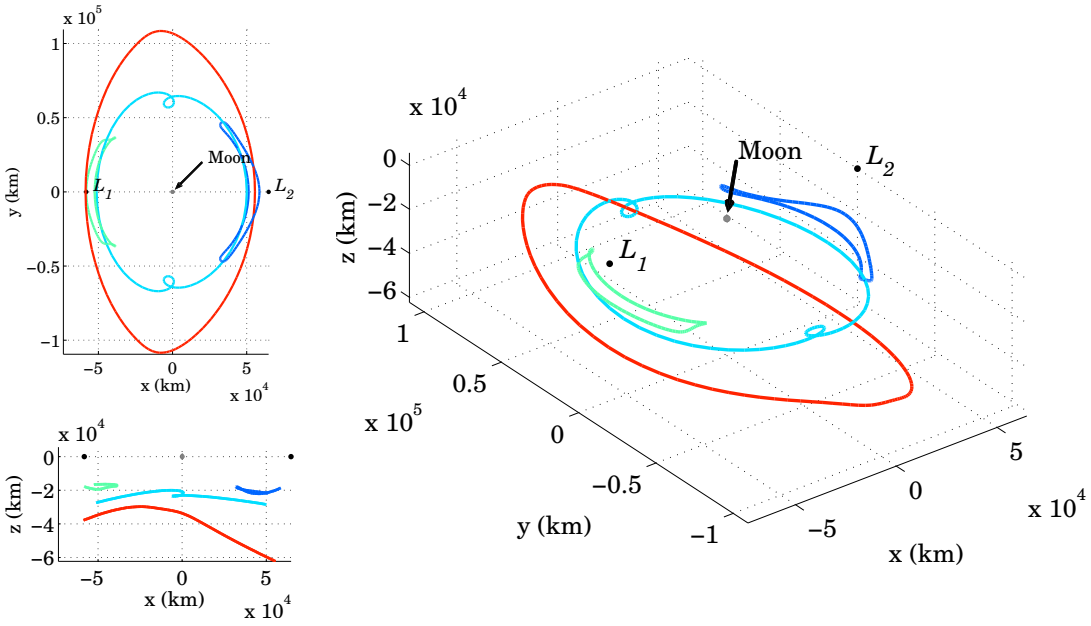


Fig. 3.12: Three views of four sample orbits generated by an augmented finite-difference method.

approximately 10,000 km. Likewise, the red orbits differ by approximately 1700 km, the green by approximately 2600 km, and the light-blue under L_2 by almost 3000 km. These differences are small considering that the initial guess for each pair of orbits is either a luni-axial circle (for those orbits centered below the Moon) or a stationary point (for the orbits offset below the Lagrange points) and generally not similar to the resulting arcs.

All numerical methods include some round-off and truncation error. Because both the FDM and the collocation strategies, as well as any shooting approach, are numerically based, solutions that are generated from these various numerical processes result in orbits that nearly—that is, do not exactly—solve the equations of motion. A goal is to produce solutions that approximate a true solution, if such a solution is available. However, when selecting a numerical procedure, the accuracy of the results must be weighed against the complexity of the procedure and the ability of that

procedure to converge on these approximate solutions. The FDM approach supplies solutions that are comparable to other, more sophisticated numerical approaches.

3.3.4 Structure of the Jacobian matrix

As mentioned, the Jacobian matrix is sparse. A diagram of the sparsity pattern based on partial derivatives of $\mathbf{F}(\mathbf{X})$ with respect to $\mathbf{X}$ for the FDM-RV formulation appears in Fig. 3.13. Diagonals corresponding to the specific element sets in $\mathbf{F}(\mathbf{X})$

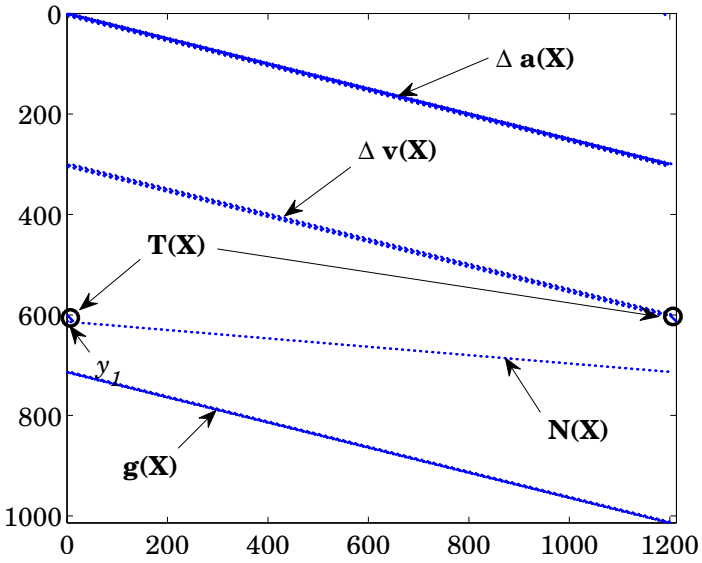


Fig. 3.13: Sparsity pattern for the Jacobian $D\mathbf{F}(\mathbf{X})$ in the FDM-RV formulation.

are labeled in the figure. For a scenario employing 101 nodes, the size of the Jacobian matrix is 1013 rows by 1212 columns where approximately 0.38% of the elements are non-zero. The sparsity pattern for the FDM-R formulation is similar to the matrix resulting from the FDM-RV formulation, except that the diagonal corresponding to $\mathbf{v}(\mathbf{X})$ does not exist in the FDM-R option. The size of the Jacobian matrix for the FDM-R option associated with 101 nodes is 710 rows by 909 columns; approximately 0.62% of the entries are non-zero.

The advantage of a finite-difference approach is its ease of implementation and the speed improvements enabled by its simplicity. Partial derivatives are easily accessible analytically, especially for an idealized force model, for example, the representation in Eq. (2.6). As the complexity of the formulation increases—certainly true with collocation, for example—analysts rely on either numerical or automatic differentiation to generate partial derivatives. MATLAB® supplies the function NUMJAC.M, which numerically approximates a derivative using a forward difference approximation [134]. Third-party software for MATLAB® exists for automatic (a.k.a. algorithmic) differentiation. TOMLAB® Optimization's MAD (MATLAB® Automatic Differentiation) suite employs automatic differentiation [135], and Shampine's PMAD (Poor Man's Automatic Differentiation) exploits complex step differentiation [136]. Both MAD and PMAD result in highly accurate derivatives, on the order of machine precision when compared to the partial derivatives determined analytically, while derivatives computed in NUMJAC.M are accurate to 10^{-8}. Because analytical derivatives are directly encoded into the algorithm, computation times are significantly faster when compared to other options. A comparison appears in Table 3.1. The computation

Table 3.1: Computation times for various differentiation strategies

Strategy	Time
Analytical	0.003269 seconds
PMAD	0.013798 seconds
NUMJAC	0.014924 seconds
TOMLAB/MAD	0.051849 seconds

times are compared for generating the matrix of partial derivatives, $\frac{\partial \boldsymbol{\nabla} U}{\partial \mathbf{r}}$, the most computationally intensive step in constructing an entry in the Jacobian matrix, for the four differentiation strategies. Clearly, if analytical derivatives are easily available, they appear to deliver the most time-efficient option. Furthermore, if the subroutine

to generate the derivatives analytically is programmed in a lower-level language (e.g., C or FORTRAN) and then complied as a MATLAB® Executable (MEX), further time savings are realized (with compiled code, the computation of $\frac{\partial \boldsymbol{\nabla} U}{\partial \mathbf{r}}$ requires less than 0.001 seconds). The formulation of the FDMs for this analysis exploits the computational advantage of analytical derivatives and compiled MEX code.

In summary, a finite-difference approach yields a solution for the path by replacing the path derivatives with their finite-difference approximations. The differences between the approximate derivatives and the derivatives determined by evaluating the differential equations at a specific point along a path are minimized by iteratively solving a linear system of equations. Inequality path constraints are added to this set of equality constraints by way of slack variables, and subsequently augmenting the constraint vector and linear system of equations, resulting in a feasible path.

4. DESIGN SPACE

The design space, that is, the solution space or feasible region, is one that meets all problem constraints and includes all candidate solutions. The feasible region contains one or more locally optimal solution and many other sub-optimal solutions. For traditional spacecraft that rely on chemical or electric propulsion, the cost that must be optimized is almost always propellant (or spacecraft mass). However, optimal solutions are often adjusted to accommodate spacecraft- and mission-design considerations not addressed when posing the original optimization problem. Examples include "soft" considerations such as the relative merits of one scientific plan or one operational strategy versus another.

Complex optimization problems with a plethora of locally optimal solutions require good initial guesses to seed their respective numerical-solution algorithms. If the solution space in general is known, an appropriate initial design is selected for further refinement. The solution space for the LSP coverage problem is not well understood for solar sails. Similar pole-sitter problems have been examined through analytical approaches [16, 119] as well as numerical techniques [30, 129]. However, the design space for the LSP coverage problem remains relatively unexplored. A numerical survey is a reasonable strategy to gain insight into this problem, whereby a large set of initial guesses for potential periodic solutions is employed to initialize a numerical process. A variety of initial guess combinations is investigated for the path, the nominal attitude profile, and the sail characteristic acceleration. The results are collected and searched for trajectories characterized by desirable features. These features include the lowest characteristic acceleration necessary to meet mission requirements, the torques needed to reorient the vehicle so that it maintains a given path, pitch angles along a trajectory, and ground-based critical metrics, such as elevation angle, altitude, and azimuth angle as viewed from an outpost at the lunar south pole.

4.1 Initial guess combinations

Not surprisingly, the initial guesses for the trajectory and the control history influence the resulting solution. Two types of initial guesses are explored for the path as well as six different types for the initial guess corresponding to the control history. The various combinations are summarized in Table 4.1. The superscript "0" indicates an initial guess for the associated variable.

Table 4.1: Summary of initial guess strategies for the trajectory and control history

IG* Trajectory	IG Control	Description
Cr		$\mathbf{r}^0$ is a retrograde circle, offset from Moon in the $-z$ direction
Cp		$\mathbf{r}^0$ is a prograde circle, offset from Moon in the $-z$ direction
P		$\mathbf{r}^0$ clustered at a point near the Moon in the southern xz half-plane
	α^*	$\hat{\mathbf{u}}_i^0$ south from Sun-line by 35.26°
	Moon	$\hat{\mathbf{u}}_i^0$ away from Moon
	EOM-NR	$\mathbf{u}_i^0$ satisfies EOM at each epoch, $\|\mathbf{u}_i^0\|$ not necessarily equal to 1
	RAA	$\hat{\mathbf{u}}_i^0$ is in the direction of the required applied acceleration at each epoch
	$\boldsymbol{\nabla} U$	$\hat{\mathbf{u}}_i^0$ points in the direction of $\boldsymbol{\nabla} U(\mathbf{r}(t_i))$
	$\hat{\boldsymbol{\ell}}$	$\hat{\mathbf{u}}_i^0$ points along the Sun-line, $\hat{\boldsymbol{\ell}}(t_i)$

*IG: Initial Guess

4.1.1 Initial guesses for the trajectory

All numerical methods for generating trajectories require an initial guess of that trajectory. The initial guess does not always need to be accurate for the algorithm to converge on a solution [30]. However, in general, accurate initial guesses lead to solutions from numerical procedures that resemble their respective initial guesses [91]. For this investigation, two families of initial guesses for the trajectory are examined: circular orbits and static points. Neither is necessarily considered accurate when initializing a numerical process, and an infinite number of other types of initial guesses are available (e.g., ellipses, sinusoidal functions, asymmetric orbits), but applying these two types of initial guesses for the trajectory leads to a greater understanding of the solution space in that an abundance of representative solutions emerge.

Circular orbits Due to the required periodicity of the converged solution, co-axial circles offset from the Moon in the z direction are selected as one option for an orbit to develop the initial guess: the x and y coordinates are defined by simple sinusoidal functions moving in a retrograde fashion and the z components are constant (cases labeled "Cr" in Table 4.1). For a retrograde orbit about the Moon, the spacecraft and the Sun are in opposition throughout the cycle. For a prograde orbit, the spacecraft is initially located between the Sun and the Moon, but moves counter-clockwise as the Sun moves clockwise about the Moon as viewed in the rotating frame (cases labeled "Cp" in Table 4.1). Associated initial guess velocities are the time derivatives of these sinusoidal functions. In some cases, the converged trajectories appear similar to their respective offset luni-axial circles; in most trials, however, the converged solution does not resemble the initial circle.

Static point Another possible option for the development of an initial guess for the trajectory is simply a static point, located initially in the xz plane as defined for the Earth–Moon CR3B system. This option is denoted as "P" in Table 4.1. Of course, only the two Lagrange points near the Moon preserve this initial guess as a

converged result, and only when both of the following conditions are satisfied: (1) if any constraint on elevation angle allows for a stationary trajectory at these points and (2) if the sail-face normal is orthogonal to the Sun-line, or "off," at all times. Nevertheless, the augmented finite-difference methods, developed by Wawrzyniak and Howell [34], converge on many different periodic solutions using this "P" strategy.

Four initial guesses for the path are illustrated in Fig. 4.1. Two initial guesses are circles, axially offset from the center of the Moon, and the two others are static-point initial guesses below L_1 and L_2. These circles and points are used as initial guesses for the sample orbits in Fig. 3.10. The xz plane below the Moon can be populated with initial guesses, potentially incorporated into a large simulation. In this analysis, for trials in which the initial trajectory guess is a point and in the xz plane, a grid spanning –75000 to 75000 km in the x direction and –75000 to 0 km in the z direction, each in 1000 km increments are used to generate the initial values, x^0 and z^0, for all points along the path. When the initial guess for the path is a concentric circle below

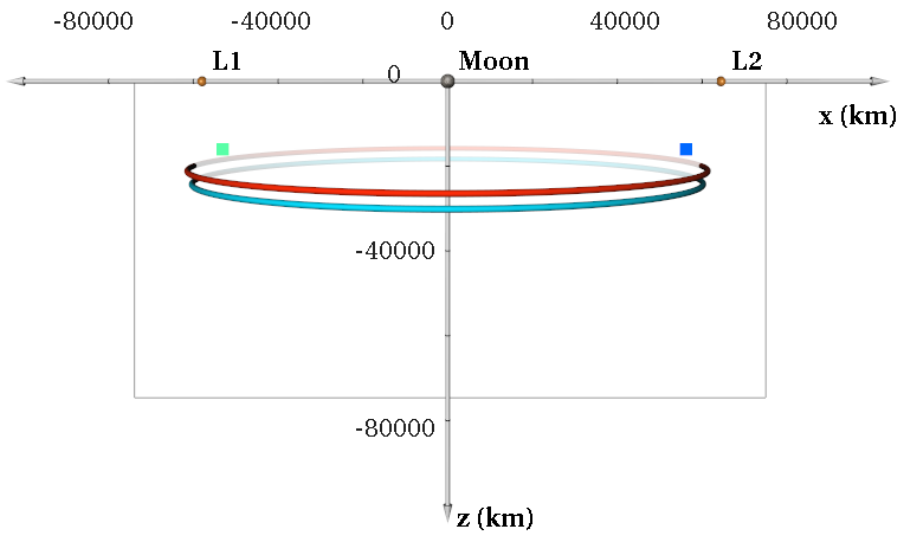


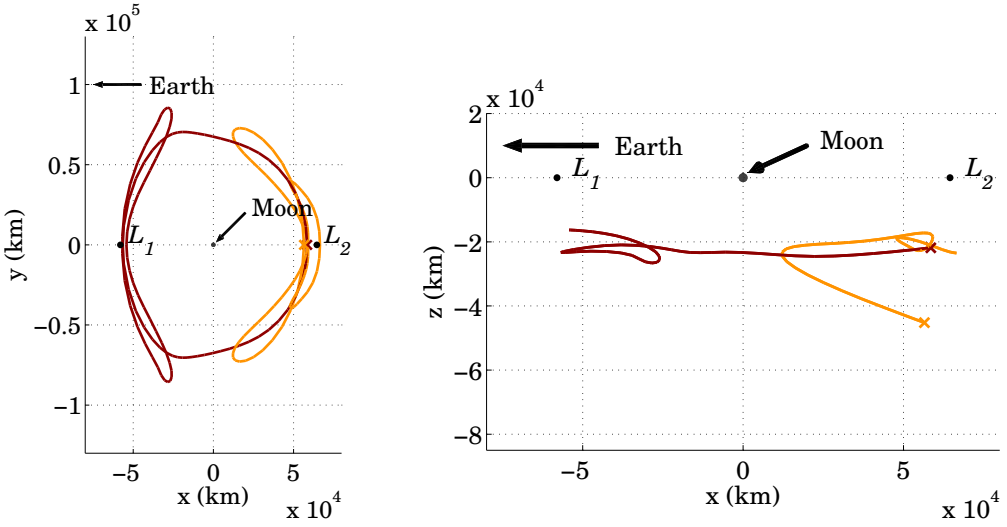
Fig. 4.1: Four sample initial guesses for the path. The origin of the coordinate system is the center of the Moon.

the Moon, a grid of radii and z-offset values span a region from 0 to 75000 km and –75000 to 0 km, respectively, in 1000 km increments. Recall that, at the initial time, the location of the Sun is aligned with the $-x$ direction. The initial state along the circular trajectories is in the xz plane to the right of the Moon.

Initial guess strategies where the circular orbit possesses a period of one half the synodic period or twice the synodic period are cursorily examined. Strategies to deliver circular orbits with periods half the synodic period of $2\pi/\Omega$ do not result in solutions with periods of π/Ω (i.e., two revolutions of the orbit per synodic cycle). When employing circular initial guesses with an initial period of $4\pi/\Omega$, the time frame of the simulation is extended to $[0, 4\pi/\Omega]$, allowing the Sun to make two revolutions about the system while the spacecraft completes one revolution.

Three types of solutions are observed when the simulation time is extended to twice the synodic period. (1) The first is where an orbit possessing a period of $2\pi/\Omega$ is simply repeated. The second and third types are actually two forms of a single-revolution solution. (2) In the second type, the path initially appears to be centered in a region below L_1, the Moon, or L_2. After $t = \pi/\Omega$, the path shifts to a different region, and then after $t = 3\pi/\Omega$, the path returns to the original region. A sample orbit illustrating this phenomenon appears in red in Fig. 4.2. Originating at the red "x" below L_2, the spacecraft moves along the path in a clockwise direction. The center of the path shifts to a region under L_1 when $t = \pi/\Omega$ and the Sun is located along the $+x$ axis. The motion after $t = 2\pi/\Omega$ mirrors the motion before that time. (3) The third type of solution behaves similar to the second, except that the center of the path remains under the same location. The orange orbit in Fig. 4.2 exhibits this behavior. For both of these sample orbits, the initial spacecraft motion is in the $-y$ direction (recall that the Sun initially moves in the $+y$ direction and completes one clockwise revolution about the system during the synodic period of $2\pi/\Omega$).

These multiple-revolution solutions demonstrate both their existence and that the augmented finite-difference methods in Wawrzyniak and Howell [34] can generate trajectories with time spans longer than one synodic period. Because the sample



(a) Top view, xy (b) Side view, xz

Fig. 4.2: Sample double orbits. Note that the initial guesses for these orbits are circular paths that have periods twice that of the synodic period.

space already exceeds 10 million combinations of initial guesses where the path is a circle with a period of $2\pi/\Omega$, initial guesses where the path is initially a circle with a period of some multiple of $2\pi/\Omega$ are not included in this investigation.

4.1.2 Initial guesses for the control history

The "control history" refers to the direction of the sail-face normal, or applied thrust vector; the control directions are initially defined at discrete points along the entire orbit. Six concepts are explored as potential initial guesses for the control history (producing a total of twelve combined initial guess strategies).

An optimal attitude: "α^*" The first control strategy simply maximizes the out-of-plane force contributed by the sail ("α^*" in Table 4.1). Derived analytically by

McInnes [6], the sail-pitch angle that maximizes the out-of-plane thrust is $\alpha^* = 35.26°$. The initial guess for the thrust vector is then

$$\hat{\mathbf{u}}_i^0 = \cos(\Omega t_i)\cos\alpha^*\hat{\mathbf{x}} - \sin(\Omega t_i)\cos\alpha^*\hat{\mathbf{y}} + \sin\alpha^*\hat{\mathbf{z}} \tag{4.1}$$

This initial guess for the control strategy, combined with a static-point-type initial guess for the trajectory, is very successfully applied by Ozimek et al. [30].

Sail-face normal directed away from the Moon: "Moon" The next strategy that serves as an initial guess option for the control is a thrust vector directed away from the center of the Moon throughout the initial guess trajectory ("Moon" in Table 4.1). Thus, if a vector, $\mathbf{r}_i^c$, is defined as the difference between the position of the Moon and the position of the spacecraft in the rotating frame at time $t = t_i$, or

$$\mathbf{r}_i^c = (x_i - 1 + \mu)\hat{\mathbf{x}} + y_i\hat{\mathbf{y}} + z_i\hat{\mathbf{z}} \tag{4.2}$$

then the initial guess for the control strategy is

$$\hat{\mathbf{u}}_i^0 = \frac{\mathbf{r}_i^c}{\|\mathbf{r}_i^c\|} \tag{4.3}$$

By definition, $\hat{\mathbf{u}}_i^0$ possesses unit magnitude.

Satisfying the equations of motion for a guessed trajectory: "EOM-NR" A third initial control strategy involves the definition of the control vector at each epoch, $\mathbf{u}_i$, by satisfying the equations of motion for either a circular or static-point-type initial guess for the path, that is,

$$\mathbf{f}(\mathbf{u}_i^0) = \widetilde{\mathbf{a}}_i + 2\left({}^I\boldsymbol{\omega}^R \times \widetilde{\mathbf{v}}_i\right) + \boldsymbol{\nabla}U(\mathbf{r}_i) - \frac{a_c}{a^*}(\hat{\boldsymbol{\ell}}(t_i)\cdot\mathbf{u}_i^0)^2\mathbf{u}_i^0 = \mathbf{0} \tag{4.4}$$

where $\widetilde{\mathbf{a}}_i$ and $\widetilde{\mathbf{v}}_i$ indicate a central-difference approximation for the acceleration and velocity, respectively, based on the initial guess for the path (a circle or a point). A Newton-Raphson iteration scheme is used to solve this nonlinear equation, with each $\mathbf{u}_i^0$ initially directed away from the Moon. The converged control history then serves as the input control history, $\mathbf{u}^0$, to the numerical algorithms. Unlike the other

strategies, the control vector does not initially possess a magnitude of one, and the numerical solution process (e.g., finite-difference method) is expected to render a viable trajectory and a control profile where each $\hat{\mathbf{u}}_i$ is unit length. This strategy is labeled "EOM-NR" in Table 4.1.

Required acceleration: "RAA" A fourth, simpler initial control strategy assumes that the initial guess trajectory is already a solution to the equations of motion, Eq. (2.6), with the caveat that the sail supplies any additional, required applied acceleration without regard to feasibility or practicality (i.e., the sail is assumed capable of unlimited, variable thrust, independent of direction). Therefore, if the applied acceleration that is required to solve the equations of motion appears as

$$\mathbf{a}_{raa,i} = \widetilde{\mathbf{a}}_i + 2\left({}^I\boldsymbol{\omega}^R \times \widetilde{\mathbf{v}}_i\right) + \boldsymbol{\nabla} U(\mathbf{r}_i) \tag{4.5}$$

where $\widetilde{\mathbf{a}}_i$ and $\widetilde{\mathbf{v}}_i$ indicate central-difference approximations of the acceleration and velocity, respectively, based on the initial guess for the path (a circle or a point). The initial guess for the control strategy is

$$\hat{\mathbf{u}}_i^0 = \frac{\mathbf{a}_{raa,i}}{\|\mathbf{a}_{raa,i}\|} \tag{4.6}$$

This strategy is labeled "RAA" for "required applied acceleration" in Table 4.1.

Parallel to the pseudo-gravity gradient: "$\boldsymbol{\nabla} U$" The fifth option in delivering the initial guess for the control history (labeled "$\boldsymbol{\nabla} U$" in Table 4.1) is similar in concept to the RAA as presented in Eqs. (4.5) and (4.6), except that only the pseudo-gravity gradient, $\boldsymbol{\nabla} U(\mathbf{r})$, is involved, that is,

$$\hat{\mathbf{u}}_i^0 = \frac{\boldsymbol{\nabla} U(\mathbf{r})}{\|\boldsymbol{\nabla} U(\mathbf{r})\|} \tag{4.7}$$

Of course, if the sailcraft moves relative to the rotating frame, for example, when the trajectory is a circle, the relative and the Coriolis acceleration terms are non-zero. Therefore, this initial guess may be more appropriate in combination with a static point initial guess for the path. The pseudo-gravity gradient for the control direction

is a critical quantity to develop equilibrium surfaces in the problem and may yield some insight (e.g., McInnes [6]).

Along the Sun-line: "$\hat{\boldsymbol{\ell}}$" The final initial guess concept for the control history is simple and completely independent of the initial guess associated with the trajectory. To seed the corrections process, the control strategy assumes that the sail-face normal is parallel to the sunlight direction (labeled "$\hat{\boldsymbol{\ell}}$" in Table 4.1).

For all initial guess combinations, the directions of the sail-face normal at the nodes along the trajectory, as well as the path itself, align to solve the equations of motion via a numerical boundary value problem solver. For this analysis, the augmented finite-difference methods presented by Wawrzyniak and Howell are employed [34]. Note that the number of nodes, n, also affects whether or not a solution converges. Throughout this analysis, $n = 101$.

4.1.3 Characteristic acceleration

For an ideal sail, the characteristic acceleration, a_c, is the parameter that encapsulates the sailcraft area, mass, and reflective properties. For a perfectly reflecting solar sail, the characteristic acceleration, from Eq. (1.8), is

$$a_c = \eta \frac{2P}{\sigma_s + m_p/A} \tag{4.8}$$

where P is the nominal solar radiation pressure at 1 AU (4.56 μN/m^2), m_p is the mass of the payload and spacecraft, excluding the sail and associated support structure, and A is the area of the sail. The $2P$ in the numerator in Eq. (4.8) assumes a perfect specular reflection, and, thus, momentum transfer from the photons, on a flat sail. An efficiency factor, η, represents absorption and non-perfect reflection and the value is typically 0.85–0.90 [6]. The sail-loading parameter, σ_s (or areal density), is simply a mass to area ratio corresponding to the sail and associated structure and is the primary metric for hardware performance.

Current values for the sailcraft characteristic acceleration are generally less than 1 mm/s^2. A recent sailcraft design for NASA's Space Technology competition (ST9) by L'Garde possesses an overall characteristic acceleration, a_c, of 0.58 mm/s^2, while the characteristic acceleration of the sail and its support structure alone is closer to 1.70 mm/s^2 (0.212 to 0.623 in units of nondimensional acceleration, respectively) [79]. The study in Ref. [29] assumes that the ST9 sailcraft can be scaled so that the overall characteristic acceleration is 1.20 mm/s^2. The IKAROS spacecraft, the first mission to demonstrate the deployment of a sail in space successfully, is designed with a characteristic acceleration of 0.364×10^{-3} mm/s^2 [137]. NASA recently launched and deployed the NanoSail-D2, which possesses a characteristic acceleration of 0.02 mm/s^2 [81]. The Planetary Society's *LightSail-1*, which is composed of three CubeSats, is expected to deliver a characteristic acceleration of 0.057 mm/s^2 [77]. These sails (and others) are designed to demonstrate the deployment of a sail in space and measure the effect of solar radiation pressure on a sail. Future solar sail flight projects may be designed to maximize the characteristic acceleration. This investigation surveys a range of characteristic accelerations to evaluate the probable level of technology that is required to support a sailcraft mission that addresses the LSP coverage problem.

4.2 Critical metrics

A trajectory designer typically constructs an orbit while considering various trade-offs. Formulating the design strategy as a single-objective optimization problem is not generally appropriate, as many variables can constitute the "cost." In extending the formulation to a multi-objective optimization problem, suitable solutions may be overlooked, and, furthermore, poor knowledge of the solution space impedes progress. In essence, the design goals typically involve optimizing "operability" while meeting mission constraints.

Path constraints, such as minimum elevation angle and maximum altitude (i.e., Eqs. (3.38) and (3.39)), are incorporated into the trajectory design when designing an

orbit to satisfy the requirements for a specific mission scenario. For a given spacecraft configuration, the elevation angle might be maximized for more margin in a visibility requirement, or, if the vehicle's motion is contained below one of the Lagrange points, to reduce the range of azimuth angles, perhaps allowing a fixed antenna at the south pole. Alternatively, a lower maximum altitude also can reduce the required antenna power.

Trajectory considerations must also be balanced with the design of the sailcraft itself. Orbits associated with lower characteristic accelerations are more feasible with near-future technology; however, fewer trajectory options are available. Orbits for an ideal sail that are associated with high pitch angles (α) along the trajectory may be infeasible when the forces on the sail are modeled with higher fidelity (e.g., optical force model), as previously noted. Finally, solutions that require smaller torques to orient the sailcraft to the required attitude may be easier to maintain. The objective in designing a solar sail spacecraft is to minimize mass while maximizing area. Incorporating an attitude control system adds mass; thus, this final metric of "turnability," along with characteristic acceleration, is of great interest.

4.2.1 Specific transverse torque, or "turnability"

The trajectories generated for this survey are based on a sailcraft that continuously reorients in inertial space. A continuous control profile requires that the attitude control system is able supply appropriate torques in the transverse body axes to reorient the vehicle throughout flight. Because the inertia properties of various sailcraft will be different, the concept of specific transverse torque is introduced to assess the "turnability" necessary for the sailcraft to follow a particular trajectory.

Some recently published results for solar sail attitude control are relevant. Sailcraft, like other space vehicles, can be spin-stabilized, three-axis stabilized, or stabilized by a gravity gradient, among other methods. Of the sailcraft that are spin-stabilized, the attitude can be modified by thrusters [138, 139], or by translating and

tilting the sail panels [138], or by reflectivity control devices that adjust the sail optical properties [140]. Three-axis stabilized sailcraft may be reoriented via (1) reaction wheels [141, 142], or control vanes [141], (2) purposefully offsetting the center of mass and the center of pressure that yields a torque via a movable boom [143] or translating mass [144, 145], (3) small thrusters at the tips of the sail masts [144], (4) magnetic torquers, or (5) a combination of two or more of these devices [146]. Other concepts yield sailcraft that are stabilized via gravity gradients and magnetic torquers [147, 148].

To model rotational motion, the sail is assumed to be a rigid body. Aligning the body axes with the principal moments of inertia, Euler's familiar vector equation of motion governs the relationship between the applied torques, the body rates, and the angular acceleration, that is,

$$\mathbf{M} = \mathbf{I} \cdot {}^{I}\dot{\boldsymbol{\omega}}^{S} + {}^{I}\boldsymbol{\omega}^{S} \times \mathbf{I} \cdot {}^{I}\boldsymbol{\omega}^{S} \tag{4.9}$$

where $\mathbf{I}$ is the central, principal inertia dyadic, ${}^{I}\boldsymbol{\omega}^{S}$ is the angular velocity of the sail body-fixed frame with respect to the inertial frame, and $\mathbf{M}$ is the vector of external torques that is required to control the spacecraft attitude. The derivation of ${}^{I}\boldsymbol{\omega}^{S}$, as well as ${}^{I}\dot{\boldsymbol{\omega}}^{S}$, as functions of Eulerian rotations appear in Appendix B. Solar sails are generally designed to be symmetric about one principal axis of inertia, such that an axial moment of inertia, I_a, is approximately twice that of the transverse moment of inertia, I_t. If the x-axis in the body frame is the axis of symmetry and the spacecraft spins about that axis at a constant rate, ω_{x0}, then Eq. (4.9), reduces to the following scalar equations

$$\begin{aligned} M_x &= 2I_t\cancelto{0}{\dot{\omega}_{x0}} \\ M_y &= I_t(\dot{\omega}_y - \omega_z\omega_{x0}) \\ M_z &= I_t(\dot{\omega}_z + \omega_{x0}\omega_y) \end{aligned} \tag{4.10}$$

where ω_{x0}, ω_y, and ω_z are the components of ${}^I\boldsymbol{\omega}^S$ in a body-fixed frame. If the sailcraft is three-axis stabilized (i.e., not spinning and $\omega_{x0} = 0$), Eqs. (4.10) further reduce to

$$\begin{aligned} M_x &= 0 \\ M_y &= I_t\dot{\omega}_y \\ M_z &= I_t\dot{\omega}_z \end{aligned} \tag{4.11}$$

It is clear from both the spinning (Eq. (4.10)) and the three-axis stabilized (Eq. (4.11)) types of motion that the ability to complete successfully attitude turns is limited by the available torques in the pitch and yaw directions (y and z axes, respectively).

Two recent studies lend insight into the relationship between the capabilities of a three-axis stabilized sailcraft and the associated maximum angular accelerations. Citing key driving performance requirements from three previous NASA mission studies where the control mechanism is a movable boom and the spacecraft is three-axis stabilized, Price et al. [143] establishes an upper limit for pitch and yaw accelerations of 2.3×10^{-9} deg/s^2. More recently, Wie [141] establishes maximum pitch and yaw accelerations of 28.1×10^{-6} deg/s^2 and a maximum turn rate of 0.02 deg/s when the control mechanism for a three-axis stabilized sailcraft is sail panel translation and rotation. The parameters for a 40-by-40 meter sail design, used by Wie, originate from a mix of ST6 and ST7 designs [141]: the sail moments of inertia are $I_x = 6000$ kg $\cdot$ m^2 and $I_y, I_z = 3000$ kg $\cdot$ m^2 with an areal density of 0.111 kg/m^2 and a characteristic acceleration of 0.0737 mm/s^2. The maximum roll-control torque for the ST6/7 sailcraft of $\pm 1.34 \times 10^{-3}$ N $\cdot$ m and the maximum pitch- and yaw-control torques are $\pm 1.45 \times 10^{-3}$ N $\cdot$ m; the corresponding maximum angular accelerations are $\pm 13.0 \times 10^{-6}$ deg/s^2 and $\pm 28.1 \times 10^{-6}$ deg/s^2, respectively.

When the spacecraft is spin-stabilized such that ω_x is constant, i.e., $\omega_x = \omega_{x0}$, the cross terms in Eq. (4.10) must be considered. In a three-axis stabilized case, Eq. (4.11), the available torque in a particular direction is employed to rotate the spacecraft about that direction. When a spacecraft is spin-stabilized about the x axis, the same turn requires additional torque about the y and z axes. A configuration

similar to the three-axis spacecraft is again employed by Wie for the analysis of a sailcraft spinning at 5 rotations per hour; in that study, Wie assumes that a torque of 0.001 $\mathrm{N}\cdot\mathrm{m}$ can be delivered to reorient the spinning spacecraft [138]. This torque corresponds to a maximum angular acceleration of approximately $\pm 20 \times 10^{-6}$ deg/s^2 in the transverse axes. The IKAROS sailcraft employs liquid crystal displays to change the reflective properties of the sail material for turning; however, its primary attitude control system is a set of thrusters on the central bus [137].

If the required rotational rates and accelerations associated with a particular trajectory are known, a specific torque about each axis can be determined. Because the labeling of the transverse y and z axes is arbitrary, the concept of a specific transverse torque is convenient. This metric characterizes the vehicle's ability to turn and is calculated from the transverse torques in either Eq. (4.10) or Eq. (4.11), that is,

$$\mathscr{M}_t = \sqrt{\left(\frac{M_y}{I_t}\right)^2 + \left(\frac{M_z}{I_t}\right)^2} \tag{4.12}$$

For a given trajectory and nominal control profile, different specific transverse torques are required that depend on the spin rate. The maximum specific transverse torques corresponding to both a three-axis stabilized sailcraft (non-spinning), $\mathscr{M}_{t,\mathrm{3AS}}$, and one spinning at 5 rotations per hour, $\mathscr{M}_{t,\mathrm{spin}}$, are examined.

4.2.2 Summary of critical metrics for "operability"

This investigation is focused on six critical metrics for assessing different sailcraft trajectories and the associated control profiles. In summary, these metrics are

- Sail characteristic acceleration, a_c
- Maximum specific transverse torque at any time along the orbit, $\mathscr{M}_t$
- Largest pitch angle at any time along the orbit, α
- Lowest elevation angle at any time along the orbit, E
- Highest altitude at any time along the orbit, A
- Largest range of azimuth angles throughout the orbit

The first two items on this list are specific to the vehicle design; the final four items are properties of the specific trajectory option. While the items in this list may be incorporated in a multi-variable optimization scheme as either parameters to optimize or as constraints, the goal of this investigation is an improved understanding of the solution space, motivating a survey. The insight gained from this survey is necessary to initialize a process that optimizes one or more of these critical metrics. Finally, additional critical metrics may be identified, or some current metrics may be disregarded at some future time. A new search would then be conducted using these same techniques.

4.3 Framework for the survey

The primary purpose of using a simple, lower-fidelity, augmented finite-difference method to propagate states and compute solutions is to examine the design space for a solar-sail spacecraft in orbit near the Moon quickly and easily. Using the MATLAB numerical computing environment, over 10 million combinations of initial guesses are used to generate trajectories for the survey. The computation employed up to eight cores on five platforms and took approximately one week. Different combinations of initial guess strategies lead to different solutions. This is especially true when comparing trajectories that arise from circular initial guesses for the path to those that arise from initial guesses that collapse the trajectory to a single point. A general survey of the solution space for all combinations of initial guesses is a first step in understanding the design options. A viable solution generated from any method possesses characteristics that can be employed to select any particular orbit for further study.

The survey is developed to incorporate two versions of a finite-difference approximation [34], three strategies to produce an initial guess for the path, and six strategies to deliver an initial control profile. This investigation also includes an examination of various combinations of elevation-angle constraints. The three path strategies in-

clude a retrograde circular orbit, a point hovering in the xz plane, and a prograde circular orbit. The circular trajectories all possess periods equal to the solar synodic period. Of the initial control profiles, the α^* and $\hat{\boldsymbol{\ell}}$ approaches are most successful in generating solutions for any type of initial path. The EOM-NR and Moon control approaches only result in converged solutions when the initial guess for the path is a circular retrograde orbit. The $\boldsymbol{\nabla} U$ and the RAA approaches are successful when the initial guess for the path is a prograde or a retrograde circle, but not when the initial guess is a point.

Any one initial control strategy is not necessarily superior to any other, however, the α^* and $\hat{\boldsymbol{\ell}}$ strategies converged more often than the other types of initial guesses for the control profile. An appropriate plan for examining the design space is the selection of a range of characteristic accelerations and path constraints and, then, the generation of solutions based on multiple combinations of initial guesses for the path and control.

4.4 Survey results pertaining to spacecraft-driven critical metrics

Critical metrics generally emerge as one of two types: those that drive the spacecraft design and those that drive the ground station design. To fly a particular path, the vehicle must possess a specific characteristic acceleration, a_c, be able to change its orientation as necessary, and produce thrust at sufficiently high pitch angles. Highlighted below are solutions from the survey that require the smallest a_c, smallest specific transverse torques, and smallest maximum pitch angles in their respective categories.

4.4.1 Minimum a_c to achieve various elevation-angle constraints

This analysis is formulated in a CR3B model. In actuality, the obliquity of the lunar orbit with respect to the Earth is 6.688°, and the Moon's orbit is a secularly precessing ellipse with respect to an inertial frame. Additional margin to accommo-

date lunar surface features and incorporate a more realistic ephemeris model is added to the obliquity to determine $E_{\min}$. For the broader survey of the design space to support a lunar communications relay, an elevation-angle constraint of $E_{\min} = 15°$ and altitude constraint of $A_{\max} = 384{,}400$ km is assumed (see Fig. 3.9 for an illustration of $E_{\min}$ and $A_{\max}$). For these constraints, the smallest sail characteristic acceleration examined in this survey that satisfies these constraints is $a_c = 1.3$ mm/s^2. Smaller values of a_c examined in this survey (i.e., 1.25 mm/s^2) are insufficient to push the vehicle sufficiently far below the Moon to satisfy the elevation-angle constraint.

In the present survey, when the elevation-angle constraint is relaxed, solutions emerge for sails with smaller characteristic accelerations. Moon-centered offset orbits conforming to the lower constraints might not be viewable at all times from an outpost near the lunar south pole. Orbits offset below L_1 and L_2 may be useful as relays for a base near lower lunar latitudes, but on the near or far side of the Moon, respectively, and may require lower elevation-angle constraints. In this investigation, the lowest overall characteristic accelerations for a spacecraft equipped with a solar sail that meet these elevation-angle constraints under the three locations are listed in Table 4.2. Since a discrete set of characteristic accelerations is considered, the actual

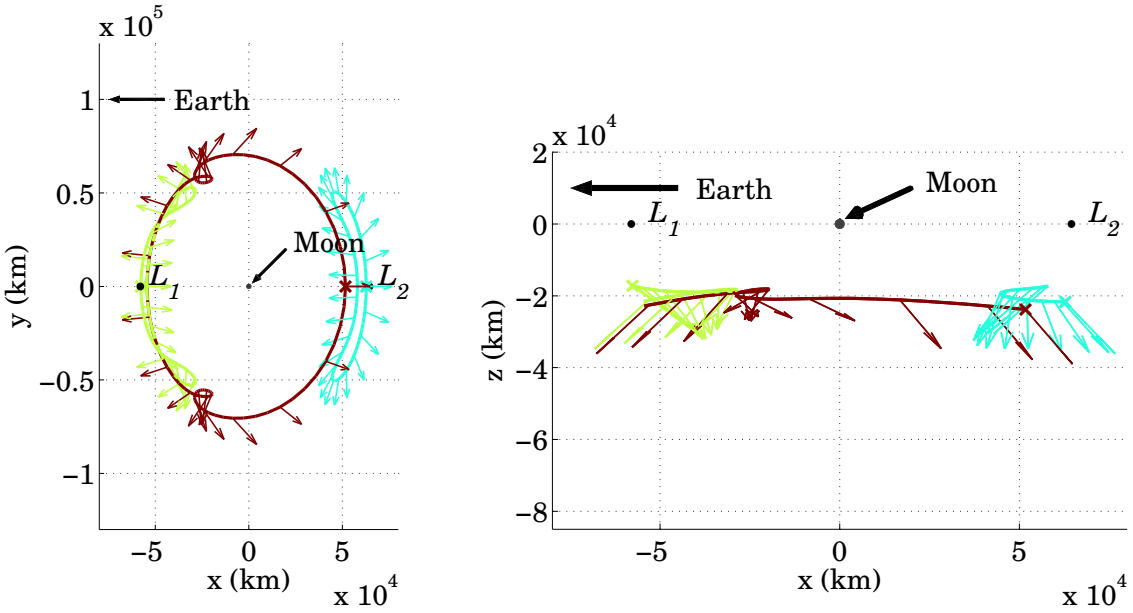
Table 4.2: Smallest a_c (mm/s^2) required for orbits located below the Moon and the Earth–Moon L_1 and L_2 points based on elevation angle

$E_{\min}$:	4°	6°	9°	12°	15°
L_1	0.58	1.00	1.25	1.50	1.60
Moon	0.55	0.75	1.00	1.50	1.60
L_2	0.40	0.55	1.00	1.25	1.30

minimum value of a_c that actually conforms to the constraint may be slightly less than the values in this table. The data in the table indicate that sails with higher characteristic accelerations are required for orbits below the Earth–Moon L_1 point or the Moon for a given elevation-angle constraint; relatively lower characteristic ac-

celerations are viable for sail trajectories below the L_2 point. Certainly, a_c could be minimized using an optimization scheme, and this broad survey simply offers a sense of the solutions available. The results from this survey yield a starting point for an optimization scheme based solely on characteristic acceleration or in combination with other metrics of interest.

A variety of solutions exists for sails with the lowest value of characteristic acceleration that is required for the three locations. One such orbit appears for each location in Fig. 4.3. As before, the Sun is defined to move clockwise about the Moon



(a) Top view, xy

(b) Side view, xz

Fig. 4.3: Orbits below L_1, the Moon, and L_2 that conform to the 15° elevation-angle constraint and possess the smallest possible characteristic acceleration, a_c.

as viewed in Fig. 4.3(a), originating at a position along the $-x$ axis. The maroon and turquoise paths originate opposite to the Sun, while the lime-green path originates on the Sun-side with respect to its center. Initial motion along all three paths proceeds in the $-y$ direction as the Sun moves in the $+y$ direction. In fact, the Sun moves in a clockwise, retrograde, fashion one half-period out of phase with a spacecraft moving along the retrograde maroon and turquoise paths. Motion along the lime-green path,

however, is prograde, as a spacecraft along that path moves in a counter-clockwise fashion. Because of the motion of the Sun, prograde paths originate on the left side of the orbit and motion is counter-clockwise while retrograde paths originate on the right side and motion is clockwise as projected on the xy plane. These trajectories, and generally all orbits conforming to the 15° elevation angles-constraint, are displaced approximately 20000 km below the Earth–Moon orbital plane. Recall that the period of each orbit is one sidereal month (29.5 days).

For a given value of a_c, an abundance of solutions exists. From a set of solutions with the lowest a_c, the trajectories appearing in Fig. 4.3 are selected based on the smallest maximum specific transverse torque ($\mathscr{M}_{t,\text{spin}}$) along the orbit for a spinning sailcraft.[1] Critical metrics, such as the maximum $\mathscr{M}_{t,\text{3AS}}$ if the sailcraft is three-axis stabilized (3AS), the maximum pitch angle (α), and the minimum elevation angle (E), along these three trajectories appear in Table 4.3. Pitch angles associated with the orbits in Fig. 4.3 are all less than 55°, indicating that the solutions may successfully transition to a higher-fidelity sailcraft SRP model.

Table 4.3: Critical metrics for orbits conforming to $E > 15°$, selected by characteristic acceleration (a_c) and $\mathscr{M}_{t,\text{spin}}$ requirements

	a_c	Max $\mathscr{M}_{t,\text{spin}}$	Max $\mathscr{M}_{t,\text{3AS}}$	Max α	Min E
L_1	1.60 mm/s^2	1.21×10^{-6} deg/s^2	2.26×10^{-9} deg/s^2	43.99°	15.00°
Moon	1.60 mm/s^2	4.39×10^{-6} deg/s^2	11.91×10^{-9} deg/s^2	48.66°	15.00°
L_2	1.30 mm/s^2	0.78×10^{-6} deg/s^2	1.33×10^{-9} deg/s^2	51.91°	15.00°

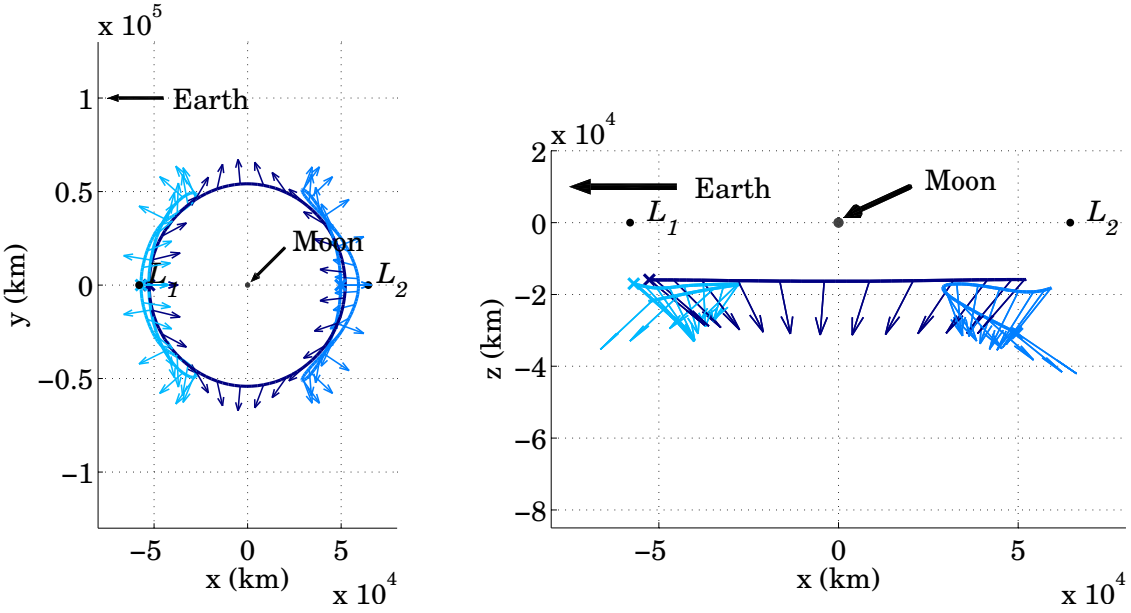
As the value of the sailcraft characteristic acceleration is increased, more trajectory solutions are available. By fixing $a_c = 1.70$ mm/s^2, a comparison with other critical metrics is possible. A sail characteristic acceleration this large is not possible with

[1] If the smallest maximum $\mathscr{M}_{t,\text{3AS}}$ along the orbit for a three-axis stabilized sailcraft is employed as the selection criterion, similar solutions appear under the Lagrange points and the same solution appears below the Moon.

current technology. However, based on the results in Table 4.2, a sail must possess a characteristic acceleration near this magnitude to satisfy the mission constraints for this scenario (e.g., $E \geq 15^\circ$ under L_1, L_2, and the Moon), and a reasonable requirement for future sail hardware technology is now established. For the rest of this investigation, to compare the broadest set of metrics, $a_c = 1.70$ mm/s^2. Note that the techniques employed to conduct this survey are applicable to missions with different path constraints and sailcraft with different physical characteristics.

4.4.2 Specific transverse torques

The next critical metric, or mission characteristic, for analysis is specific transverse torque, a measure of the attitude control authority necessary for a continuously reorienting sailcraft to follow a particular path. Sample orbits with the smallest maximum value of $\mathscr{M}_{t,\text{spin}}$ for a spin-stabilized spacecraft corresponding to a sail characteristic acceleration of $a_c = 1.70$ mm/s^2 appear in Fig. 4.4. For the dark-blue path below the Moon in Fig. 4.4, the trajectory originates on the $-x$ side of the Moon and the spacecraft moves in a counter-clockwise, prograde, fashion. The color scheme employed for all orbits in this document is based on a logarithmic jet scale, where deep blue corresponds to a $\mathscr{M}_{t,\text{spin}}$ of 2.3457×10^{-7} deg/s^2 (which is associated with the dark blue path below the Moon in Fig. 4.4) and a deep red corresponds to a $\mathscr{M}_{t,\text{spin}}$ of 4.3857×10^{-6} deg/s^2 (associated with the red orbit under the Moon in Fig. 4.3). Note that the arrows that represent the direction of the sail-face normal for the portions of the path near the x axis appear to be directed "inward" as viewed from above the xy plane (Fig. 4.4(a)) for a prograde orbit and "outward" for a retrograde orbit; the sail-face normal is constrained to be directed away from the Sun, which moves clockwise about the system. Consequently, the sunlight direction is generally aligned with the projection of the sail-face normal vector when the normal vector is projected into the xy plane. Both paths associated with the smallest maximum $\mathscr{M}_{t,\text{spin}}$ below



(a) Top view, xy

(b) Side view, xz

Fig. 4.4: Orbits under L_1, the Moon, and L_2 that possess the smallest maximum possible $M_{t,\text{spin}}$ corresponding to a characteristic acceleration of $a_c = 1.70$ mm/s^2.

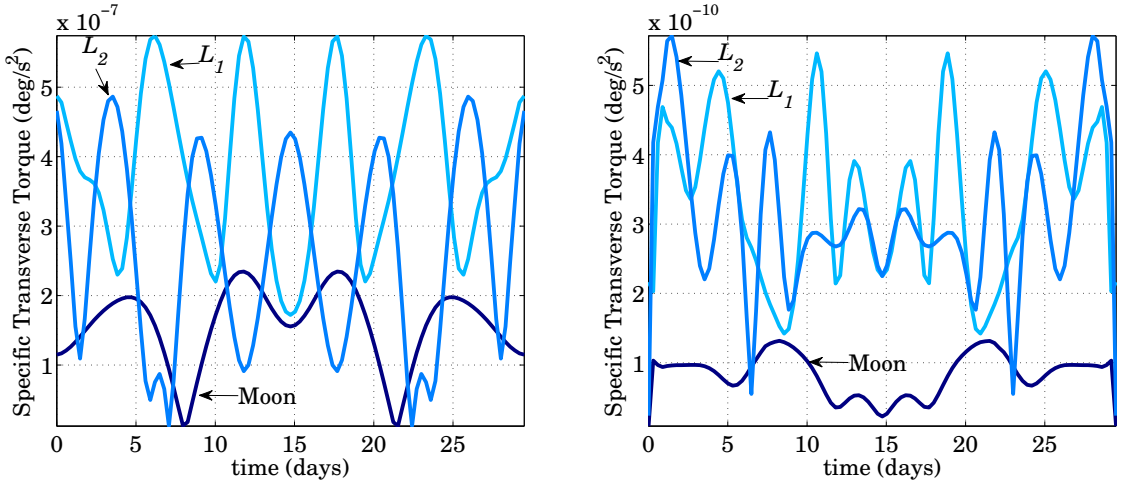
the Lagrange points in Fig. 4.4 are considered prograde orbits as well. Again, these paths are displaced approximately 20000 km below the xy plane.

Critical metrics for the orbits with the smallest maximum $\mathscr{M}_{t,\text{spin}}$ are listed in Table 4.4. Recall that $\mathscr{M}_{t,\text{spin}}$ and $\mathscr{M}_{t,\text{3AS}}$ are independently determined after the orbit is

Table 4.4: Critical metrics for orbits selected by smallest maximum $\mathscr{M}_{t,\text{spin}}$ corresponding to a characteristic acceleration of $a_c = 1.70$ mm/s^2

	Max $\mathscr{M}_{t,\text{spin}}$	Max $\mathscr{M}_{t,\text{3AS}}$	Max α	Min E
L_1	5.74×10^{-7} deg/s^2	5.46×10^{-10} deg/s^2	48.56°	15.01°
Moon	2.35×10^{-7} deg/s^2	1.32×10^{-10} deg/s^2	49.42°	15.07°
L_2	4.86×10^{-7} deg/s^2	5.71×10^{-10} deg/s^2	54.71°	15.30°

generated and are based on the angles α and δ that orient the sail-face normal direction as well as the spin rate of the vehicle. The equation for the body rates appears in Appendix B. The maximum $\mathscr{M}_t$ for a spin-stabilized vehicle and the maximum value for a three-axis stabilized spacecraft configuration do not necessarily occur at the same locations along the orbit, as is apparent in Fig. 4.5. Furthermore, the orbit



(a) $\mathscr{M}_{t,\text{spin}}$ (b) $\mathscr{M}_{t,\text{3AS}}$

Fig. 4.5: Specific transverse torque profiles corresponding to orbits in Fig. 4.4, selected by smallest maximum $\mathscr{M}_{t,\text{spin}}$.

with the smallest maximum $\mathscr{M}_{t,\text{spin}}$ is not necessarily the solution with the smallest maximum $\mathscr{M}_{t,\text{3AS}}$ from the survey. However, the set of solutions below L_1, the Moon, and L_2 that correspond to the smallest maximum $\mathscr{M}_{t,\text{3AS}}$ also resemble the orbits that are associated with the smallest maximum $\mathscr{M}_{t,\text{spin}}$ from the survey. Associated critical metrics when the selection criterion is based on the smallest maximum $\mathscr{M}_{t,\text{3AS}}$ are listed in Table 4.5. The results from the survey indicate that the trajectories below L_2 that are associated with the smallest maximum specific transverse torques ($\mathscr{M}_{t,\text{spin}}$ and $\mathscr{M}_{t,\text{3AS}}$) correspond to a maximum pitch angle along the path with a value greater than 55°. The results listed in Tables 4.4 and 4.5, as well as the associated paths

appearing in Fig. 4.4 and specific transverse torques in Fig. 4.5, correspond to the set of solutions where the maximum pitch angle is restricted to 55°.

Table 4.5: Critical metrics for orbits selected by smallest maximum $\mathscr{M}_{t,\mathrm{3AS}}$ corresponding to a characteristic acceleration of $a_c = 1.70$ mm/s^2

	Max $\mathscr{M}_{t,\mathrm{spin}}$	Max $\mathscr{M}_{t,\mathrm{3AS}}$	Max α	Min E
L_1	6.25×10^{-7} deg/s^2	3.43×10^{-10} deg/s^2	46.44°	15.41°
Moon	2.35×10^{-7} deg/s^2	1.29×10^{-10} deg/s^2	49.36°	15.07°
L_2	5.58×10^{-7} deg/s^2	3.77×10^{-10} deg/s^2	53.82°	15.78°

4.4.3 Pitch angle

As discussed previously, depending on the realistic optical and shape properties of a solar sail, the assumption of a perfectly reflecting, flat, ideal sail diverges from a realistic solar sail model for pitch angles greater than 50° to 60°. Furthermore, sail effectiveness is severely attenuated at high pitch angles [6, 149]. Therefore, it is desirable to examine trajectories below L_1, the Moon, and L_2 that are associated with the smallest maximum pitch angles. The three solutions from the survey associated with the smallest maximum pitch angles for sails with $a_c = 1.70$ mm/s^2 appear in Fig. 4.6. Note that the distance from the xy plane is generally 20000 km or greater for these orbits. The associated pitch angle histories for these three trajectories appear in Fig. 4.7. The corresponding critical metrics for these three paths are listed in Table 4.6.

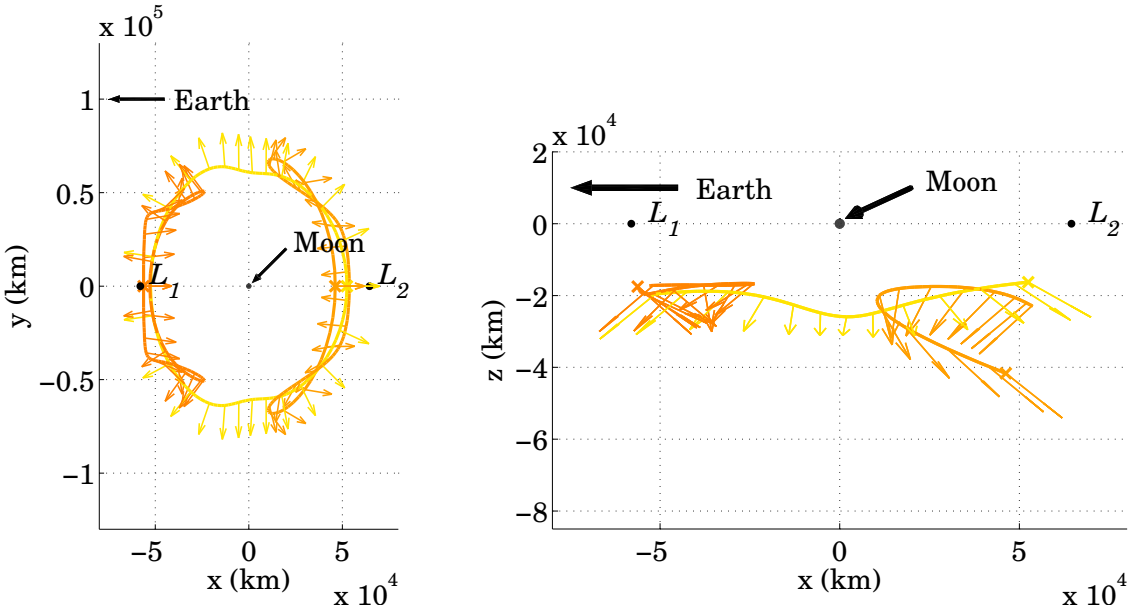
4.5 Survey results pertaining to ground-based critical metrics

If the range of motion or field of view associated with a radio antenna located at the lunar south pole (LSP) is limited, sail trajectories exist that may accommodate

Table 4.6: Critical metrics for orbits selected by smallest maximum pitch angle, α, corresponding to a characteristic acceleration of $a_c = 1.70$ mm/s^2

	Max $\mathscr{M}_{t,\text{spin}}$	Max $\mathscr{M}_{t,\text{3AS}}$	Max α	Min E
L_1	2.09×10^{-6} deg/s^2	5.86×10^{-9} deg/s^2	41.95°	15.00°
Moon	1.59×10^{-6} deg/s^2	1.28×10^{-9} deg/s^2	38.83°	15.00°
L_2	1.93×10^{-6} deg/s^2	1.53×10^{-9} deg/s^2	46.71°	15.00°

this constraint. As mentioned, an elevation-angle constraint of 15° is established in the survey. However, solutions for a sailcraft with a characteristic acceleration value of $a_c = 1.70$ mm/s^2 exist that do not activate this constraint. It is also useful to



(a) Top view, xy

(b) Side view, xz

Fig. 4.6: Sample orbits offset below L_1, the Moon, and L_2 that conform to the 15° elevation-angle constraint and possess the smallest maximum pitch angles at any point along the trajectory.

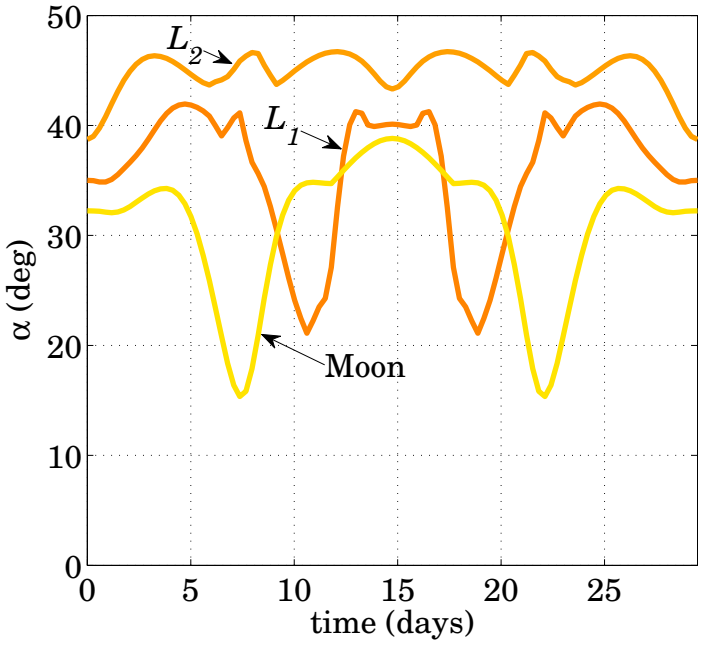


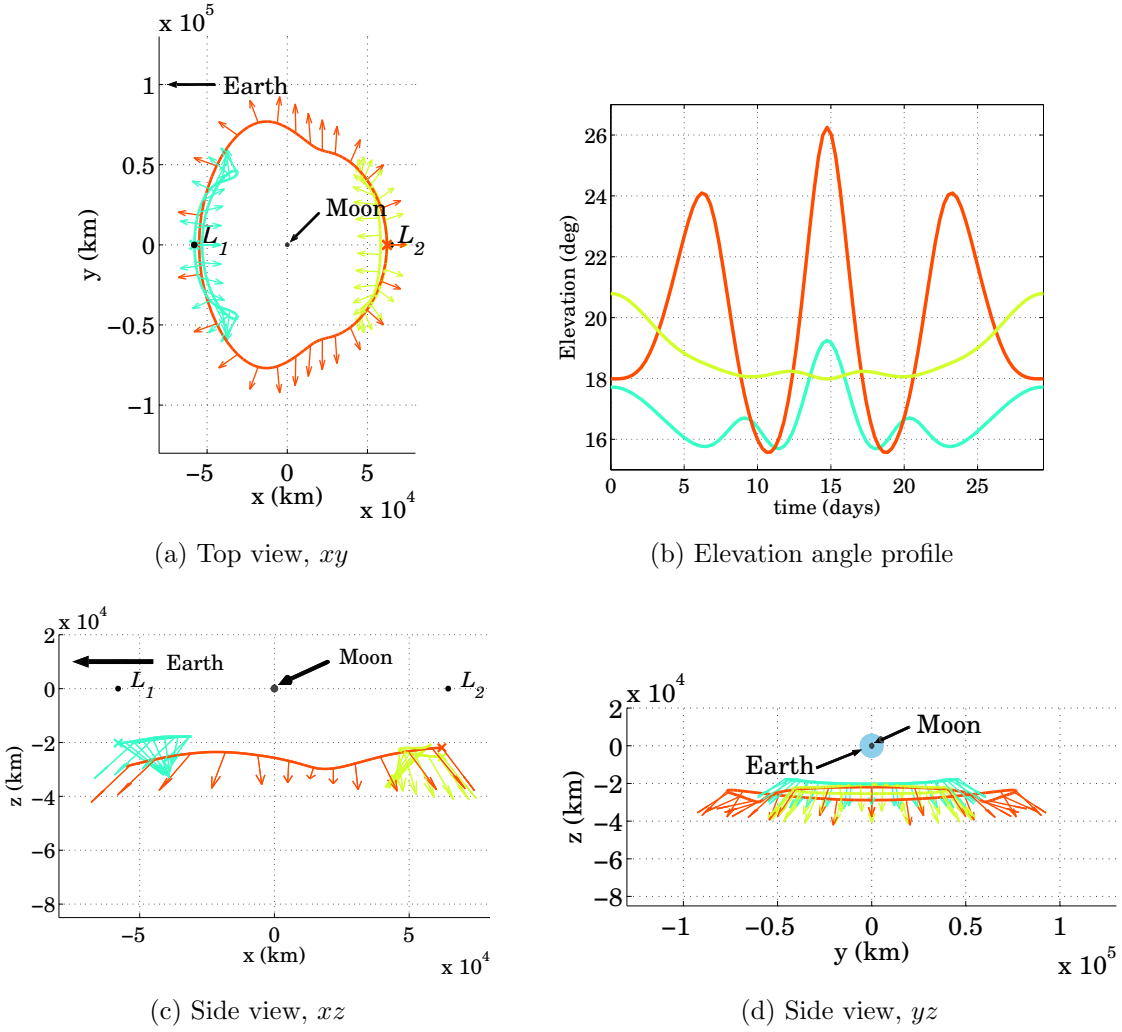
Fig. 4.7: Pitch-angle histories for orbits in Fig. 4.6.

identify the range of altitudes necessary for the orbits associated with a particular sailcraft characteristic acceleration. Finally, the design of any facility at the LSP benefits from information on the field of view. These ground-based critical metrics are examined for a sailcraft with $a_c = 1.70$ mm/s^2.

4.5.1 Elevation angle

The first critical metric for ground-station design is the elevation angle. While a constraint of 15° is imposed in this survey, any solutions not activated by this constraint lend insight into other available options. The three sample orbits below L_1, the Moon, and L_2 that appear in Fig. 4.8 possess the associated critical metrics as listed in Table 4.7.

Unfortunately, little extra margin is available for orbits under L_1 and the Moon. However, solutions do exist below L_2 that possess a minimum elevation angle along



(a) Top view, xy

(b) Elevation angle profile

(c) Side view, xz

(d) Side view, yz

Fig. 4.8: Orbits under L_1, the Moon, and L_2 possessing the largest minimum elevation angle, E.

the path of nearly 18°.[2] The relationship between a_c and E is understandable within the context of a trade-off between a_c and E; that is, a lower a_c is required for trajectories under L_2, as demonstrated in Section 4.4.1. Alternatively, a greater minimum elevation angle is available under L_2 as compared to L_1, or even the Moon, for the same a_c. Displacements from the xy plane for these orbits are generally between

[2]The path under L_2 is selected in post-processing and originates from the set of solutions with $\alpha \in [0, 55°]$. Orbits with larger minimum elevation angles exist that require larger maximum pitch angles along the path.

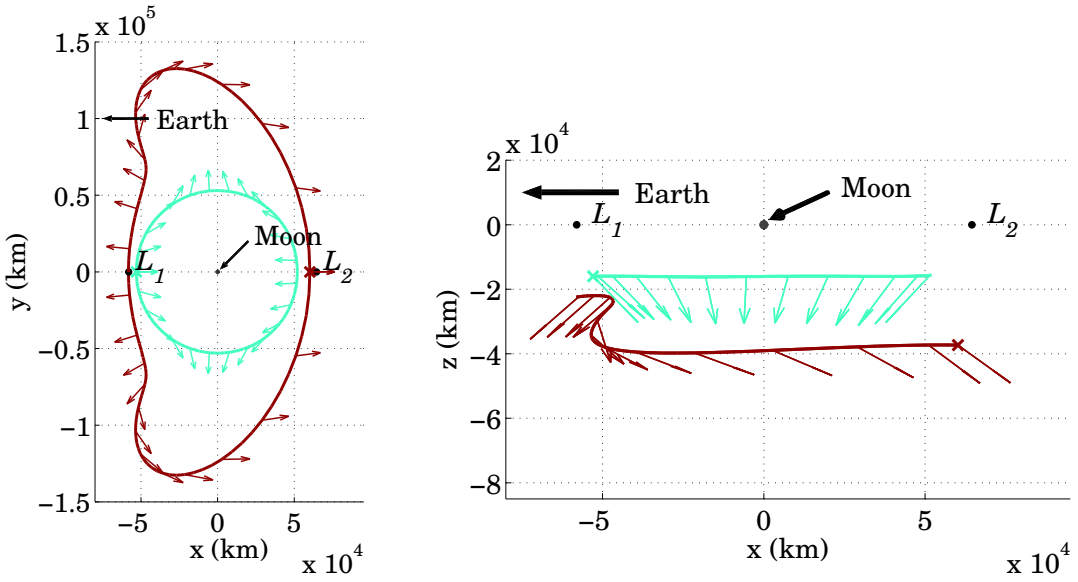
Table 4.7: Critical metrics for orbits selected by largest minimum elevation angle, E, corresponding to a characteristic acceleration of $a_c = 1.70$ mm/s^2

	Max $\mathscr{M}_{t,\text{spin}}$	Max $\mathscr{M}_{t,\text{3AS}}$	Max α	Min E
L_1	0.82×10^{-6} deg/s^2	0.88×10^{-9} deg/s^2	46.55°	15.69°
Moon	2.45×10^{-6} deg/s^2	2.94×10^{-9} deg/s^2	50.55°	15.57°
L_2	1.29×10^{-6} deg/s^2	0.71×10^{-9} deg/s^2	54.62°	17.98°

20000 and 30000 km and are associated with an increase in altitude with respect t the lunar south pole.

4.5.2 Altitude

The distance of a vehicle from a station on the ground drives the power requirements for the transmitting and receiving antennas both on the ground and on the spacecraft. Therefore, it is useful to assess the range of distances from the lunar south pole for a family of solutions when designing a mission and communications system. The altitude of the spacecraft in orbit below the Moon depends on the dynamics of the system, which is based on the gravity of the primaries and the characteristic acceleration, a_c, of the sail. Sailcraft with larger values of a_c may orbit either closer to the Moon or at higher elevation angles. For a characteristic acceleration of 1.70 mm/s^2, the smallest maximum altitude along any trajectory from the survey is 55009 km, while the largest maximum altitude is 141310 km. The aqua path is displaced from the xy plane by less than 20000 km, while the red path is displaced by 20000 to 40000 km. These two paths, both below the Moon, are plotted in Fig. 4.9, and their critical metrics are listed in Table 4.8. The red path originates below L_2 and is a retrograde orbit, while the aqua path originates below L_1 and is a prograde orbit. Both possess a minimum elevation angle of 15.00°. For the red path, this minimum



(a) Top view, xy

(b) Side view, xz

Fig. 4.9: Orbits possessing the largest and smallest altitudes from a base at the LSP.

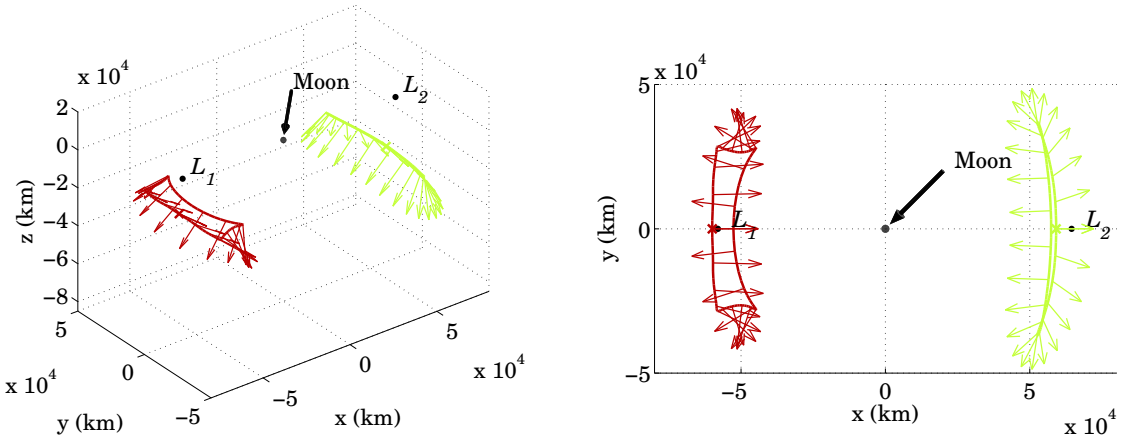
Table 4.8: Critical metrics for orbits selected by altitude corresponding to a characteristic acceleration of $a_c = 1.70$ mm/s^2

	Max $\mathscr{M}_{t,\text{spin}}$	Max $\mathscr{M}_{t,\text{3AS}}$	Max α	Max A
Farthest	4.19×10^{-6} deg/s^2	9.26×10^{-9} deg/s^2	54.97°	141310 km
Nearest	0.85×10^{-6} deg/s^2	0.75×10^{-9} deg/s^2	50.01°	55009 km

elevation angle is encountered when the spacecraft is at the extremities of the orbit in the $\pm y$ direction. Note that the red path in Fig. 4.9 is selected from a set of solutions where the maximum pitch angle at any point along the trajectory is less than 55°.

4.5.3 Azimuth angle ranges

For simplicity, assume that the antenna is fixed on the lunar surface and is always directed toward the sailcraft relay. If the sailcraft is located below L_1, the smallest swath width required of a ground antenna is 64.4°. The smallest swath width associated with a sail below L_2 is 70.2°. The two trajectories associated with these swath widths appear in Fig. 4.10, and their respective critical metrics are listed in Table 4.9.



(a) 3-D view

(b) Top view, xy

Fig. 4.10: Orbits under L_1 and L_2 possessing the smallest ranges of azimuth angles.

Table 4.9: Critical metrics for orbits selected by azimuth angle, Az, range corresponding to a characteristic acceleration of $a_c = 1.70$ mm/s^2

	Max $\mathscr{M}_{t,\text{spin}}$	Max $\mathscr{M}_{t,\text{3AS}}$	Max α	Min E	Az range
L_1	3.74×10^{-6} deg/s^2	9.33×10^{-9} deg/s^2	48.77°	15.00°	64.42°
L_2	1.24×10^{-6} deg/s^2	0.68×10^{-9} deg/s^2	53.99°	16.60°	70.22°

4.6 Discussion of survey results

Some general observations are notable, based on the results from the survey in the Earth–Moon system. First, trajectory options for a spacecraft relay, to support a facility at the lunar south pole, exist for sailcraft with overall characteristic accelerations of 1.3 mm/s^2 or greater. Although solar sailing has only recently been demonstrated in flight, the recent design of the sailcraft for the ST9 ground demonstration delivers an overall characteristic acceleration of 0.58 mm/s^2 (including payload and attitude control system) while the characteristic acceleration supplied solely by the sail and structure is 1.7 mm/s^2 [79]. However, a solar sail alone is not the only option for a single-vehicle relay to deliver LSP coverage. Other researchers propose hybrid propulsion systems that are based on a combination of a solar sail and low-thrust technologies [16, 150–155]. While the survey techniques developed for this analysis are applied to solar sails, the formulation of these techniques is also adaptable for hybrid or other systems where a continual thrust component is available.

Second, if the vehicle is either spin-stabilized or three-axis stabilized, the trajectory options resulting from this survey require attitude control authority within the assumptions employed in other investigations of solar sail attitude dynamics and control. The maximum specific transverse torques for a sailcraft spinning at 5 rotations per hour, which are hundreds of times larger than the maximum specific transverse torques for a three-axis stabilized spacecraft, are less than the 28.1×10^{-6} deg/s^2 assumption published by Wie for a three-axis stabilized spacecraft [141]. To remain in an orbit offset below the Moon, a sailcraft must continually reorient itself. The control profiles for these sample orbits are essentially continuous, as returned by the numerical BVP solver employed to generate the trajectory. In reality, it may be advantageous to command the spacecraft in a "turn-and-hold" scheme, whereby the orientation is held in an inertially fixed attitude for some length of time (e.g., 2–3 days), and then reoriented to a new attitude [156]. In this scenario, it is presumed

that lower specific transverse torques from the continuous case translate to lower specific transverse torques in a turn-and-hold scheme.

Returning to the assumption of a vehicle that continually reorients, it is observed that larger specific torques from the three-axis stabilized attitude scheme are apparently correlated with the larger specific torques from the spin-stabilized attitude scheme. Statistics from the tens of millions of sample orbits generated in this survey suggest a power relationship between the maximum $\mathscr{M}_{t,\text{spin}}$ and $\mathscr{M}_{t,\text{3AS}}$ for the spin rate of 5 rotations per hour employed in this analysis, that is,

$$\mathscr{M}_{t,\text{3AS}} = (\mathscr{M}_{t,\text{spin}})^K \qquad (4.13)$$
$$K = 1.50 \pm 0.075$$

when $\mathscr{M}_{t,\text{spin}}$ and $\mathscr{M}_{t,\text{3AS}}$ are measured in deg/s^2 and the error bounds are three-standard deviations.[3] Recall that the magnitudes of the maximum specific transverse torques are less than one and that a larger specific transverse torque is required to reorient a spinning vehicle.

Finally, some notable observations about the orbits are summarized. Generally, prograde lunar orbits are "flatter" and evolve at smaller distances relative to the lunar south pole when compared to retrograde lunar orbits. However, retrograde orbits with variable elevation angles may be attractive for some mission-specific considerations. Trajectories below L_2 require smaller sail characteristic accelerations when compared to trajectories located under L_1 and the Moon. As demonstrated in Section 4.1.1, trajectories exist that naturally transfer from a Moon-centered path to a path located below a Lagrange point. Because sail optical properties degrade over time [157], a sailcraft orbiting below the Moon can be shifted to an orbit under L_2 as a_c decreases. All of the orbits surveyed are unstable because of their dependencies on a solar sail. Furthermore, all trajectories require some form of flight-path control. Controllability of the flight path is not necessarily related to any particular critical metric employed in this analysis and is a subject for future investigation.

[3] $K = 1.39 \pm 0.06$ when $\mathscr{M}_{t,\text{spin}}$ and $\mathscr{M}_{t,\text{3AS}}$ are measured in rad/s^2.

5. INSTANTANEOUS EQUILIBRIUM SURFACES

"The most celebrated of all dynamical problems," as described by Whittaker, is the three-body problem [158, 159]. After Newton developed a mathematical formulation for motion within a gravity field consisting of multiple bodies, he [160] and others [159] investigated the coupled motion of the Moon, the Earth, and the Sun. In 1772, Euler first proposed the restricted assumption of the three-body problem, in which the mass of one body is considered infinitesimal compared to those of the other two bodies, and formulated the problem in a rotating reference frame [159].

Contemporaneously, Euler and Lagrange discovered multiple equilibrium solutions to the circular restricted three-body (CR3B) problem in a frame attached to, and rotating with, the two larger primaries; the motion of the primaries traces out concentric circles in inertial space. If a stationary particle is located at one of these equilibrium points, it theoretically remains there indefinitely unless perturbed. Euler identified the collinear solutions, while Lagrange determined the locations for both the collinear and the equilateral points [159]. The discovery of periodic solutions in the CR3B problem motivated numerous advances in mathematics and dynamical analysis since the problem's formulation. Laplace, Jacobi, Gyldén, Lindstedt, Hill, Poincaré, Darwin (the son of naturalist Charles Darwin), Levi-Civita, Sundman, Birkhoff, Komolgorov, Arnold, and Moser all figure prominently in the development of the understanding of motion in the three-body system [82, 159]. Poincaré merits special recognition for using the CR3B system to introduce the study of chaotic phenomena and for his conjecture that an infinite number of periodic solutions exist for the CR3B problem [159].

The equilibrium points are often a starting point for investigations of motion within a nonlinear system, such as the CR3B regime. Both periodic and quasi-periodic orbits in the vicinity the Lagrange points and families of such orbits are the subject of various numerous investigations. A "family" is a collection of orbits

that possess similar dynamical properties and are often generated via a continuation technique in which a previous solution initializes the search for a subsequent orbit. A search for these families is often initialized by analytical approximations of an orbit in the vicinity of a Lagrange point. Analytical approaches for generating periodic and quasi-periodic orbits are discussed and developed by Moulton [161], Szebehely [82], Farquhar and Kamel [162], and Richardson and Cary [163], among many others. Computing power has enabled an array of numerical techniques for generating orbits and their associated families near the Lagrange points as well. A full catalog of the development of analytical tools for generating orbits in the vicinity of the Lagrange points is beyond the scope of this document. However, a few examples are worth mentioning: tools developed for a numerical investigation by Breakwell and Brown [164] are employed to generate and study families of halo orbits near the collinear Lagrange points [165, 166]. Such investigations have led to numerous subsequent studies of periodic and quasi-periodic motion in the CR3B problem [167]. A more thorough discussion of families of orbits in the vicinity of the collinear libration points, along with the history of the numerical developments for generating those orbits, is given by Howell [168], and a selected bibliography of libration point missions is summarized by Farquhar [169].

Investigations into mission design within the context of a restricted four-body problem also start with equilibrium solutions [170–178]. However, with the incorporation of another massive body (e.g., the Sun in the Earth–Moon system or the Moon in the Sun–Earth system), the equilibrium solutions are no longer stationary, and their locations must be calculated for each instantaneous time epoch. Effectively, the five equilibrium solutions from the CR3B model evolve into an infinite set of instantaneous equilibrium solutions with the inclusion of a moving third massive body.

The addition of a solar sail introduces another force into the system, and the sail can be positioned and oriented differently, yielding an infinite number of instantaneous equilibrium solutions. In the context of the CR3B regime, these instantaneous

equilibrium solutions appear as surfaces. Initial development of these surfaces occurs as recently as the early 1990s [6]. Any single artificial Lagrange point depends on the characteristic acceleration and orientation of the sail. Just as periodic solutions are available in the vicinity of the five classical Lagrange points, solutions in the vicinity of these surfaces are also available and are the basis for investigations of solar sail trajectories in the Sun–Earth, the Earth–Moon, and a Sun–asteroid CR3B regime [3, 4, 7, 9, 13, 44, 45, 117, 118, 153, 179–181]. In the case of the Earth–Moon system, the Sun moves in a clockwise direction as viewed from above and the equilibrium surfaces are instantaneous. After a brief introduction to the mathematical determination of the Lagrange points in the CR3B system, observations regarding the relationship of solar sail trajectories to instantaneous equilibrium surfaces are offered.

5.1 Equilibrium in circular restricted systems

While a general solution to the CR3B problem is not easily accessible, particular solutions (i.e., the equilibrium points) are straightforward. Extending the search for instantaneous equilibrium solutions in a restricted four-body model is more difficult, but remains tractable. Developing such solutions is useful in undertaking the more challenging task of generating instantaneous equilibrium surfaces for solar sail spacecraft.

5.1.1 Equilibrium in the circular restricted three-body system

For clarity, some fundamental definitions from the CR3B problem are summarized. The locations of the equilibrium points in the CR3B model were known to Euler and Lagrange over 200 years ago. The collinear points are labeled L_{1-3} and the equilateral points are labeled $L_{4,5}$ in Fig. 5.1. In the figure, the larger primary is P_1, the smaller is P_2, and the barycenter, B, establishes the origin of the system. The nondimensional distance from B to P_1 is μ, where μ is the ratio of the mass of P_2 to the combined masses of P_1 and P_2, as developed in Eq. (2.5). The distance from B to P_2, as well as

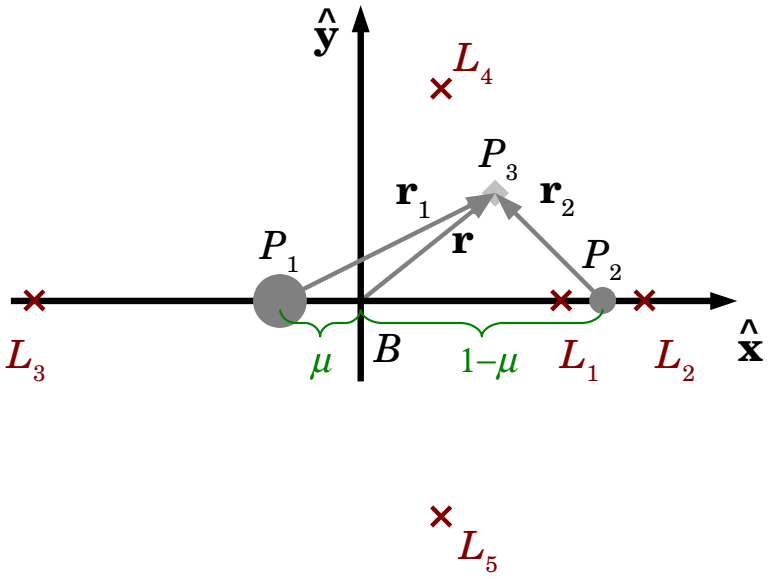


Fig. 5.1: Locations of the five Lagrange points in the CR3B problem.

the nondimensional mass of P_1, is then $1-\mu$. For completeness, the vectors from P_1 to the infinitesimal mass at P_3, $\mathbf{r}_1$, from B to P_3, $\mathbf{r}$, and from P_2 to P_3, $\mathbf{r}_2$, are also included in the figure. The collinear equilibrium point between the primaries is labeled L_1 (this point in often denoted the "cis-lunar" Lagrange point when referring to the Earth–Moon system), while the exterior collinear points are L_2 ("trans-lunar" in the Earth–Moon system) and L_3. The system is assumed to rotate counter-clockwise, and L_4 leads the rotation while L_5 lags. These equilibrium solutions are often termed the Lagrange or libration points.

The equations of motion in the CR3B problem appear in Eqs. (2.13)–(2.15), or $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ in first-order form. Equilibrium solutions are always determined from $\mathbf{f}(\mathbf{x}) = \mathbf{0}$, thus, these locations are computed by setting the acceleration and velocity terms in Eqs. (2.13)–(2.15) to zero. Note that no additional acceleration terms (e.g., $\mathbf{a}_s(t)$) are applied. Only centripetal and gravitational acceleration terms remain, and, essentially, equilibrium occurs when these two sets of accelerations balance in

a rotating coordinate system. Assuming a nondimensional system (i.e., $\omega = 1$), the resulting algebraic relationships are

$$-x^e + \frac{(1-\mu)(x^e+\mu)}{r_1^3} + \frac{\mu(x^e-(1-\mu))}{r_2^3} = 0 \tag{5.1}$$

$$-y^e + \frac{(1-\mu)y^e}{r_1^3} + \frac{\mu y^e}{r_2^3} = 0 \tag{5.2}$$

$$\frac{(1-\mu)z^e}{r_1^3} + \frac{\mu z^e}{r_2^3} = 0 \tag{5.3}$$

where $\left\{ \begin{array}{ccc} x & y & z \end{array} \right\}^{\mathrm{T}}$ constitute $\mathbf{r}$ and the superscript " ^{e} " indicates an equilibrium solution. It is immediately apparent that all equilibrium solutions require that $z^e = 0$. From this fact, two types of solutions emerge.

The first type of equilibrium solution consists of the collinear Lagrange points, L_{1-3}. For these three points $y^e = 0$. As a result, $r_1 = (x^e - \mu)$ and $r_2 = (x^e - (1-\mu))$. Consequently, Eq. (5.1) becomes

$$-x^e + \frac{\mu_1(x^e+\mu_2)}{|x^e-\mu_2|^3} + \frac{\mu_2(x^e-(1-\mu_2))}{|x^e-(1-\mu_2)|^3} = 0 \tag{5.4}$$

where $\mu_1 = 1-\mu$, the mass of P_1, and $\mu_2 = \mu$, the mass of P_2. Note the inclusion of absolute values in the denominators of the terms in Eq. (5.4). Iterative techniques are used to establish the value of the real root in the region of interest quickly. An alternative approach involves solving the roots of a quintic polynomial and selecting only the real, positive roots [82]. The approximate distances of the x coordinates for the L_1 and L_2 points from the smaller primary, P_2, are listed in Table 5.1 for various systems. Corresponding nondimensional values normalized by the distance from P_1 to P_2 are also listed in the table. The Sun–Earth/Moon system assumes that the Sun is P_1 and the Earth and the Moon are a composite body, P_2. A CR3B model can be constructed from any combination of two massive primaries that move in concentric circular orbits. While not a standard configuration, the Sun–Moon system also meets this criteria. Realistically, in a more complete model, any particle at a Sun–Moon Lagrange point will quickly depart.

In the second type of equilibrium solution, $y \neq 0$. It is clear that Eq. (5.2) is satisfied when $r_1 = r_2 = 1$ in nondimensional units. Because the distance from P_1 to

Table 5.1: Approximate distances from P_2 to L_1 and L_2

	Earth–Moon	Sun–Earth	Sun–Earth/Moon	Sun–Moon
L_1 ($\times 10^3$ km)	58.0	1491.6	1497.6	345.2
L_2 ($\times 10^3$ km)	64.5	1501.5	1507.7	345.7
L_1 (nondim)	0.151	0.00997	0.00997	0.00231
L_2 (nondim)	0.168	0.01004	0.01004	0.00231

P_2 is also equal to 1, these equilibrium points are labeled the equilateral or triangular Lagrange or libration points and are located at $L_{4,5}$. Using simple trigonometry, it is easy to establish that $y^e = +\sqrt{3}/2$ for L_4, that $y^e = -\sqrt{3}/2$ for L_5, and that $x^e = \mu - 1/2$ for both points [82].

5.1.2 The bi-circular restricted four-body system

Many forms of a restricted four-body model are available. In general, these various formulations are classified as coherent and non-coherent. For the former, the locations of the three massive primaries conform to Newton's law of gravitation and are often represented via a series approximation [173, 182]. For the latter, the locations of the bodies do not satisfy Newton's law of gravitation, but do provide an approximation for the locations of the massive primaries that is sufficient to yield useful results. For the present discussion, a non-coherent bi-circular restricted four-body (BCR4B) system is adequate.

The BCR4B model, like the CR3B model, assumes that two primaries are fixed in a reference frame that rotates with respect to inertial space [173, 182, 183]. A third, massless particle is free to move within the system. In the BCR4B regime, a fourth massive primary moves in a circular fashion in a frame fixed to the first two primaries and rotating with respect to inertial space. A standard representation of the BCR4B

model appears in Fig. 5.2. In the model, a rotating reference frame, R, comprising

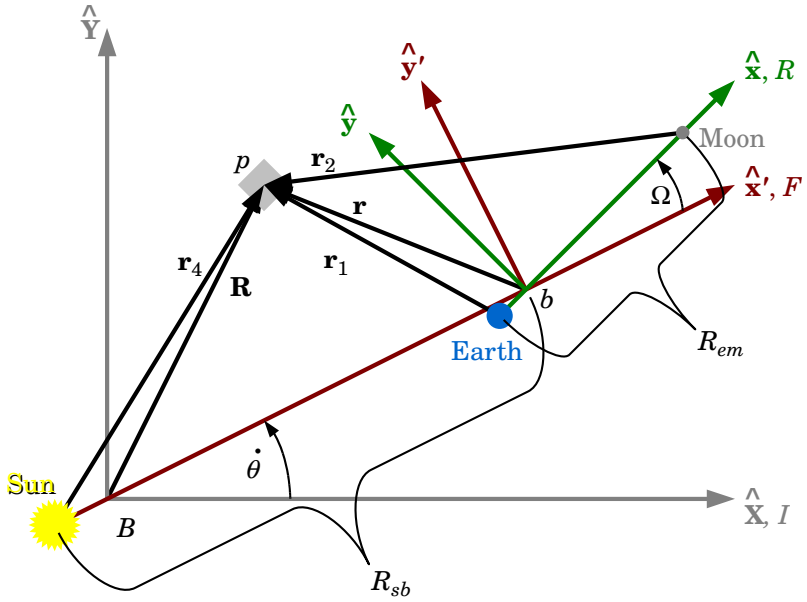


Fig. 5.2: Reference frames for the BCR4B model.

$\left\{ \begin{array}{ccc} \hat{\mathbf{x}} & \hat{\mathbf{y}} & \hat{\mathbf{z}} \end{array} \right\}^{\mathrm{T}}$, is attached to the Earth and the Moon, and rotates about the Earth–Moon barycenter, b. A second frame is also defined with an origin at point b and rotates with the Sun about the system barycenter B; this frame is labeled F and comprises $\left\{ \begin{array}{ccc} \hat{\mathbf{x}}' & \hat{\mathbf{y}}' & \hat{\mathbf{z}}' \end{array} \right\}^{\mathrm{T}}$. The R frame completes one rotation with respect to the F frame each synodic month (29.5 days) and one rotation with respect to the inertial frame, I, each sidereal month (27.3 days). The rotation rate of R with respect to F is labeled Ω, and the rotation rate of F with respect to I is labeled $\dot{\theta}$; both rates are assumed constant. For consistency, the vector from the Sun to the particle is labeled $\mathbf{r}_4$ in this and all subsequent models. The vectors from the Earth to the particle and

the Moon to the particle are $\mathbf{r}_1$ and $\mathbf{r_2}$, respectively. Following Guzmán [182], the equations of motion for a particle, p, with respect to point B in the F frame are

$$ {}^F\ddot{x}' - \ddot{\theta}y' - 2\dot{\theta}({}^F\dot{y}') + \mathscr{U}_{x'} = f_{x'} \tag{5.5} $$

$$ {}^F\ddot{y}' + \ddot{\theta}x' + 2\dot{\theta}({}^F\dot{x}') + \mathscr{U}_{y'} = f_{y'} \tag{5.6} $$

$$ {}^F\ddot{z}' + \mathscr{U}_{z'} = f_{z'} \tag{5.7} $$

where an applied force, $\mathbf{f}$, equals $\mathbf{0}$, unless otherwise noted. Recall that the left superscript F indicates a derivative in that rotating frame. The position vector of the particle, as expressed in terms of coordinate frame F, is $\mathbf{R} = \left\{ \begin{array}{ccc} x' & y' & z' \end{array} \right\}^{\mathrm{T}}$. In the bi-circular model, $\ddot{\theta} = 0$. A pseudo-potential term $\mathscr{U}$ possesses the gradients

$$ \begin{aligned} \mathscr{U}_{x'} = -\dot{\theta}^2 x' + \frac{\mu_s}{r_4^3}\left(x' + \frac{1}{1+\mu_s}\right) + \frac{1-\mu}{r_1^3}\left(x' - \frac{\mu_s}{1+\mu_s} + \mu\cos\Omega t\right) \\ + \frac{\mu}{r_2^3}\left(x - \frac{\mu_s}{1+\mu_s} - (1-\mu)\cos\Omega t\right) \end{aligned} \tag{5.8} $$

$$ \mathscr{U}_{y'} = -\dot{\theta}^2 y' + \frac{\mu_s}{r_4^3}y' + \frac{1-\mu}{r_1^3}\left(y' + \mu\sin\Omega t\right) + \frac{\mu}{r_2^3}\left(y' - (1-\mu)\sin\Omega t\right) \tag{5.9} $$

$$ \mathscr{U}_{z'} = \left(\frac{\mu_s}{r_4^3} + \frac{1-\mu}{r_1^3} + \frac{\mu}{r_2^3}\right) z' \tag{5.10} $$

System parameters are nondimensionalized such that the Earth–Moon distance is $R_{em} = 1$, the rotation rate of the R frame with respect to the I frame is 1, and the sum of the gravitational parameters of the Earth $(1-\mu)$ and the Moon (μ) is also 1. Consequently, $\dot{\theta} = 1 - \Omega$, the period of one revolution of the R frame with respect to the I frame is 2π, the distance from the Sun to the Earth–Moon barycenter is $R_{Sb} = 389.17$ nondimensional distance units, or 1 AU (astronomical unit), and the gravitational parameter of the Sun is $\mu_s = 328900.6$ nondimensional units.

To explore the effect of the Sun on the Earth–Moon $L_{1,2}$ points, the BCR4B model serves as a framework. For this application, the equations of motion are formulated in the R frame, where the Earth and the Moon are fixed and Sun moves in a circle, clockwise about the system, at a constant rate Ω, as illustrated in Fig. 5.3. The Sun is assumed to be sufficiently far from the Earth–Moon system such that $-\mathbf{r}_4$ is effectively parallel to $\mathbf{r}_{bB}$ and $\mathbf{r}_{bS}$. The vector equation of motion is similar to Eq. (2.6), but

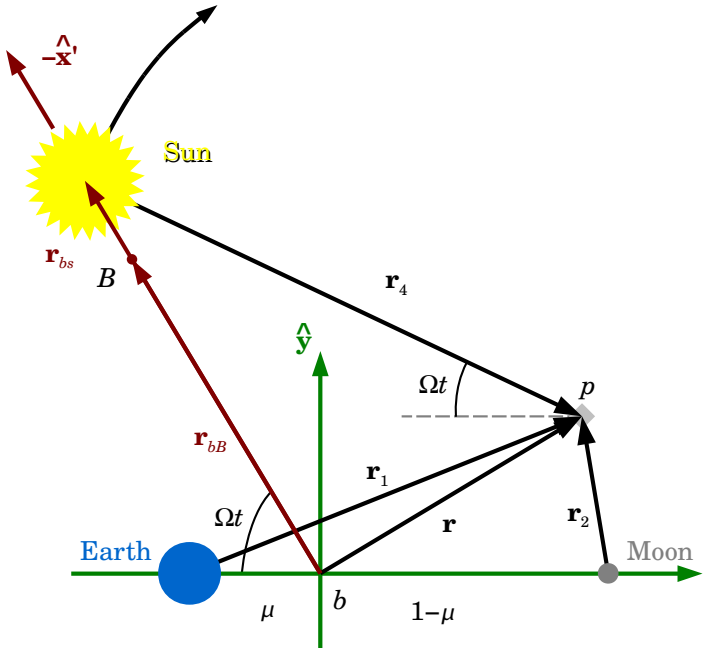


Fig. 5.3: Bi-circular restricted four-body model in the Earth–Moon rotating frame.

includes an additional term to accommodate the gravitational attraction from the Sun, that is,

$$ {}^{R}\mathbf{a} + 2\left({}^{I}\boldsymbol{\omega}^{R} \times {}^{R}\mathbf{v}\right) + \boldsymbol{\nabla} U(\mathbf{r}) + \mathbf{a}_4 = \mathbf{f} \tag{5.11} $$

where $\boldsymbol{\nabla} U$ is defined in Eq. (2.7), the applied force $\mathbf{f} = \mathbf{0}$, unless otherwise noted, ${}^{R}\mathbf{a} = \left\{ \begin{array}{ccc} {}^{R}\ddot{x} & {}^{R}\ddot{y} & {}^{R}\ddot{z} \end{array} \right\}^{\mathrm{T}}$, ${}^{R}\mathbf{v} = \left\{ \begin{array}{ccc} {}^{R}\dot{x} & {}^{R}\dot{y} & {}^{R}\dot{z} \end{array} \right\}^{\mathrm{T}}$, and the gravitational contribution from the Sun is

$$ \mathbf{a}_4 = \mu_s \left(\left(\frac{x_4}{r_4^3} + \frac{r_{bB_x}}{r_{bB}^3} \right) \hat{\mathbf{x}} + \left(\frac{y_4}{r_4^3} + \frac{r_{bB_y}}{r_{bB}^3} \right) \hat{\mathbf{y}} + \left(\frac{z_4}{r_4^3} + \frac{r_{bB_z}}{r_{bB}^3} \right) \hat{\mathbf{z}} \right) \tag{5.12} $$

Note that motion described by the vector differential equation in Eq. (5.11) is referenced to b. In the CR3B model, b is fixed. In the BCR4B model, b moves with respect to the inertial frame. The pair of terms in each component of in Eq. (5.12) (for example, $\frac{x_4}{r_4^3} + \frac{r_{bB_x}}{r_{bB}^3}$) incorporates the effect of solar gravity on a particle moving about b. The first term in each pair is the *direct* effect of solar gravity on the particle and the second is the *indirect* effect, because the Sun affects the motion of b. This

formulation is analogous to the familiar relative equation of two-body motion for the movement of a particle about a central body and where the effects of additional bodies are accommodated as perturbations [83, 85, 184]. In Eq. (5.12),

$$\mathbf{r}_{bB} = \frac{\mu_s}{1+\mu_s}\mathbf{r}_{bS} \tag{5.13}$$

$$\begin{aligned}\mathbf{r}_4 &= \mathbf{r} - \mathbf{r}_{bS} \\ &= (x + \mathrm{AU}\cos\Omega t)\,\hat{\mathbf{x}} + (y - \mathrm{AU}\sin\Omega t)\,\hat{\mathbf{y}} + z\hat{\mathbf{z}}\end{aligned} \tag{5.14}$$

$$\mathbf{r}_{bs} = -\mathrm{AU}\cos\Omega t\hat{\mathbf{x}} + \mathrm{AU}\sin\Omega t\hat{\mathbf{y}} + 0\hat{\mathbf{z}} \tag{5.15}$$

where AU is one astronomical unit, approximately 149.6 million kilometers.

To determine instantaneous equilibrium points, and consistent with the CR3B system, the acceleration and velocity terms, $\mathbf{a}$ and $\mathbf{v}$, are set to zero in Eq. (5.11) to determine equilibrium solutions in the BCR4B system. Employing the locations of the Lagrange points from the CR3B model as initial guesses, the instantaneous Lagrange points in the BCR4B model are determined via an iterative procedure. Note that, unlike in the CR3B model, a particle positioned at an instantaneous equilibrium solution will not remain stationary, and, as such, that position is not a true equilibrium solution. However, these solutions are labeled "equilibria" for purposes of this analysis. As expected, the solar gravity perturbs the location of the Lagrange points. The equilibrium solutions for Eq. (5.11) are functions of the location of the Sun as it completes its cycle about the Earth and the Moon. The instantaneous locations of the Lagrange points in the vicinity of the Moon as traced by the blue path in Fig. 5.4 are particularly relevant. The path for the instantaneous equilibrium solutions completes two cycles for every one cycle of the Sun about the system. When $t = 0$ and the Sun is positioned in the $-\hat{\mathbf{x}}$ direction, the perturbed equilibrium solutions are to the left of $L_{1,2}$, as indicated by the red 'x' in each plot. As the Sun moves clockwise, the instantaneous equilibrium solution shifts counter-clockwise, as indicated by the black arrow in Fig. 5.4. As the Sun completes a quarter cycle ($\Omega t = 90^\circ$ and is positioned in the $+\hat{\mathbf{y}}$ direction), the solutions are on the $\hat{\mathbf{x}}$ axis to the right of the $L_{1,2}$ points. When the Sun completes one half revolution and is aligned on the positive $\hat{\mathbf{x}}$

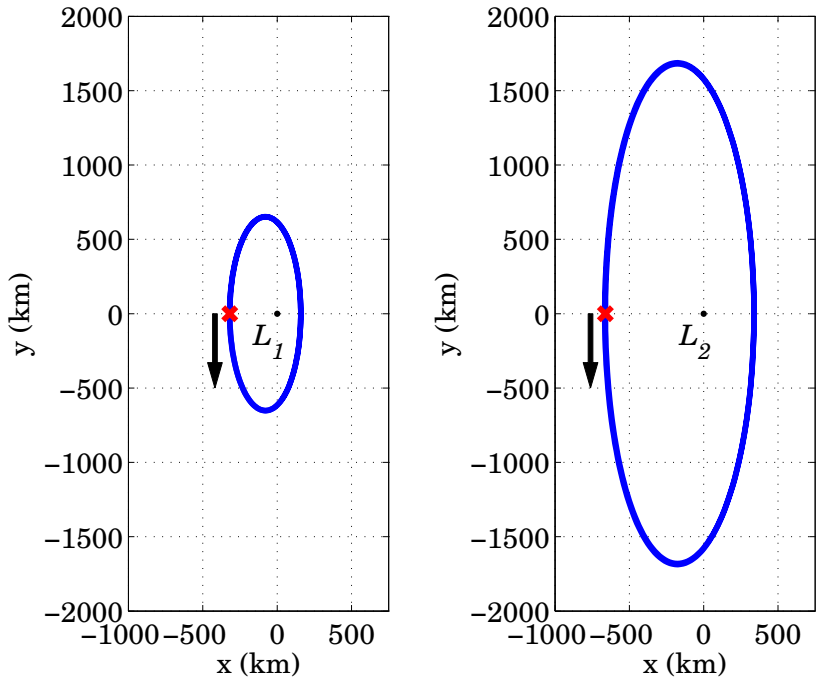


Fig. 5.4: Paths of instantaneous equilibria for the Earth–Moon $L_{1,2}$ points as perturbed by solar gravity centered on the $L_{1,2}$ locations from CR3B model. Note that the Sun moves in a clockwise direction in this frame.

axis ($\Omega t = 180°$), the instantaneous equilibrium solutions are nearly at their original locations on the $-\hat{\mathbf{x}}$ side of the path. A zoom view of the pertinent regions appears in Fig. 5.5. The location of the instantaneous solution at $t = 0$ is indicated by a red 'x' in each plot. Overall, the distance from each respective Lagrange point, L_1 and L_2, to the instantaneous equilibrium solutions in the perturbed model are less than 3% of the relative L_1 and L_2 distances to the Moon (as listed in Table 5.1).

This double cycle occurs because of the interaction of the direct and indirect acceleration terms from solar gravity as noted in Eq. (5.12). In the Earth–Moon system, these terms are nearly equal and opposite in sign for solar gravity, and the net acceleration from Eq. (5.12) due to solar gravity is approximately two orders of magnitude smaller than the acceleration due to the gravity of the Earth or the Moon on a particle in the vicinity of the Lagrange points, as listed in Table 5.2.

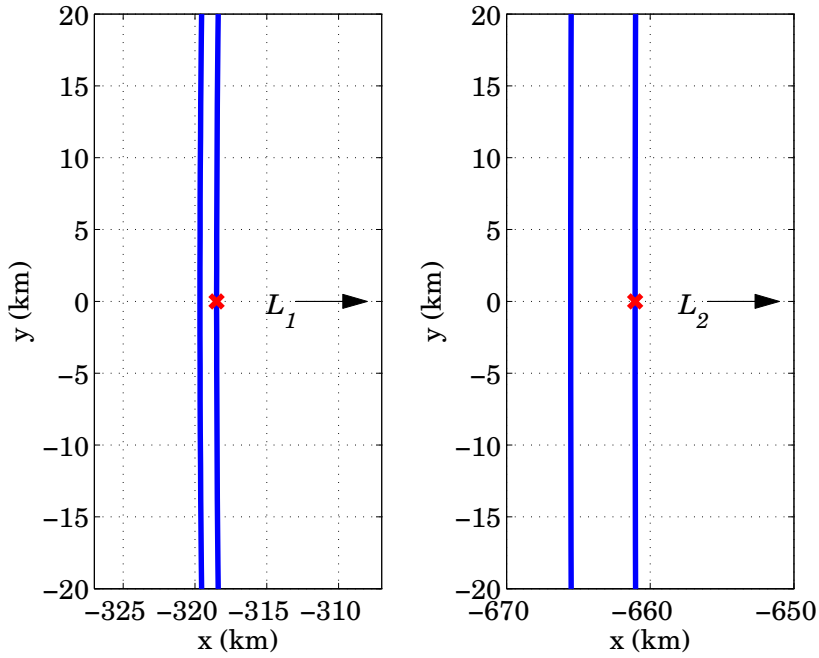


Fig. 5.5: Zoomed view of the paths of instantaneous equilibria for the Earth–Moon $L_{1,2}$ points as perturbed by solar gravity to compare the solar-angle cases of 0° and 180°.

Table 5.2: Comparison of gravitational accelerations due to the Earth, Moon, and Sun in Eq. (5.11) on a particle at the Earth–Moon CR3B $L_{1,2}$ points

	L_1	L_2
Earth	3.7419 mm/s^2	1.9779 mm/s^2
Moon	1.4565 mm/s^2	1.1779 mm/s^2
Sun (average)	0.0203 mm/s^2	0.0280 mm/s^2

Because of the relative formulation for the solar gravity in Eq. (5.12), the paths for the instantaneous equilibrium solutions in Fig. 5.4 are counter-clockwise when the motion of the Sun in this Earth–Moon frame is clockwise. From this perturbation analysis, it is concluded that trajectories developed in the vicinity of the Moon in the

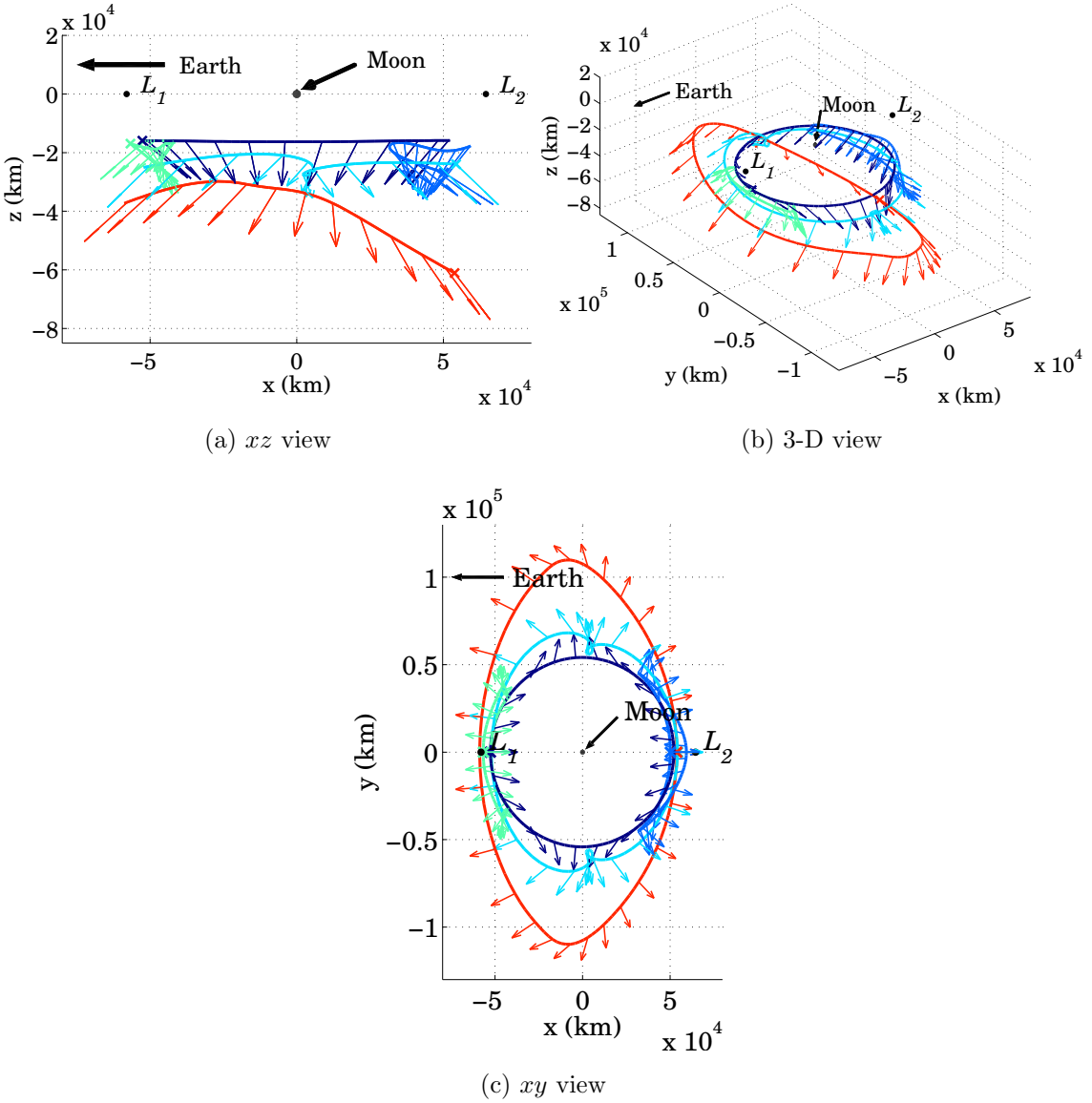
Earth–Moon CR3B model are not significantly affected by the gravitational attraction of the Sun, and, therefore, it is reasonable to neglect solar gravity in this regime when developing solutions of lower fidelity, including scenarios that involve a solar sail.

5.2 Sample reference trajectories

Five sample reference paths generated via finite-difference methods are selected to explore the relationship between instantaneous equilibrium surfaces and solar sail trajectories. These paths are generated within the context of the CR3B model in reference frame R and appear in Fig. 5.6. Four of the five reference trajectories are from Fig. 3.10, while a fifth from Fig. 4.4 is included to illustrate a sample counter-clockwise orbit. In Fig. 5.6, the Moon appears to scale and the two Earth–Moon Lagrange points near the Moon are included for reference. The initial state of the spacecraft corresponding to each trajectory is defined to occur when the Sun is along the $-x$ axis, per Eq. (2.11) and is marked with an '×' in Fig. 5.6. In the royal-blue orbit below L_2, the '×' is located at the southern-most, or minimum z, position along the orbit. The cyan and red-orange orbits both originate in the region below L_2, while the aqua and dark-blue orbits originate below L_1. The initial motion for all five trajectories is in the $-y$ direction. The arrows along the trajectories represent the direction of the sail-face normal, $\hat{\mathbf{u}}$, at a given epoch and form positive dot products with the sunlight vector, $\hat{\boldsymbol{\ell}}(t)$. For convenience, these five orbits are assigned numeric identifiers that appear in Table 5.3. The general vicinity of the orbit, along with the

Table 5.3: Identification numbers for the five sample trajectories

ID No.	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
Location	L_1	L_2	Moon	Moon	Moon
Rotation	CW	CW	CCW	CCW	CW



(a) xz view

(b) 3-D view

(c) xy view

Fig. 5.6: Five sample reference trajectories below the Moon.

direction of revolution (clockwise, CW, or counter-clockwise, CCW), is listed in the table as well.

Each trajectory in Fig. 5.6 requires a characteristic acceleration of 1.70 mm/s^2. From recent explorations of the design space, it is apparent that a net characteristic acceleration of approximately 1.50 to 1.70 mm/s^2 is required for an ideal sail

to maintain the 15° elevation-angle constraint in the LSP coverage problem consistently [30, 35, 129]. Only two spacecraft to date, IKAROS and NanoSail-D2, have achieved orbit and employed a solar sail as their sole source of propulsion [67, 71]. However, as previously noted, a sailcraft that possesses a net characteristic acceleration of 0.58 mm/s^2 was designed and ground tested by L'Garde for the New Millennium Program's Space Technology 9 competition at NASA [79]; the characteristic acceleration of the L'Garde sail and structure alone is 1.70 mm/s^2. In reality, solar sailing technology will likely evolve to complement some other type of propulsive device and leverage a hybrid propulsion system [185]. Nevertheless, if the LSP coverage problem is addressed solely with a sailcraft, some sail technology advancements are required to deliver an appropriate characteristic acceleration.

The sailcraft orbits in Fig. 5.6 include minimum elevation angles as viewed from the LSP of slightly more than 15°, generally near the extreme y values along each orbit. Elevation-angle histories as a function of time along each reference trajectory appear in Fig. 5.7. The dark-blue orbit possesses the most consistent elevation-angle

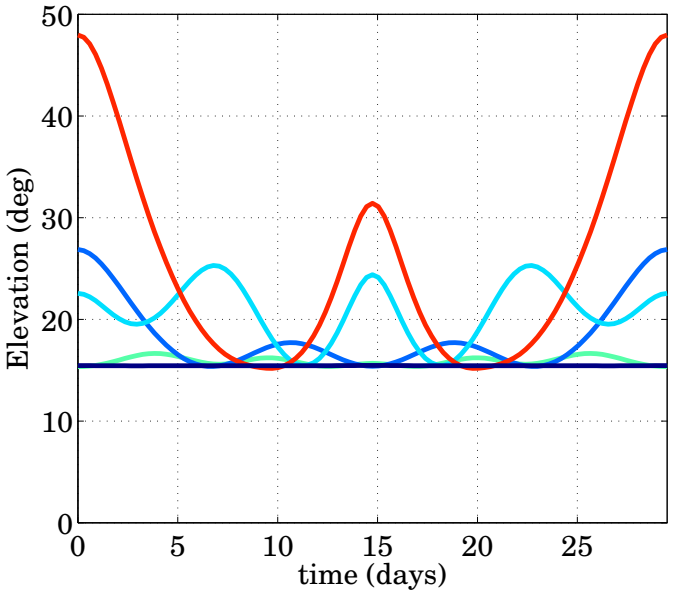


Fig. 5.7: Elevation history for the five reference orbits.

profile, ranging from 15.43° and 15.46°, while the red-orange orbit's elevation angles range from 15.18° to 48°.

Application of the FDM algorithm in the CR3B system results in a control profile expressed in terms of unit vectors defined in the rotating Earth–Moon frame, R. These vectors are decomposed into pitch and clock angles, α and δ, respectively, with respect to the Sun-line. Recall that

$$\hat{\boldsymbol{\ell}}(t) = \cos\Omega t\hat{\mathbf{x}} - \sin\Omega t\hat{\mathbf{y}} + 0\hat{\mathbf{z}} \tag{5.16}$$

in the R frame. As illustrated in Fig. 2.3 from Section 2.1, the pitch and clock angles are evaluated as

$$\alpha = \cos^{-1}(\hat{\boldsymbol{\ell}}(t) \cdot \hat{\mathbf{u}}) \tag{5.17}$$

$$\delta = \tan^{-1}\left(\frac{\ell_1 u_2 - \ell_2 u_1}{u_3}\right) \tag{5.18}$$

where $\tan^{-1}$ is the four-quadrant arctangent function, $\ell_{1,2}$ are the first two elements of the Sun-line vector, $\hat{\boldsymbol{\ell}}(t)$, and u_{1-3} are the three components of $\hat{\mathbf{u}}$. The pitch and clock angles for the five sample orbits from Fig. 5.6 appear in Fig. 5.8. Note that δ is periodic and is equal to 180° at each half cycle of the orbit, or at the times when the Sun lies along the $\hat{\mathbf{x}}$ axis, forming a solar-sail mirror configuration as described in Section 2.2.

These sample trajectories are generated by the FDM process in the Earth–Moon system. Solar gravity is not incorporated into the model. However, transforming these solutions into a Sun–Moon rotating frame does lend insight into the dynamics. Note that the solutions in the Sun–Moon rotating frame are based on gravity from the Earth and the Moon, as well as solar radiation pressure, but *not* gravity from the Sun. These Sun–Moon views of the reference trajectories appear in Fig. 5.9. In this frame, the sunlight direction remains fixed. Recall that Orbits 1 (aqua) and 2 (royal-blue) are located under the Earth–Moon L_1 and L_2 points, respectively. As the Earth–Moon system rotates in the Sun–Moon frame, these locations and the orbits associated with them rotate as well. The motion of these orbits is counter-clockwise as evident

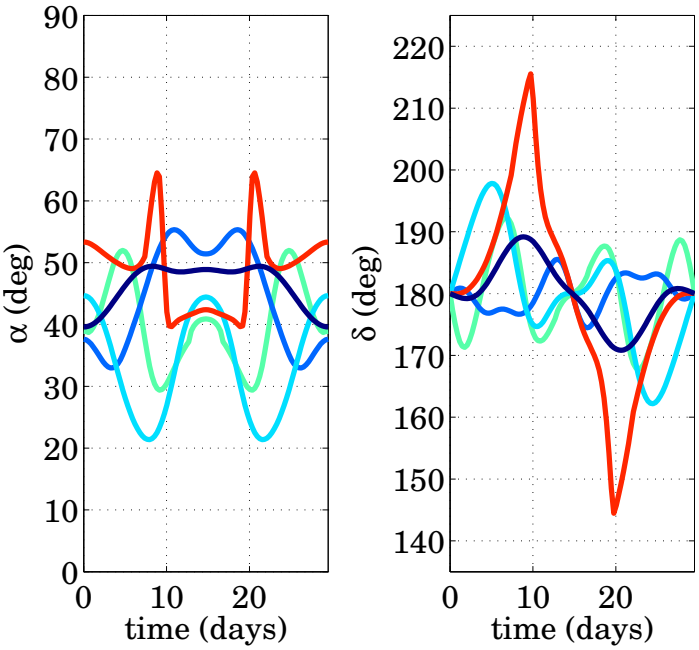
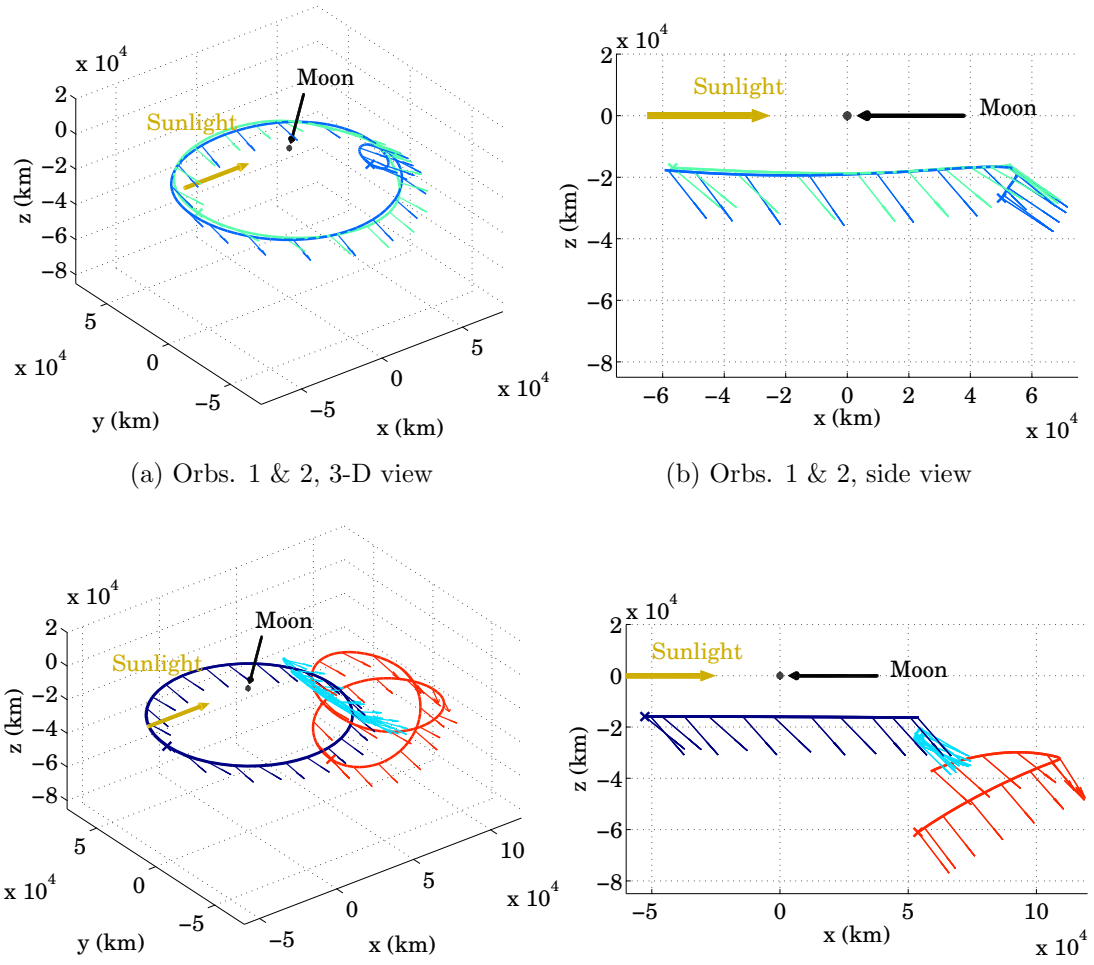


Fig. 5.8: Pitch and clock angles for the five sample orbits.

in Fig. 5.9(a). A spacecraft moving along Orbit 3 or 4 remains on the opposite side of the Moon from the Sun and moves in a clockwise direction in the Earth–Moon frame, as depicted in Fig. 5.6(c). In the Sun–Moon rotating frame, these orbits appear to be located under a fictitious L_2 point in the Sun–Moon system. Further exploration of this observation is beyond the scope of this document but may lead to additional insights. The dark-blue orbit, which completes one counter-clockwise revolution in the Earth–Moon frame, completes two revolutions in the Sun–Moon frame. Finally, note that the arrows that indicate the direction of the sail-face normal are generally away from the sunlight direction.

5.3 Instantaneous equilibrium in the Earth–Moon rotating frame

Periodic trajectories for space vehicles in the CR3B system are often designed in the vicinity of the five equilibrium solutions to the nonlinear differential equations that govern motion in the system. When a solar sail possessing a small characteristic



(a) Orbs. 1 & 2, 3-D view

(b) Orbs. 1 & 2, side view

(c) Orbs. 3–5, 3-D view

(d) Orbs. 3–5, side view

Fig. 5.9: Sample trajectories in the Sun–Moon rotating frame.

acceleration is incorporated into a CR3B regime, these five equilibrium point solutions evolve into five three-dimensional artificial equilibrium surfaces originating from the original Lagrange points. As the sail characteristic acceleration is increased the surface attached to L_3 connects to the surfaces attached to L_4 and L_5, depending on the location of the Sun. Further connections of individual surfaces depend on geometry and the value of the sail characteristic acceleration [4, 181]. On these surfaces, the sail-face normal is aligned in a direction opposite to that of the gravity-gradient vector associated with the primaries.

In the Sun–Earth system, this equilibrium surface is fixed relative to the frame rotating with the Sun–Earth line. In the Earth–Moon system, however, the Sun's rays are constantly moving with respect to the Earth–Moon gravitational field, generating a surface of instantaneous equilibrium locations that are time variant. Similar to the instantaneous solutions in the BCR4B system, these solutions are labeled "equilibria" for purposes of this analysis, despite not being true equilibrium solutions in the stationary sense of the term. McInnes discusses short-term (approximately 3-hour) solutions on an Earth–Moon equilibrium surface by assuming they are fixed over such a time span [4]. Forcing a spacecraft to follow this instantaneous equilibrium surface for longer times necessarily violates the concept of equilibrium.

Some investigations into the motion of solar sails in the Earth–Moon regime examine these instantaneous equilibrium surfaces as starting points for generating periodic orbits [4, 181]. In the present investigation, solutions are developed via numerical techniques without regard to equilibrium solutions. However, overlaying the resulting trajectories on corresponding instantaneous equilibrium surfaces provides insight into the underlying dynamical structure that governs the motion in the system.

Generating an instantaneous equilibrium surface for solar sails in the Earth–Moon CR3B regime assumes that the velocity and acceleration terms in Eq. (2.6) are equal to zero, that is,

$$\boldsymbol{\nabla} U(\mathbf{r}) = \mathbf{a}_s(t) \tag{5.19}$$

Equation (5.19) is combined with Eq. (2.8) and rearranged to solve for the required characteristic acceleration as a function of position and time, that is,

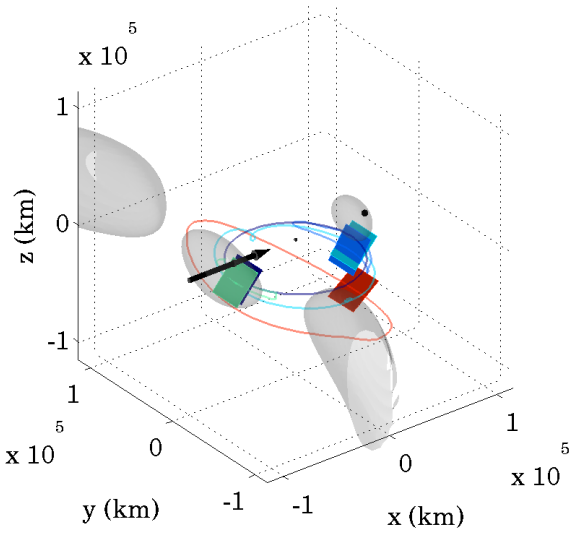
$$a_c(\mathbf{r}, t) = \frac{|\boldsymbol{\nabla} U(\mathbf{r})|^3 a^*}{\left(\hat{\boldsymbol{\ell}}(t) \cdot \boldsymbol{\nabla} U(\mathbf{r})\right)^2} \tag{5.20}$$

where a^* is the system characteristic acceleration from Eq. (2.4). Using a grid of locations for $\mathbf{r}$, the results from applying Eq. (5.20) are then contoured, consistent with various values of a_c, to reveal equilibrium solutions for sails with different characteristics. Contouring in three dimensions reveals that the set of instantaneous equilibrium solutions for a given a_c is best described as a surface.

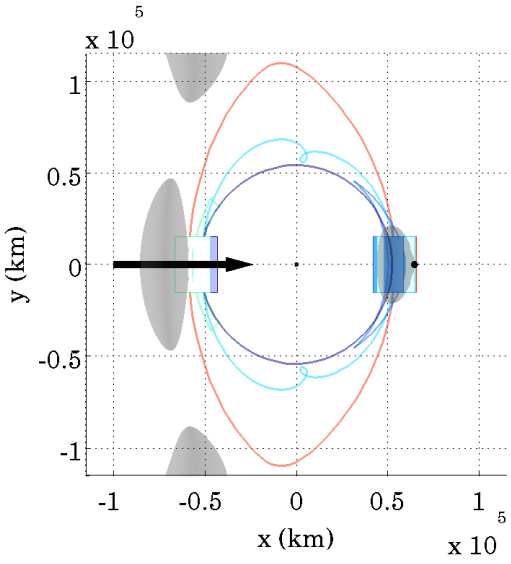
For any non-zero value of a_c, the surface evolves as the Sun direction changes in the Earth–Moon frame. To demonstrate this evolution, a_c is selected to be 1.70 mm/s^2. Initially, at $\Omega t = 0$, an in-plane equilibrium torus wraps around the Earth.[1] A second, smaller, equilibrium surface exists on the far side of the Moon. Motion in the vicinity of the Moon is of particular interest in this investigation, and the five sample trajectories from Section 5.2 overlay the instantaneous equilibrium surfaces appearing in Figs. 5.10 to 5.15. In the figures, the sunlight direction (originating at the Sun), or the Sun-line, is represented by the large black arrow, the Moon is centered at the origin of the coordinate system. The location and orientation of a sailcraft appears as a colored plane along each trajectory at the time associated with the specified direction of the Sun-line. The Moon, as well as the L_1 and L_2 points from the Earth–Moon CR3B regime, are included for scale. A view of the evolution of these surfaces for $a_c = 1.70$ mm/s^2 that includes the Earth and the Moon appears in Appendix C.

Some observations regarding the motion of a sailcraft as related to the evolution of these equilibrium surfaces are apparent from the figures. First, observe that the motion of the vehicle along the reference orbits is initially in the $-\hat{\mathbf{y}}$ direction, as is the "motion" of the equilibrium surface. When the angle of the Sun-line approaches and recedes from 90° and 270°, both the vehicle and the surface are located on the opposite side of the Moon from the Sun. In fact, the location of the vehicle appears to track the "bulkier" portion of the surface. The vehicle is not necessarily connected to the surface at any time, but it is, in general, located near a point tangential to the southern ($-z$) region of the surface at most instants. Although not included in this document, larger values of a_c correspond to larger volumes enclosed by these surfaces, and smaller characteristic accelerations correspond to smaller volumes. Established in Section 4.4.1, larger minimum elevation angles correspond to larger values of a_c.

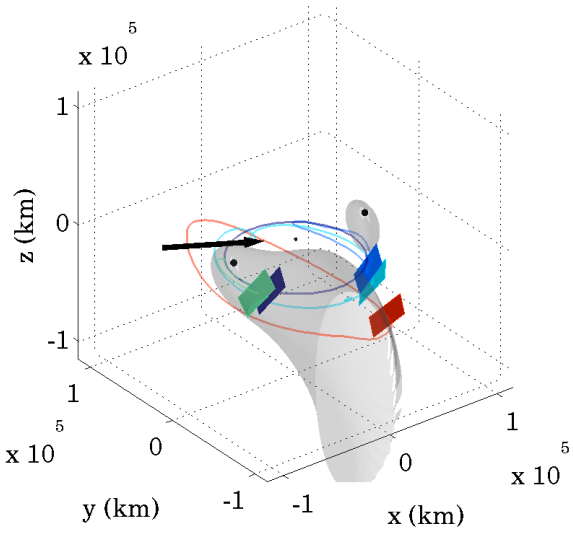
[1]For larger values of a_c, this torus extends in the vertical directions on the Sun side of the Earth–Moon barycenter [181].



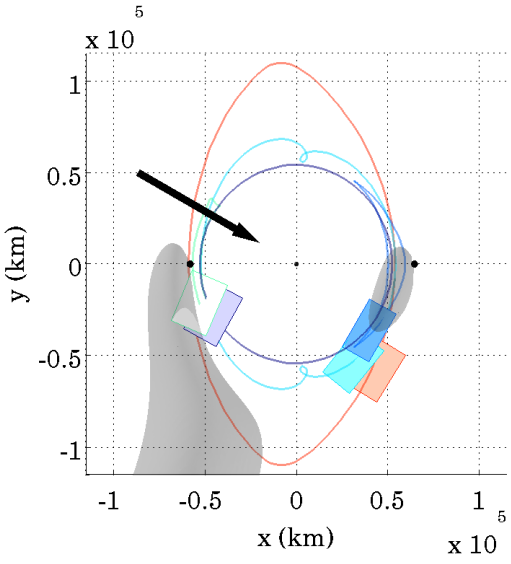
(a) $\Omega t = 0°$, 3-D



(b) $\Omega t = 0°$, xy view

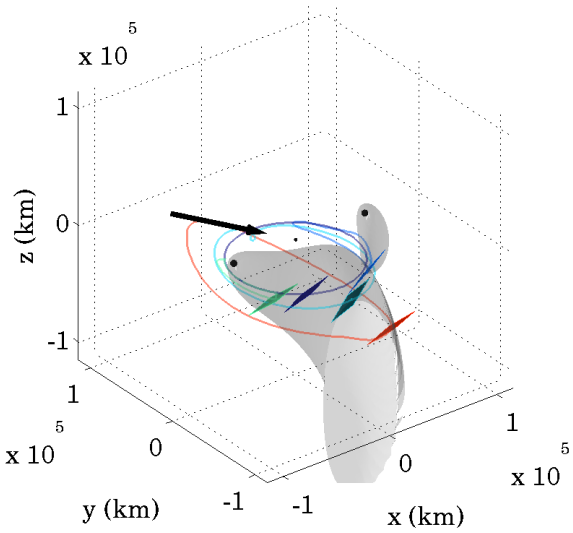


(c) $\Omega t = 30°$, 3-D

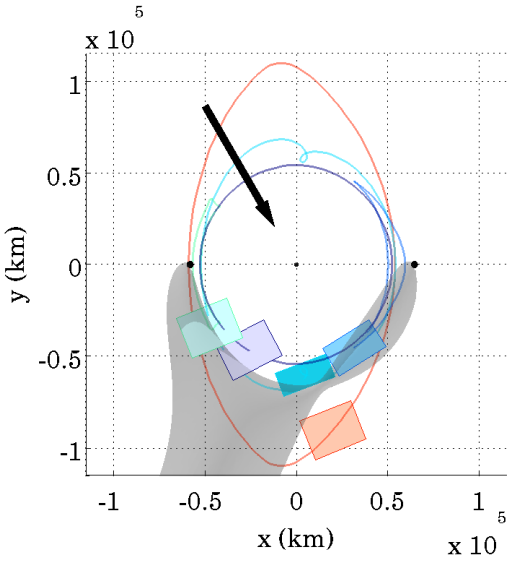


(d) $\Omega t = 30°$, xy view

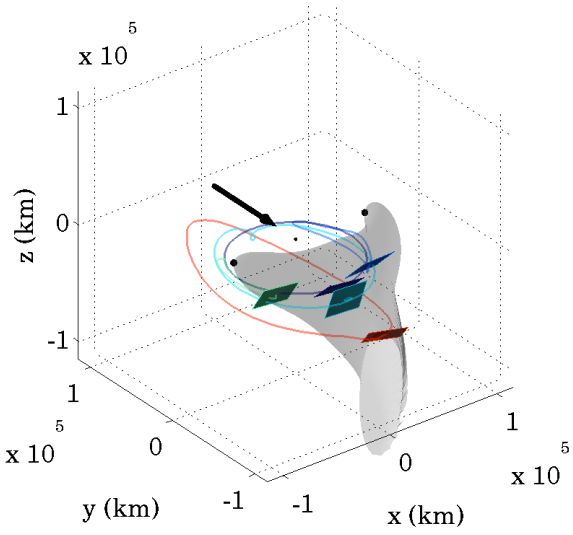
Fig. 5.10: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 0° and 30°.



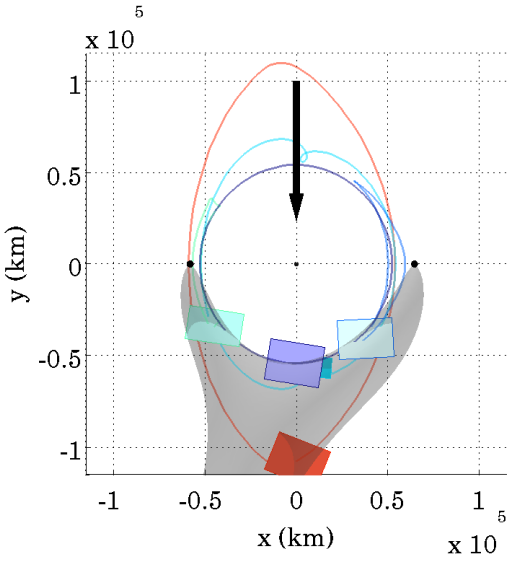
(a) $\Omega t = 60°$, 3-D



(b) $\Omega t = 60°$, xy view

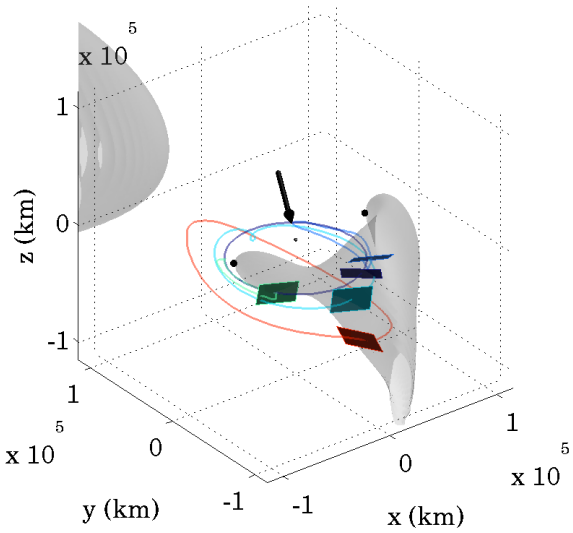


(c) $\Omega t = 90°$, 3-D

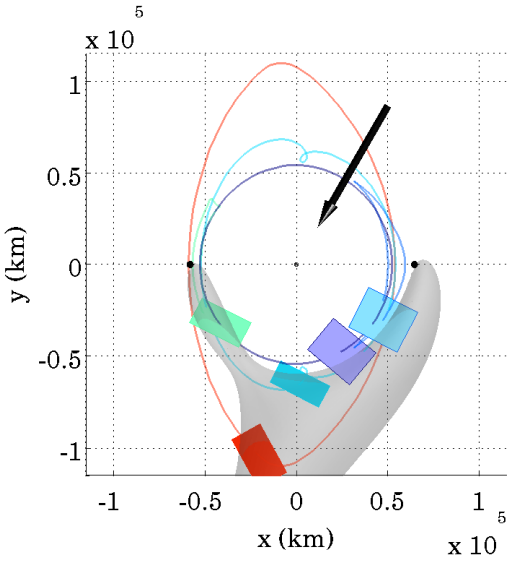


(d) $\Omega t = 90°$, xy view

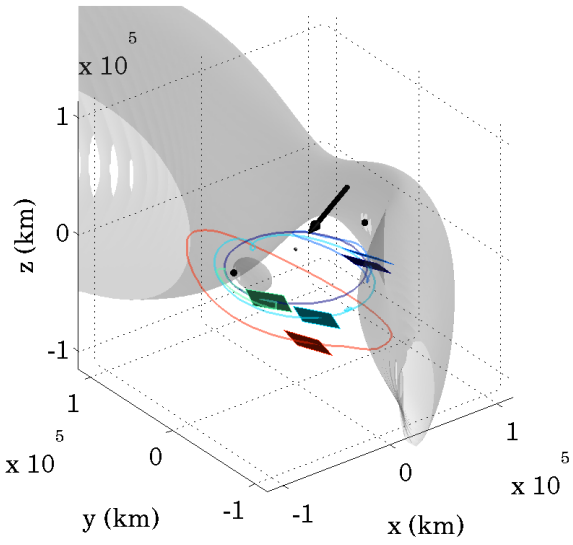
Fig. 5.11: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 60° and 90°.



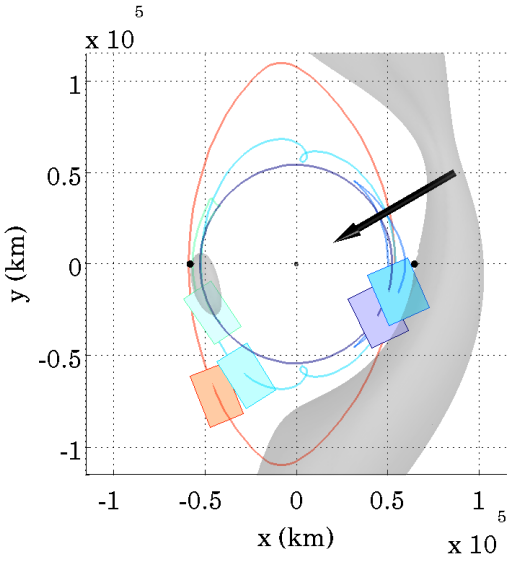
(a) $\Omega t = 120°$, 3-D



(b) $\Omega t = 120°$, xy view

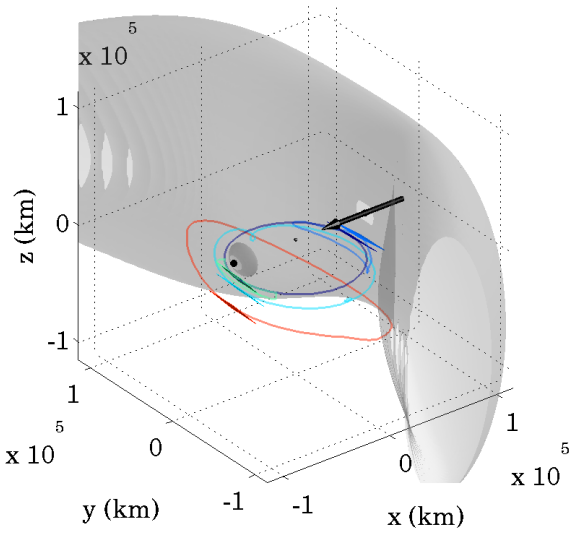


(c) $\Omega t = 150°$, 3-D

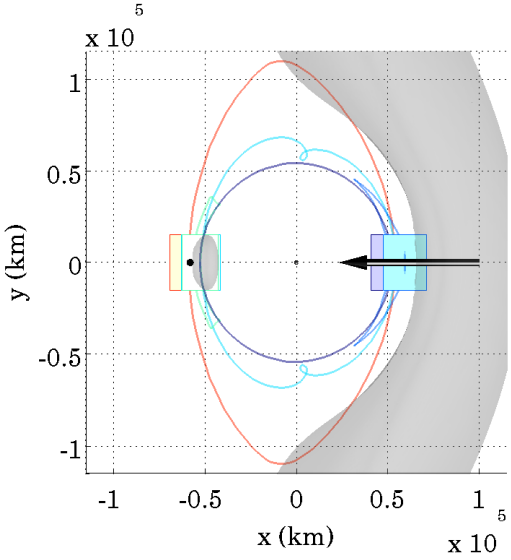


(d) $\Omega t = 150°$, xy view

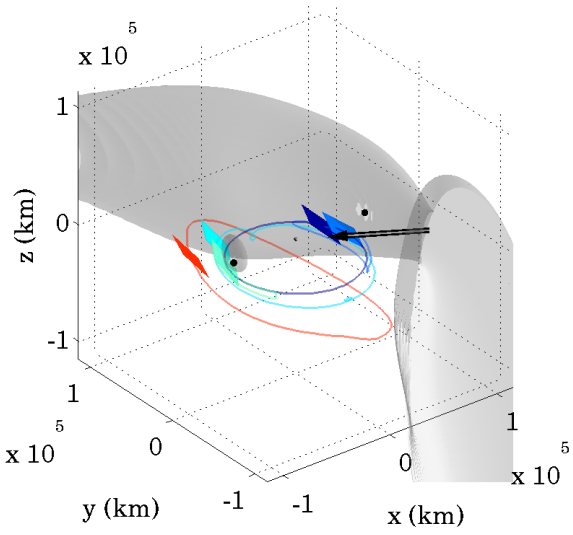
Fig. 5.12: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 120° and 150°.



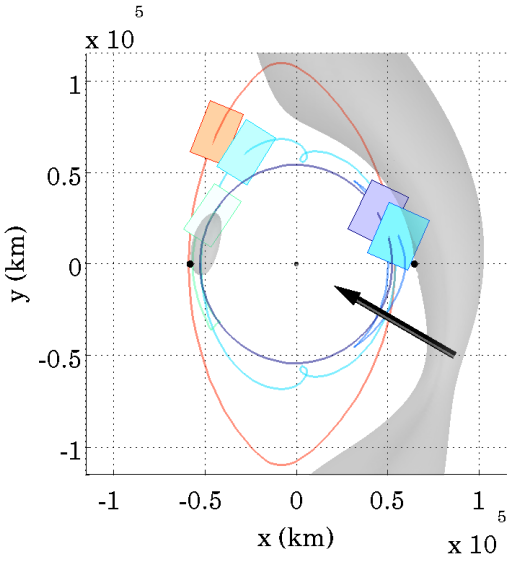
(a) $\Omega t = 180°$, 3-D



(b) $\Omega t = 180°$, xy view

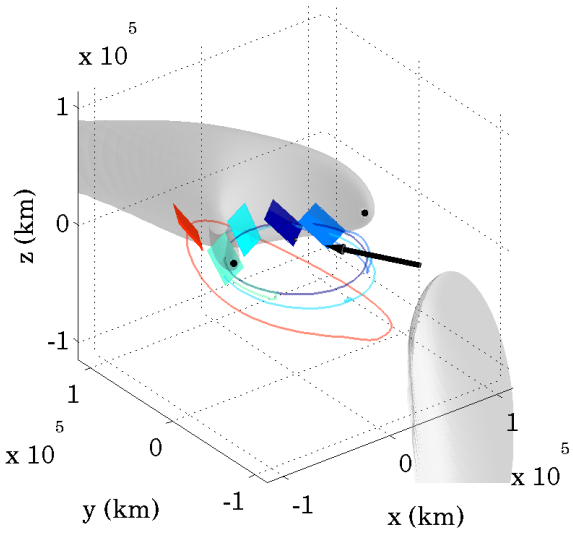


(c) $\Omega t = 210°$, 3-D

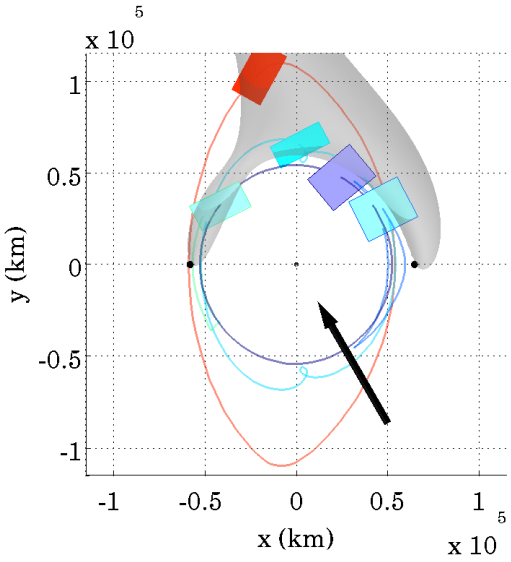


(d) $\Omega t = 210°$, xy view

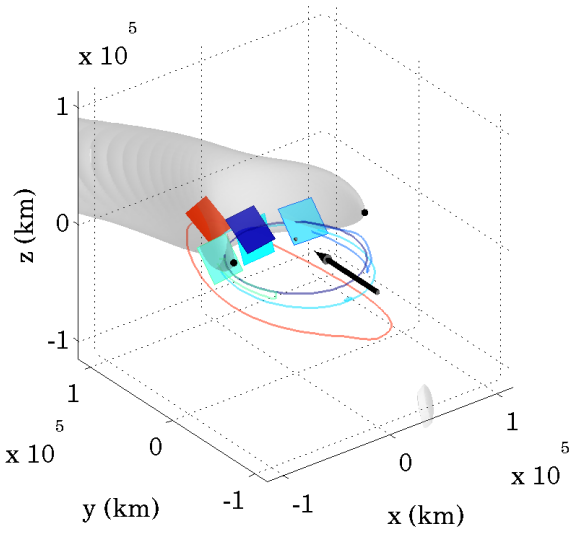
Fig. 5.13: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 180° and 210°.



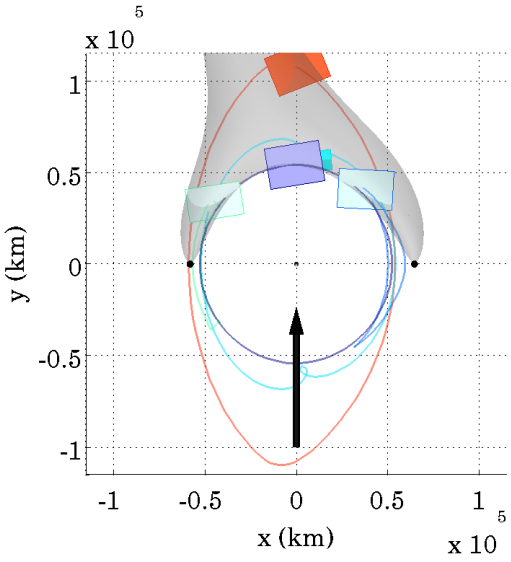
(a) $\Omega t = 240°$, 3-D



(b) $\Omega t = 240°$, xy view

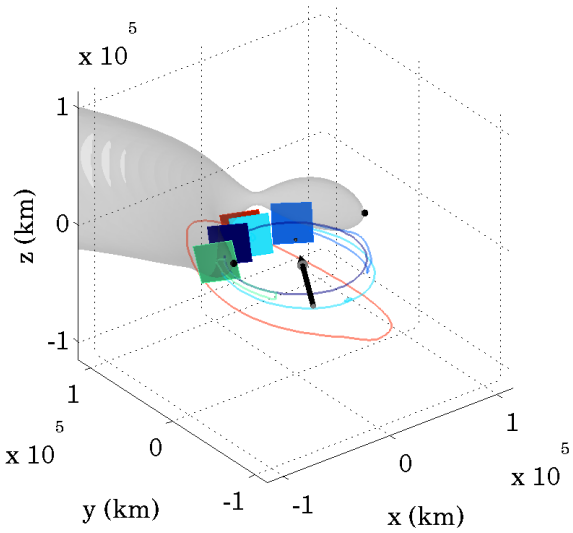


(c) $\Omega t = 270°$, 3-D

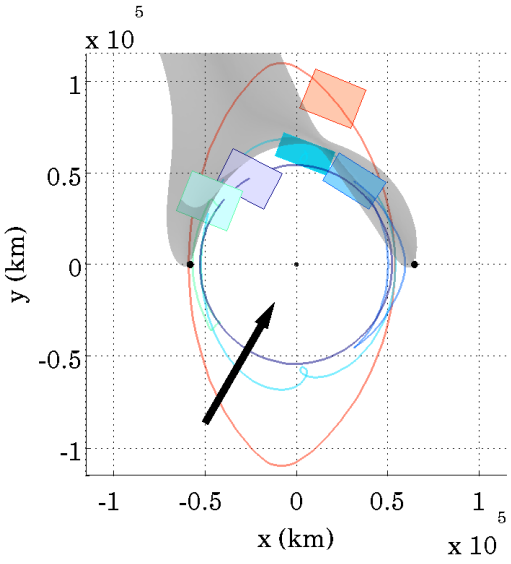


(d) $\Omega t = 270°$, xy view

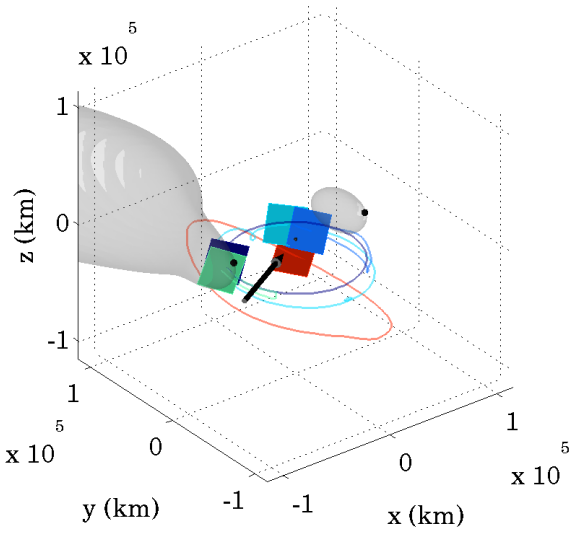
Fig. 5.14: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 240° and 270°.



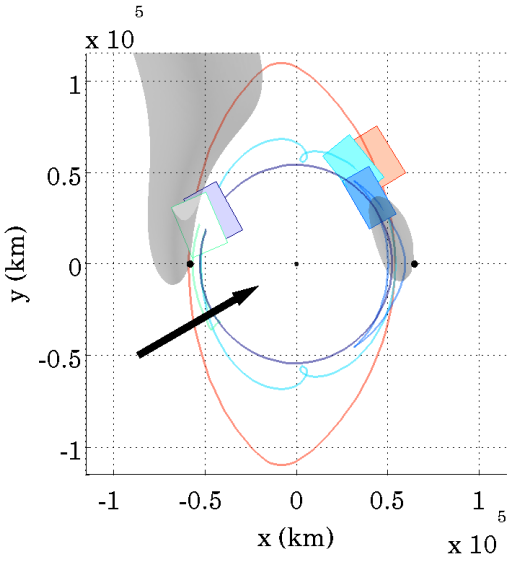
(a) $\Omega t = 300°$, 3-D



(b) $\Omega t = 300°$, xy view



(c) $\Omega t = 330°$, 3-D



(d) $\Omega t = 330°$, xy view

Fig. 5.15: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 300° and 330°.

The characteristic acceleration, the minimum elevation angle and the extent of the equilibrium surface south of the Earth–Moon plane all appear to be related.

Instantaneous equilibrium surfaces that incorporate solar gravity, in addition to the effect of a solar sail, in a BCR4B model are nearly identical to the surfaces highlighted in Figs. 5.10 through 5.15. It is not expected that trajectories developed in a BCR4B system will differ significantly from those developed in the CR3B system. For an appropriate comparison to Figs. 5.10 through 5.15, the sample trajectories should be recalculated in the BCR4B model and overlayed on the solar-gravity surfaces. Recall, however, that the acceleration from solar gravity is approximately 100 times smaller than that from the Earth, the Moon, or for a sail that possesses a characteristic acceleration of $a_c = 1.70$ mm/s^2. Given this limited but increased understanding of the solution, trajectories must be simulated for actual flight. Experience in this model suggests that a direct transition to an ephemeris-based model may be possible. The use of solutions generated in the Earth–Moon CR3B model to initialize an FDM process for generating solutions in a formulation based on the ephemeris positions of solar system bodies is explored in the next chapter.

6. TRANSITION TO AN EPHEMERIS-BASED MODEL

When developing spacecraft trajectories to support various mission scenarios, designers frequently use experience and straightforward physical models as the basis for an orbit concept. Keplerian ellipses are appropriate approximations in two-body systems with small perturbations. Dynamical systems theory can be employed to supply a framework in a multi-body regime where transfers between orbits in the vicinity of the Lagrange points are designed within the context of the circular restricted three-body (CR3B) system [101,186]. Both of these approaches are transitioned to full ephemeris dynamical environments, and the resulting trajectories resemble trajectories from the lower-fidelity models.

The trajectory design problem is further complicated when the additional force from a solar sail is incorporated. To address the challenge of solar sail mission design in a multi-body regime, various authors employ numerical techniques to generate both the path and the associated control profile for the solar sail while relying on relatively crude initial guesses for both the trajectory and the maintenance strategy [30,34,129]. The finite-difference method (FDM) approach to the solving the lunar south pole coverage problem with a solar sail spacecraft within the context of a CR3B dynamical environment including terrestrial and lunar gravity, as well as solar radiation pressure, is documented in Chapter 3.

The solutions from Chapter 3, however, are incomplete. To model the motion of the vehicle for a mission scenario properly, solutions from the simpler model must be transitioned to a higher-fidelity model, such as one based on ephemeris positions and the gravitational effects of the Earth, the Moon, and the Sun [30]. Because the effect of solar gravity on the instantaneous equilibrium surfaces in the Earth–Moon system is minor, a direct transition to an ephemeris-based model and bypassing a bicircular restricted model that incorporates solar gravity is possible. Additional effects

from other solar system bodies may also be included. Furthermore, higher-fidelity reflectivity models for the sail can also be introduced. For the current investigation, the sample trajectories identified in Section 5.2 are used to initialize an FDM process that is based on the ephemeris locations of the Sun, the Earth, and the Moon. The procedure employed to transition the solutions from the Earth–Moon CR3B system to the higher-fidelity model, as well as the results of that transition, are detailed below. Error properties for the FDM scheme discussed in Section 3.3.1 apply to the FDM scheme developed for the inertial, ephemeris model as well. The FDM algorithm is demonstrated to be a viable tool for generating trajectories in high-fidelity models.

6.1 Transition to inertial, ephemeris model

The advantage of developing trajectories in the Earth–Moon CR3B system is that periodicity is constrained and the resulting orbits are bounded. Orbits that are exactly periodic are not possible within the context of an ephemeris model. However, the periodic orbits from the simpler system can be employed to initialize an FDM process in the higher-fidelity model. Two challenges are encountered during an implementation of such a transition. First, the CR3B solution must be transformed into the appropriate path in the inertial frame. Given the new path, it is equally important to formulate an FDM algorithm properly to generate a viable solution.

6.1.1 Transforming the CR3B solution into an inertial, ephemeris frame

Prior to transitioning a CR3B system to an ephemeris-based model, an appropriate ephemeris must be selected. For this investigation, the Jet Propulsion Laboratory (JPL) planetary and lunar ephemeris DE 421 is employed, which supplies the orbits of the Sun, planets (including Pluto), and the Moon for the years 1900–2050 [187]. This ephemeris is an improvement on previous ephemerides such as the DE 418 [188] and the widely used DE 405 [187] and is "recommended for lunar missions" [189];

lunar ephemerides in the DE 421 ephemeris are available to sub-meter accuracy. However, any of these three ephemerides are sufficient to demonstrate the process.

The ephemerides are loaded into the algorithm using MATLAB®-based SPICE routines, or MICE, available from the JPL Navigation and Ancillary Information Facility (NAIF) [190]. Other publicly available SPICE kernels are also employed in this analysis and are listed in Table 6.1 in the order in which they are loaded into the MATLAB® environment.[1]

Table 6.1: SPICE kernels used in the high-fidelity simulation

Kernel	Description
naif0009.tls	Leap second kernel file
pck00009.tpc	General planetary constants kernel (PCK) file
de-403-masses.tpc	GM/mass kernel file (based on DE 403, no update for DE 421)
de421.bsp	Binary planetary and lunar ephemeris
moon_pa_de421_1900-2050.bpc	Binary PCK file for the Moon (lunar orientation, etc.)
moon_080317.tf	Kernel containing lunar frame specifications compatible with moon_pa_de421_1900-2050.bpc

To begin the transformation from the solution in the Earth–Moon CR3B regime into an ephemeris-based model, a date is identified when the Moon is in opposition (for example, 15 July 2011, 06:15:41.796 ephemeris time, ET; this time is denoted as t_0 subsequently) to most closely match the initial conditions of the orbit from the CR3B model. Multiple revolutions from the CR3B orbit are generated by repeating the periodic path any number of synodic months. Although different characteristic

[1]All files are current as of 2 July 2011 and are located in the subdirectories of ftp://naif.jpl.nasa.gov/pub/naif/generic_kernels/.

lengths and masses are employed to generate the CR3B trajectories, the distance from the Earth to the Moon, $L^* = 385156.4$ km, is frozen at this value once this date is selected. Additionally, all other reference dimensions and gravitational parameters are obtained from the files appearing in Table 6.1. Based on *these* dimensions and parameters, the nondimensional path from the CR3B model is dimensionalized and the nondimensional time associated with each state is converted to seconds past the J2000 epoch (1 January 2000, 12:00:00 ET). If appropriate, the states from the reference path can be interpolated.

Rotating the reference path to the ephemeris-based inertial frame employs the instantaneous position and velocity of the Moon with respect to the Earth from the ephemeris data. Let H be the inertial reference frame corresponding to a particular set of ephemerides (i.e., the Earth-mean-equator or Earth-mean-ecliptic at the J2000 epoch) and let $\mathbf{r}^H_{me,i}$ and $\mathbf{v}^H_{me,i}$ be the position and velocity vectors, respectively, of the Earth with respect to the Moon obtained from ephemeris data in this inertial frame at each epoch, t_i, corresponding to the reference path data. A set of unit vectors are constructed based on $\mathbf{r}^H_{me,i}$ and $\mathbf{v}^H_{me,i}$, that is,

$$\hat{\mathbf{X}}^H_i = -\frac{\mathbf{r}^H_{me,i}}{\|\mathbf{r}^H_{me,i}\|}, \quad \hat{\mathbf{Z}}^H_i = \frac{\mathbf{r}^H_{me,i} \times \mathbf{v}^H_{me,i}}{\|\mathbf{r}^H_{me,i} \times \mathbf{v}^H_{me,i}\|}, \quad \hat{\mathbf{Y}}^H_i = \hat{\mathbf{Z}}^H_i \times \hat{\mathbf{X}}^H_i \tag{6.1}$$

The instantaneous rotation rate vector of the R frame attached to the Earth and the Moon with respect to the H frame is

$$ {}^H\boldsymbol{\omega}^R_i = \frac{\mathbf{r}^H_{me,i} \times \mathbf{v}^H_{me,i}}{\|\mathbf{r}^H_{me,i}\|^2} \tag{6.2}$$

$$= \frac{\|\mathbf{r}^H_{me,i} \times \mathbf{v}^H_{me,i}\|}{\|\mathbf{r}^H_{me,i}\|^2} \hat{\mathbf{Z}}^H_i \tag{6.3}$$

Finally, a matrix that rotates a spacecraft state consisting of the position and velocity at t_i from the R frame to the H frame is

$$ {}^H\mathbf{T}^R_i = \begin{bmatrix} \mathbf{C}_i & \mathbf{0} \\ \frac{d}{dt}\mathbf{C}_i & \mathbf{C}_i \end{bmatrix} \tag{6.4}$$

where

$$\mathbf{C}_i = \left[\begin{array}{ccc} \hat{\mathbf{X}}_i^H & \hat{\mathbf{Y}}_i^H & \hat{\mathbf{Z}}_i^H \end{array} \right] \tag{6.5}$$

$$\frac{d}{dt}\mathbf{C}_i = \left[\begin{array}{ccc} {}^H\boldsymbol{\omega}_i^R \times \hat{\mathbf{X}}_i^H & {}^H\boldsymbol{\omega}_i^R \times \hat{\mathbf{Y}}_i^H & {}^H\boldsymbol{\omega}_i^R \times \hat{\mathbf{Z}}_i^H \end{array} \right] \tag{6.6}$$

A rotation from the H frame to the R frame uses the inverse of ${}^H\mathbf{T}_i^R$. These rotations are necessary because the orientation with Earth–Moon orbital plane is offset from the ecliptic plane, ranging between 4.99° and 5.30° for the three years after t_0 (or between 18.94° and 23.36° with respect to the Earth-mean-equatorial plane for the same period). An assumption that the Earth, the Moon, and the Sun all lie in the same plane leads to a less accurate initialization of the numerical process (e.g., the FDM algorithm) that is formulated in terms of ephemeris locations of the specified bodies. All vectors highlighted in this chapter are expressed in the H frame, and, for simplicity, superscript H is dropped from the notation for the remainder of this chapter.

After rotating the states from the CR3B reference paths to the inertial frame and prior to executing the FDM process to refine those paths based on ephemeris information, the locations of the Sun, the Earth, and the Moon are extracted from the ephemeris file at each epoch, t_i, and the position vectors are repositioned to originate at a base point at the instantaneous location of the Moon. Additionally, the direction from the center of the Moon to the lunar south pole at each epoch is determined by invoking `cspice_pxform()`. This SPICE function produces a position-based transformation matrix based on the relative orientation of two reference frames. If the arguments into the `cspice_pxform` function are structured so that the result of `cspice_pxform` produces a transformation matrix from an inertial frame to a Moon-fixed body frame, the third column of the matrix is the vector to the lunar north pole. For this application, either a Moon-fixed frame based on the principle lunar inertia axes or a frame based on the north mean lunar rotation axis is employed. The angle between these two frames is 0.02885°, equivalent to approximately 875 m at the lunar poles. In addition to the average 5.16° offset between the Earth–Moon orbital plane

and the ecliptic, the angle between the equatorial plane of the Moon and the ecliptic ranges from 1.49° to 1.60° for the three years following t_0. The angle between the Earth–Moon orbital plane and the equatorial plane of the Moon ranges from 3.70° to 3.91° during that same period.

6.1.2 Formulating the FDM for an ephemeris-based trajectory

The formulation for the FDM algorithm presented in Section 3.3 requires only a few significant changes for use in the ephemeris-based inertial frame. The first modification is the change in the constraint equations. The second change in the reformulation of the FDM algorithm is an update of the differential equations that govern motion in the system.

The FDM accommodates both equality and inequality path constraints in both the CR3B formulation and the ephemeris-based formulation. When the reference trajectories are generated in the CR3B model, the elevation-angle constraint includes a margin to accommodate differences in frame alignment and dynamical characteristics in comparison to an ephemeris-based model. In the high-fidelity formulation, the elevation-angle constraint is relaxed. The form of the constraint in described by Eqs. (3.38) is also generalized to incorporate the correct alignment of the lunar south pole direction as observed from the inertial frame. The new constraint equation is

$$E_{\min} \leq \arcsin\left(\frac{\mathbf{r}_{op} \cdot \hat{\mathbf{r}}_{mo}}{r_{op}}\right) \tag{6.7}$$

where the term on the left is the new elevation angle, $\mathbf{r}_{mo}$ is the vector from the center of the Moon to the lunar south pole, and $\mathbf{r}_{op}$ is the vector from the lunar south pole to the vehicle. The minimum elevation angle is $E_{\min}$. These vectors and angles appear in Fig. 6.1. The attitude constraint from Eq. (3.40) and the unit magnitude constraint on the control vector from Eq. (3.36) remain, but the altitude constraint is neglected because locations along all of the reference paths are sufficiently close to the Moon for this investigation. Additionally, because the solution will not be periodic in the ephemeris-based inertial system, the periodicity constraint is also eliminated.

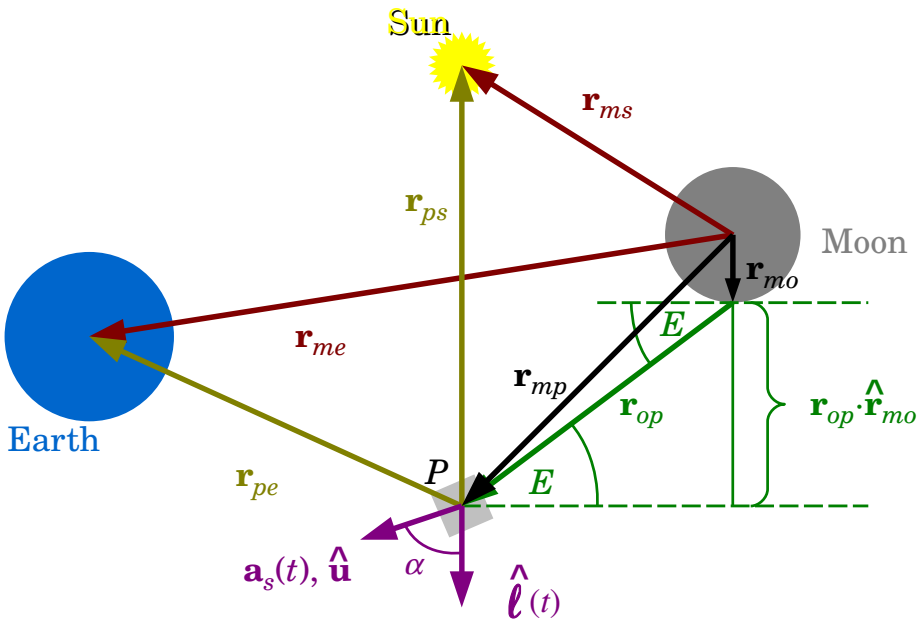


Fig. 6.1: Vectors employed in the Moon-centered equations of motion and vectors for the elevation-angle constraint (in green).

A new formulation of the equations of motion is also required when searching for trajectories in an ephemeris-based model. To establish this new dynamical model, the Moon is defined as the origin of the coordinate system, and the Sun and the Earth are assumed to move relative to that point. The vector form of the familiar two-body relative equation of motion with perturbations from the gravitational effects from other celestial bodies [83,85,184], as well as the acceleration due to an ideal solar sail, is

$$\ddot{\mathbf{r}}_{mp} + \frac{G(m_p + m_m)}{r_{mp}^3}\mathbf{r}_{mp} = Gm_e\left(\frac{\mathbf{r}_{pe}}{r_{pe}^3} - \frac{\mathbf{r}_{me}}{r_{me}^3}\right) + Gm_s\left(\frac{\mathbf{r}_{ps}}{r_{ps}^3} - \frac{\mathbf{r}_{ms}}{r_{ms}^3}\right) + a_c\left(\frac{\mathrm{AU}}{r_{ps}}\right)^2 \cos^2\alpha\hat{\mathbf{u}} \tag{6.8}$$

where G is Newton's gravitational constant and m_p is the mass of the sail, which is negligible compared to the mass of the Moon, m_m. The mass of the Earth and the Sun are m_e and m_s, respectively. All vectors in Eq. (6.8) appear in Fig. 6.1;

note that the path is written in terms of $\mathbf{r}_{mp}$ after the transformation to the inertial coordinate frame. Only the gravitational effects of the Sun, the Earth, and the Moon are incorporated in the equations of motion in this demonstration. Effects from other bodies are considered negligible. The average gravitational contributions due to the Moon and other possible, significant solar system bodies on the path of a particle following the dark-blue orbit are listed in Table 6.2 to highlight the scale of the

Table 6.2: Average gravitational acceleration contributions from selected solar system bodies for a particle along the dark-blue reference trajectory in Fig. 5.6.

Moon	1.58 mm/s^2	0.213 (nondimensional units)
Earth	0.58 mm/s^2	0.580
Sun	3.24×10^{-3} mm/s^2	1.192×10^{-3}
Jupiter	3.61×10^{-8} mm/s^2	1.329×10^{-8}
Venus	1.64×10^{-9} mm/s^2	6.016×10^{-10}

contributions of the three bodies used in this simulation on the motion of the sail compared to neglected solar system bodies. It is noted that the model is not truly coherent without these effects. State-variable and constraint vectors are formulated similar to Eqs. (3.17) and (3.32), respectively, and the new Jacobian, $D\mathbf{F}(\mathbf{X})$, is developed accordingly.

The period of each trajectory is no longer constrained, and an extrapolation of neighboring numerically determined accelerations is employed for calculating $\widetilde{\mathbf{a}}_1$ and $\widetilde{\mathbf{a}}_n$ (e.g., $\widetilde{\mathbf{a}}_1$ is a function of $\mathbf{r}_1$ to $\mathbf{r}_4$ and $\widetilde{\mathbf{a}}_n$ is a function of $\mathbf{r}_{n-3}$ to $\mathbf{r}_n$). Partial derivatives of these accelerations with respect to position elements at neighboring nodes are tractable analytically for a linear extrapolation of the numerically derived acceleration values at two neighboring nodes. Recall that each $\widetilde{\mathbf{a}}_i$ is based on the position vectors $\mathbf{r}_{i-1}$, $\mathbf{r}_i$, and $\mathbf{r}_{i+1}$; therefore, $\widetilde{\mathbf{a}}_1$ is a function of $\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$, and $\mathbf{r}_4$ and

$\widetilde{\mathbf{a}}_n$ is a function of $\mathbf{r}_{n-3}$, $\mathbf{r}_{n-2}$, $\mathbf{r}_{n-1}$, and $\mathbf{r}_n$ when these numerical accelerations are linearly extrapolated.

An important observation concerning the FDM scheme is notable when computing the Jacobian matrix. For the FDM scheme in this high-fidelity simulation, as well as the corrections process in the CR3B simulation, developing the partial derivatives for the Jacobian matrix analytically is generally straightforward; these partial derivatives are often available in a compact form. In the current simulation, $\widetilde{\mathbf{a}}_1$ and $\widetilde{\mathbf{a}}_n$ are extrapolated based on higher-degree polynomials and more node points than those employed for a linear extrapolation. Only in the case where higher-degree polynomials are employed are any partial derivatives, specifically, $\frac{\partial \widetilde{\mathbf{a}}_1}{\partial \mathbf{r}}$ and $\frac{\partial \widetilde{\mathbf{a}}_n}{\partial \mathbf{r}}$, determined by numerical techniques, which increase convergence times compared to formulations based purely on analytically calculated the partial derivatives. In this application, however, the computational cost is not significant and numerically calculating these specific partial derivatives is warranted to improve the accuracy of the solution.

A final pre-processing step is recommended prior invoking the FDM algorithm, that is, the nondimensionalization of all quantities from the planetary ephemerides and the spacecraft trajectory. With nondimensional quantities, partial derivatives of the accelerations with respect to positions, $\frac{\partial \mathbf{a}}{\partial \mathbf{r}}$, possess orders of magnitudes similar to the corresponding acceleration level. Using dimensional values leads to a poorly conditioned Jacobian matrix, $D\mathbf{F}(\mathbf{X})$, in which the orders of magnitude of the elements corresponding to acceleration terms in the matrix are scaled approximately by a factor of 10^{-6} when compared to a similar nondimensional formulation, and, thus, the algorithm does not converge on a solution when using dimensionalized values.

6.2 Sample trajectories transitioned to an ephemeris-based model

One advantage of developing trajectories in the context of the Earth–Moon CR3B regime is that the model incorporates the gravitational contributions from the Earth and the Moon as well as the effect of solar radiation pressure on the sailcraft. The

use of this framework reflects an assumption that this idealized CR3B system is a reasonable simplification of a higher-fidelity model. One such model is based on the ephemeris positions of the Earth and the Moon as well as solar ephemeris positions and solar gravity. A test of the utility of the CR3B assumption is to use a solution developed in that regime to initialize a trajectory generation process in a higher-fidelity model such that the resulting high-fidelity solution resembles the original CR3B solution.

Each of the five sample orbits from Section 5.2 is employed successfully to generate new trajectories based on an ephemeris model. Using the inverse of ${}^{H}\mathbf{T}_{i}^{R}$ from Eq. (6.4) at each epoch, the results from the FDM scheme are transformed into a Moon-centered frame that is based on the instantaneous position and velocity of the Earth with respect to the Moon. Results for a one-month simulation initiating on 15 July 2011 06:15:41.796 ET appear in Fig. 6.2. The locations of L_1 and L_2 from the CR3B formulation are included in the figure for a comparison of the ephemeris-based results to their corresponding solutions in the CR3B model. Note that the locations of L_1 and L_2 are instantaneous in the Moon-centered rotating frame. Because periodicity is not constrained, the orbits in Fig. 5.6 do not close during the one-month simulation. An "x" marks the position of the spacecraft at this initial epoch along each trajectory.

While each solution converges and appears similar to the initial guess supplied by the CR3B result, key orbital characteristics differ between the ephemeris-based results and the CR3B solutions used to initialize the FDM algorithm based on the high-fidelity model. The elevation-angle history for the five reference orbits appears in Fig. 6.3, and the associated pitch and clock angle profiles appear in Fig. 6.4. Recall the definitions of sail pitch and clock from Eqs. (5.17) and (5.18). Simulations do not converge for any sample orbit when the elevation-angle constraint is 15°. Because of the higher-fidelity geometric information is available when incorporating the ephemeris data, a lower elevation-angle constraint is allowable. For this simulation, all orbits reconverge in the ephemeris-based model with a 9° elevation-angle constraint.

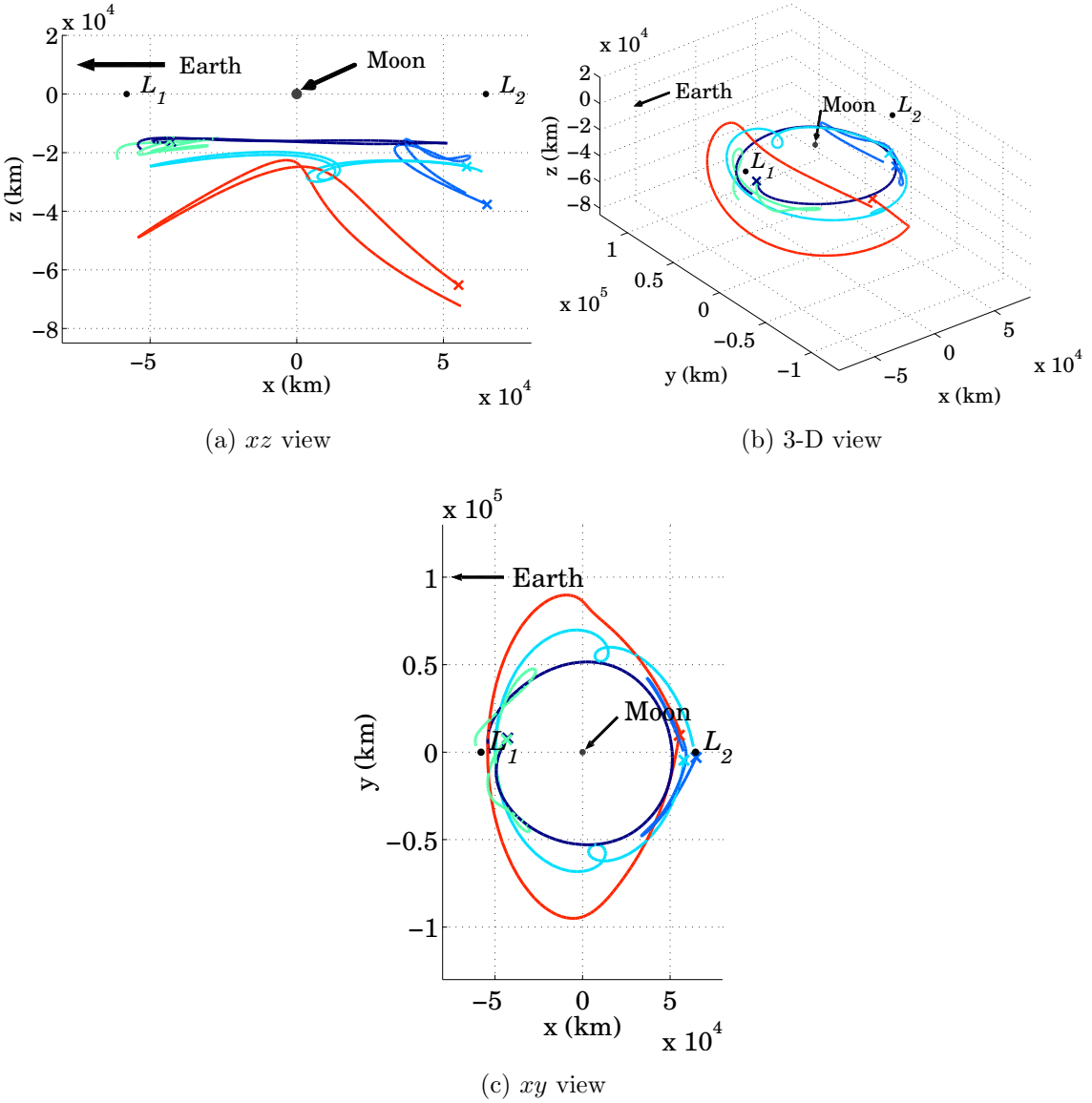


Fig. 6.2: One-month simulation of five sample trajectories as viewed relative to a Moon-centered rotating frame (*cf.* Fig 5.6).

Sail pitch and clock angles as determined in an ephemeris model also differ when compared to the respective angles from the CR3B results. Sail pitch is still defined as the angle between $\hat{\boldsymbol{\ell}}(t_i)$ and $\hat{\mathbf{u}}_t$ and δ is measured from the instantaneous angular momentum vector of the Earth–Moon system about the Sun-line. Note that $\hat{\boldsymbol{\ell}}(t_i)$

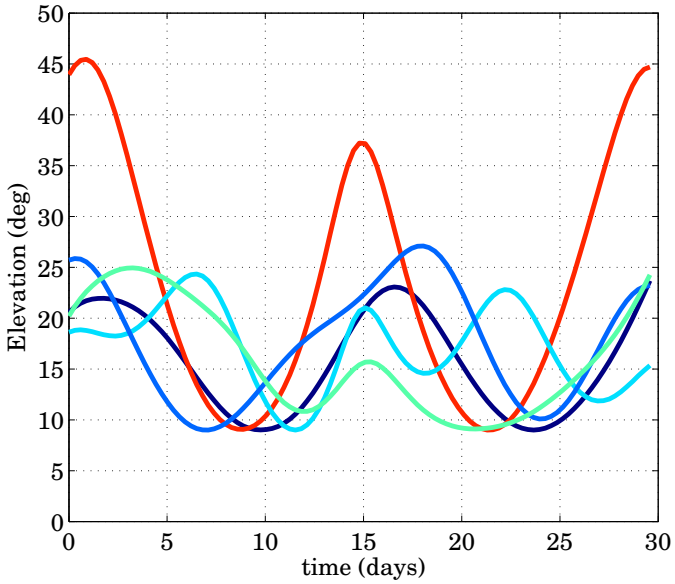


Fig. 6.3: Elevation history for the five reference orbits from the one-month simulation (*cf.* Fig. 5.7).

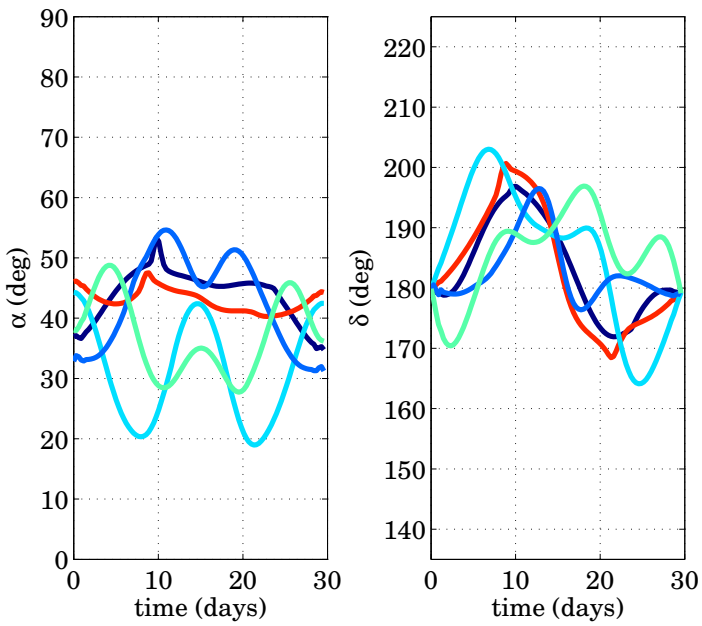
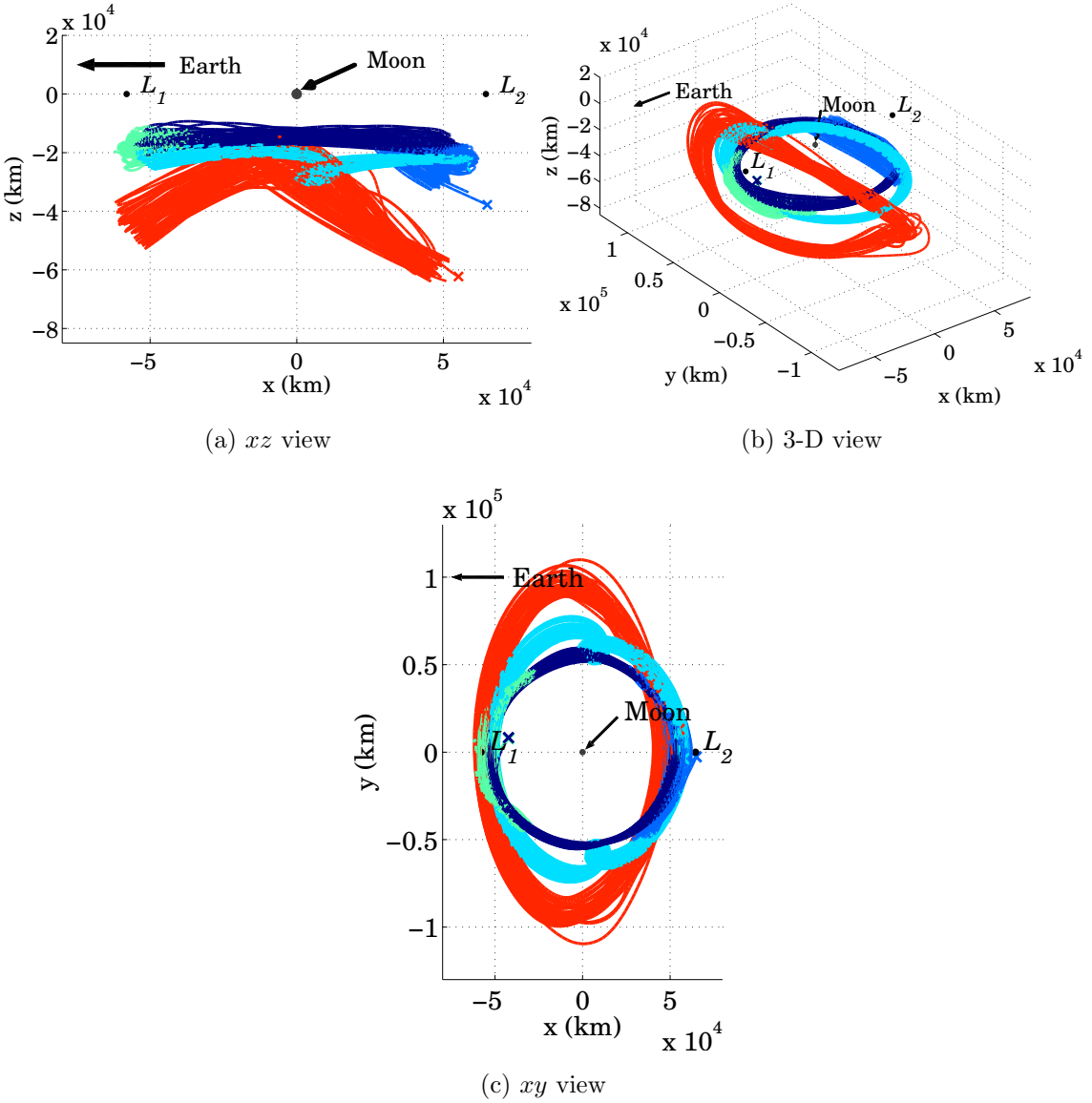


Fig. 6.4: Pitch and clock angles for the five sample orbits from the one-month simulation (*cf.* Fig. 5.8).

and the instantaneous angular momentum vector of the Earth–Moon system are not necessarily orthogonal. The general trend of the profiles for α and δ in Fig. 6.4 are similar to those appearing in Fig. 5.8. However, the maximum pitch angle associated with the red-orange orbit is improved (i.e., reduced) in the ephemeris-based model which may increase the viability of this orbit.

The FDM algorithm is adaptable for any simulation length. Results from a thirty-six-month simulation appear in Figs. 6.5 through 6.7. To illustrate the elevation angles, as well as the pitch and clock angles, only the first 180 days (6 months) of data appear in the Figs. 6.6 and 6.7. One immediate observation from a longer simulation involves the first and last revolutions. The "bounding" revolutions possess segments that extend beyond the quasi-periodic region that defines the balance of the trajectory, even when a more accurate numerical model for the acceleration components is incorporated at the first and the last nodes in the FDM algorithm. It is common practice to include an extra revolution prior to and after the time of interest when performing the simulation when the first and last points are unconstrained, then to discard those revolutions as superfluous to a high-fidelity reference path. The algorithm uses those initial and final locations to increase flexibility, and transfer trajectories are connected to locations along the reference path interior to these nodes.

Further improvement in model fidelity is possible. Effects from additional solar system bodies are easily incorporated into the differential equations and subsequent partial derivatives of the finite-difference method. However, prior experience suggests that these effects are minor in comparison to improvements in the sail reflectivity model and the incorporation of disturbances due to lunar and terrestrial spherical harmonics. An optical model (Appendix A) represents the next level of fidelity in the spacecraft model. Further improvements to the solar sail reflectivity model incorporate sail deformation and manufacturing defects [6, 149, 191]. Additionally, higher-fidelity numerical methods improve the numerical precision of the solution [30,95,129]. The solutions from this ephemeris model demonstrate that the FDM algorithm is



(a) xz view

(b) 3-D view

(c) xy view

Fig. 6.5: Thirty-six-month simulation of five sample trajectories as viewed relative to a Moon-centered rotating frame.

useful for improving the accuracy of lower-fidelity solutions. Despite the accuracy or precision employed in the determination of a reference trajectory, however, the flight path of a spacecraft following that reference orbit must be controlled.

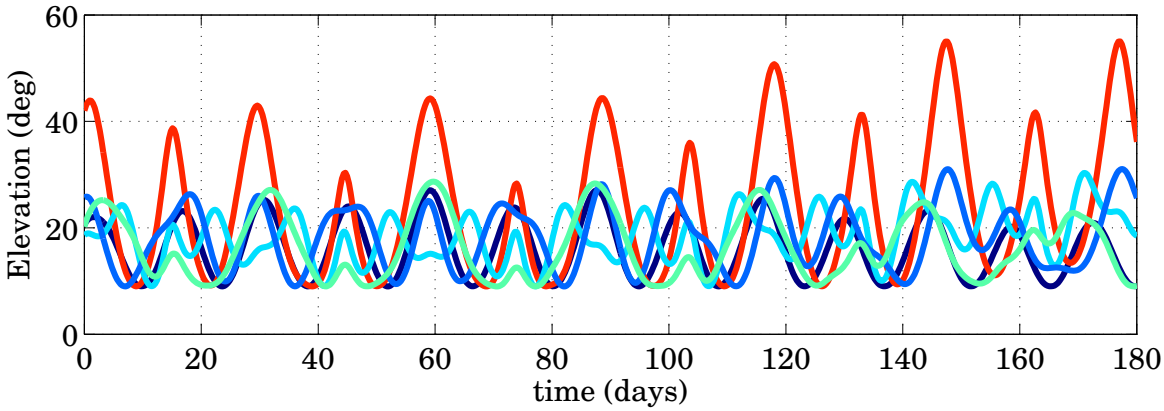


Fig. 6.6: Elevation history for the five reference orbits from the thirty-six-month simulation; only the first 180 days are plotted.

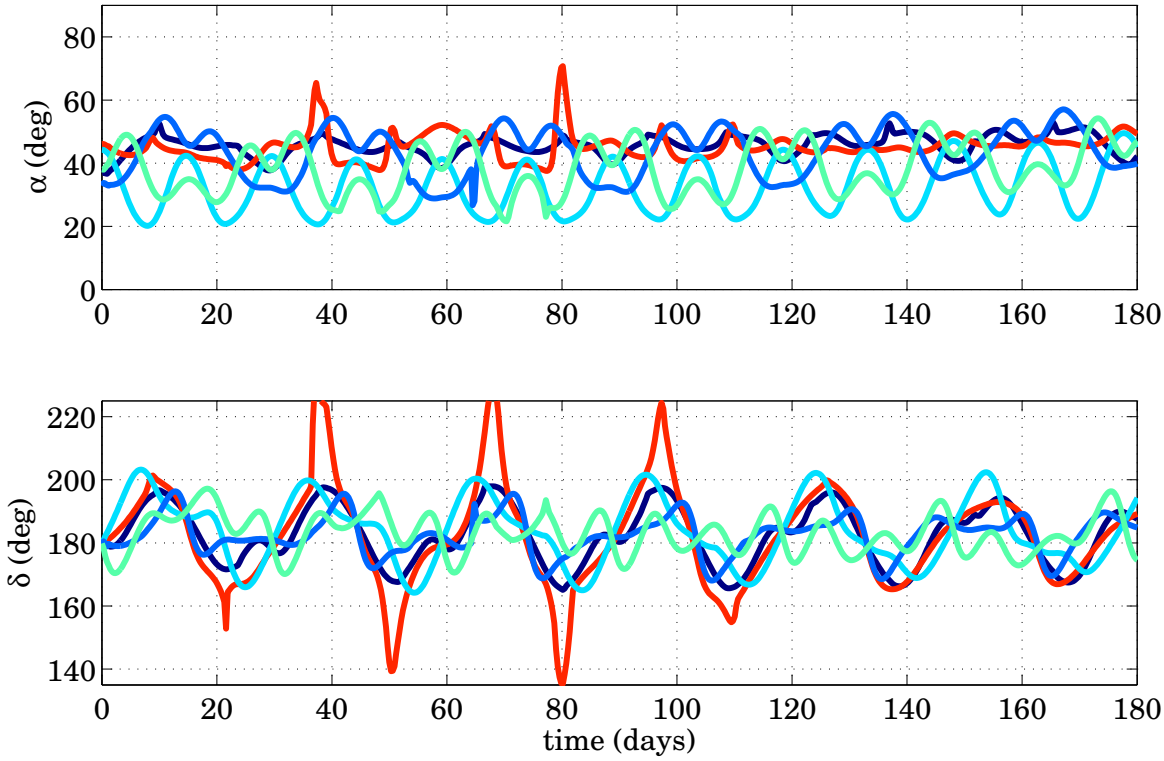


Fig. 6.7: Pitch and clock angles for the five sample orbits from the thirty-six-month simulation; only the first 180 days are plotted.

7. FLIGHT-PATH CONTROL FOR SOLAR SAIL SPACECRAFT

Designing reference trajectories and understanding their fundamental dynamics are necessary first steps toward flying sailcraft in dynamically complicated regimes, such as an environment simultaneously influenced by the gravity of multiple bodies. However, the existence of a reference orbit alone is insufficient for flight. The spacecraft must be controlled to follow the reference trajectory. Regardless of the accuracy attained for a path, or the precision of the numerical method employed, no spacecraft exactly follows the intended path. Model uncertainty, mis-modeling, and model fidelity associated with both the space vehicle and its environment all contribute to perturb the spacecraft trajectory. Furthermore, unless the equations of motion are analytically solvable, any numerical strategy is limited and only solves the equations of motion approximately. When the orbit is unstable, true for an offset lunar solar sail orbits as determined in this investigation, the sailcraft must be controlled to the follow the reference path; otherwise, a slight deviation from the path often leads to a trajectory that rapidly diverges from the desired course. Three families of strategies for sailcraft flight-path control are examined in this Chapter.

The five sample trajectories introduced in Section 5.2 are employed to test the various controllers. In this analysis, a finite-difference method (FDM) is employed to generate reference trajectories. It is presumed that if a controller maintains an actual trajectory to follow a reference path that is generated by an FDM in a CR3B regime, it may lend insight into designing a controller for a more realistic, highly accurate trajectory modeled in a higher-fidelity ephemeris regime as well as an expanded model that incorporates other errors.

7.1 Solar sail attitude in an inertial frame

For all of the control strategies in this investigation, the goal is an attitude that remains fixed inertially over a specified time interval. Thus, the orientation of the spacecraft is best represented as a set of angles measured with respect to an inertially fixed frame. Recall that the acceleration contributed by the sail is evaluated as

$$\mathbf{a}_s(t) = \frac{a_c}{a^*}(\hat{\boldsymbol{\ell}}(t) \cdot \hat{\mathbf{u}})^2 \hat{\mathbf{u}} \tag{7.1}$$

In terms of coordinates defined for the inertial frame, I, the sunlight vector can be written as

$$\hat{\boldsymbol{\ell}}(t) = \cos(t - \Omega t)\hat{\mathbf{X}} + \sin(t - \Omega t)\hat{\mathbf{Y}} + 0\hat{\mathbf{Z}} \tag{7.2}$$

where the normalized rotation rate of the Earth–Moon frame with respect to the inertial frame is equal to one, i.e., $\|{}^I\boldsymbol{\omega}^R\| = 1$, Ω is the ratio of the sidereal to synodic periods, and time, t, remains nondimensional such that the period of primary motion in the Earth–Moon system is 2π.[1] The sail pointing vector is also expanded in terms of inertially fixed unit vectors using spherical coordinates, that is,

$$\hat{\mathbf{u}} = u_X\hat{\mathbf{X}} + u_Y\hat{\mathbf{Y}} + u_Z\hat{\mathbf{Z}} \tag{7.3a}$$

$$= \cos\phi\cos\theta\hat{\mathbf{X}} + \cos\phi\sin\theta\hat{\mathbf{Y}} + \sin\phi\hat{\mathbf{Z}} \tag{7.3b}$$

where ϕ and θ are latitude and longitude angles, respectively, defined in terms of the fixed inertial frame as illustrated in Fig. 7.1. In terms of the rotating working frame, R, the sail pointing vector is

$$\hat{\mathbf{u}} = (u_X\cos t + u_Y\sin t)\,\hat{\mathbf{x}} + (u_Y\cos t - u_X\sin t)\,\hat{\mathbf{y}} + (u_Z)\,\hat{\mathbf{z}} \tag{7.4}$$

where u_X, u_Y, and u_Z are the three inertial components of the pointing vector in Eq. (7.3). The sail acceleration from Eq. (7.1) is formulated in the working (rotating) frame using inertially defined pointing vectors,

$$\mathbf{a}_s(t) = \frac{a_c}{a^*}\left(u_X\cos(t-\Omega t) + u_Y\sin(t-\Omega t)\right)^2 \left\{ \begin{array}{c} u_X\cos t + u_Y\sin t \\ u_Y\cos t - u_X\sin t \\ u_Z \end{array} \right\} \tag{7.5}$$

[1] The sunlight vector $\hat{\boldsymbol{\ell}}(t)$ is parallel to $\mathbf{r}_4$ from Fig. 5.2.

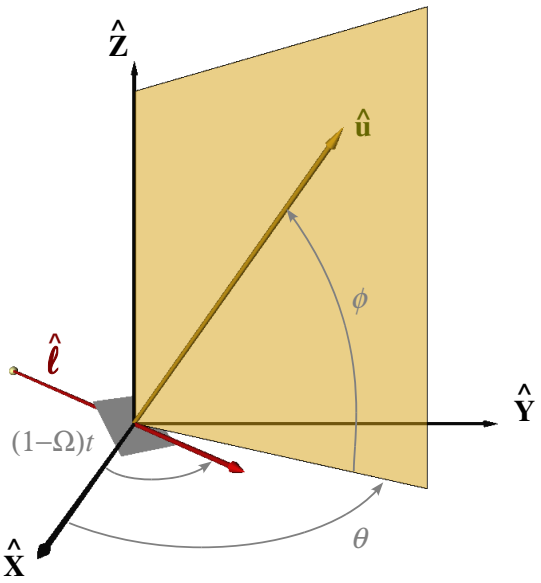


Fig. 7.1: Sail pointing vector in terms of inertially defined latitude and longitude angles (*cf.* Fig. 2.3).

The attitude profiles in inertial latitude and longitude angles for the five reference orbits from Section 5.2 are plotted in Fig. 7.2. Note that longitude angle, θ, increases by 29.12° degrees over one period for each orbit, consistent with the motion of the Earth–Moon system about the Sun over one month. The continuous profile associated with the red-orange orbit developed in the CR3B regime exhibits the greatest complexity in that its pitch angle (α from Fig. 3.11) may not be achievable with a realistic solar sail model (see Appendix A) and that its attitude rates are higher than those from the other reference orbits. However, it is included in this analysis to demonstrate the control schemes.

7.2 Turn-and-hold control strategies

The resulting solutions from the augmented finite-difference method described in Section 3.3 and the collocation solutions from Ozimek, Grebow, and Howell [30, 129] couple a trajectory that meets path constraints with an attitude profile that changes

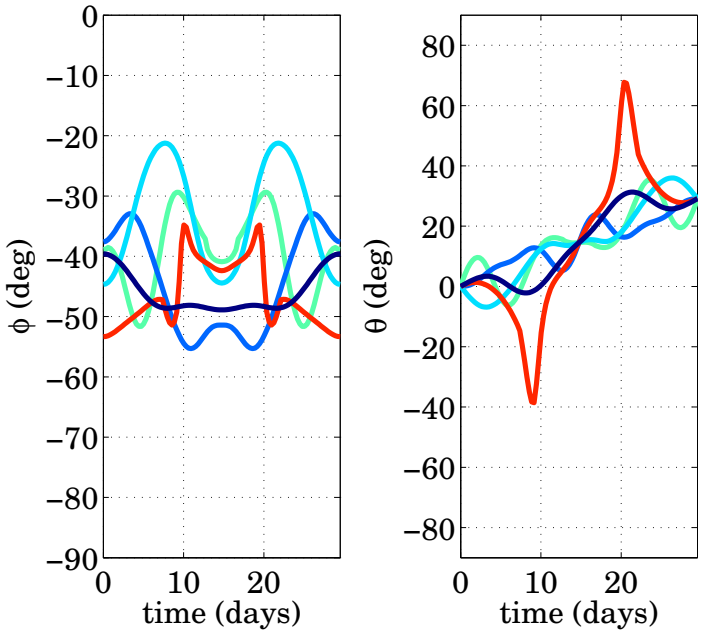


Fig. 7.2: Inertially defined latitude, ϕ, and longitude, θ, angles for the three reference orbits from Section 5.2.

over time. These attitude profiles are continuously varying because the algorithm used in their design did not constrain the attitude to be constant over the period of the orbit or even along shorter segments of the orbit. The dynamics of solar sails in this regime is not well-understood, and any constraint on the attitude, especially one that required an inertially fixed attitude, might restrict or prevent the numerical algorithms from yielding solutions.

With reference trajectories now available, new strategies may be employed to control a sailcraft along various reference paths. These approaches are employed as "turn-and-hold" strategies whereby the vehicle remains in an inertially fixed orientation for some predetermined duration before reorienting to a new attitude. These turn-and-hold strategies employ numerical techniques to solve two-point boundary value problems. In essence, a discrete controller is applied to a continuous system.

The definitions of some terms used in the numerical formulations are specified for clarity. A future state along the reference path is selected as a *target*. The portion of

the trajectory between the present state and the target is called an *arc*. An arc may be decomposed into one or more *segments* or *subarcs*, where the boundaries of a subarc occur when a turn is executed. Thus, the length of an arc is the number of turns (segments or subarcs) multiplied by the time between turns. The length of a subarc is measured by the amount of time between the turns that define its boundaries. The time at which a turn is executed is denoted an *event*. Because the reference paths from the FDM are reported at discrete intervals, an Akima cubic-spline [192] is employed to interpolate the reference attitude profile. The set of orientations, as well as their associated event times, necessary to achieve the target state at the end of the arc is labeled a turn *sequence* or *profile*.

Along a particular subarc, the sail attitude is fixed (as a consequence, the attitude is time-varying along the entire arc segment in the Earth–Moon frame). The attitude can be expressed in terms of latitude, ϕ_i, and longitude, θ_i, angles for subarc i. For convenience, $\phi_i \equiv \phi(t_i)$ and $\theta_i \equiv \theta(t_i)$. The initial guess for the latitude, ϕ_i, and the longitude, θ_i, for each subarc are the average values from the continuous values of ϕ and θ generated over arc segment i from the reference solution. Two strategies for averaging ϕ and θ are employed. For the first, a trapezoidal integration from t_i to t_{i+1} is calculated numerically to generate an initial guess for the fixed latitude and longitude values for an attitude held from t_i to t_{i+1},

$$\phi_i^0 = \frac{1}{t_{i+1} - t_i} \int_{t_i}^{t_{i+1}} \phi(t) dt \tag{7.6a}$$

$$\theta_i^0 = \frac{1}{t_{i+1} - t_i} \int_{t_i}^{t_{i+1}} \theta(t) dt \tag{7.6b}$$

For the second strategy, an average is calculated simply based on the values of ϕ and θ at t_i and t_{i+1}, that is,

$$\phi_i^0 = \frac{\phi(t_{i+1}) + \phi(t_i)}{t_{i+1} - t_i} \tag{7.7a}$$

$$\theta_i^0 = \frac{\theta(t_{i+1}) + \theta(t_i)}{t_{i+1} - t_i} \tag{7.7b}$$

Like the reference path, the reference attitude profile from the FDM solution is reported at discrete time intervals that are not necessarily coincident with the epoch

associated with a turn. Again, an Akima cubic-spline [192] is employed to interpolate the reference attitude profile. The results from the two sets of averages that are delivered by Eqs. (7.6) and (7.7) are nearly the same. However, a process initialized by one average or the other may be more likely to converge, depending on the application.

The attitude profiles from Fig. 7.2 averaged using Eqs. (7.7) over two-day segments appear in Fig. 7.3. While the latitude profile, ϕ, is periodic, the longitude profile, θ,

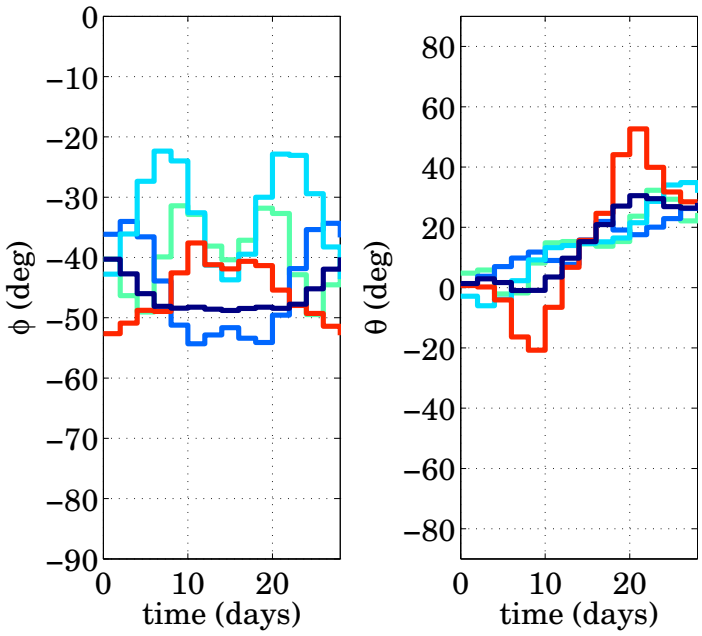


Fig. 7.3: Averaged values of latitude and longitude for two-day segments.

includes a secular component (as appears in Fig. 7.2) that must not be allowed to modulate between over 360° in a numerical algorithm. Longitude is unwrapped by changing absolute jumps greater than or equal to π to their 2π complement (MATLAB®'s `unwrap` function is employed for this purpose [123]).

The criteria for the control schemes are based on the goal to minimize the distance between the trajectory at some future time (or times) and a target state (or set of states) along a reference path at that future time(s). Ideally, this distance is zero in both position and velocity. Similar to the linearization for the FDM problem in Eq. (3.21), the nonlinear dependence of this targeting constraint on a set of

control parameters is approximated with a Taylor expansion; higher-order terms are neglected. The linear approximation appears in the form

$$\mathbf{F}(\mathbf{X}^{j+1}) \approx \mathbf{F}(\mathbf{X}^{j}) + \frac{\partial \mathbf{F}(\mathbf{X}^{j})}{\partial \mathbf{X}^{j}} \left(\mathbf{X}^{j+1} - \mathbf{X}^{j}\right) \tag{7.8}$$

where $\mathbf{X}$ is some vector combination of the controls available to deliver the appropriate attitude (e.g., attitudes necessary for targeting states along the reference trajectory, intermediate positions and velocities), $\mathbf{F}(\mathbf{X})$ is a set of targeting constraints that depends on the control strategy and dynamical environments, and $\frac{\partial \mathbf{F}(\mathbf{X}^{j})}{\partial \mathbf{X}^{j}}$ is the Jacobian relating the partial derivatives of the constraints to the available controls. The superscript "j" is an iteration index. Equation (7.8) is rearranged under the assumption that $\mathbf{X}^{j}$ is in the neighborhood where $\mathbf{F}(\mathbf{X}^{j+1}) = \mathbf{0}$, that is,

$$\mathbf{X}^{j+1} = \mathbf{X}^{j} + \delta\mathbf{X} \tag{7.9}$$

and

$$\delta\mathbf{X} = -\left[\frac{\partial \mathbf{F}(\mathbf{X}^{j})}{\partial \mathbf{X}^{j}}\right]^{+} \mathbf{F}(\mathbf{X}^{j}) \tag{7.10}$$

and the superscript "+" indicates a generalized inverse operation, or pseudoinverse.

The type of pseudoinverse depends on the problem formulation. If the number of elements in $\mathbf{F}$ is greater than the number of elements in $\mathbf{X}$ (i.e., the Jacobian contains more rows than columns), then the pseudoinverse leads to a least-squares fit of the parameters in $\mathbf{X}$ to the set of constraints in $\mathbf{F}$. If the procedure converges on a solution, the pseudoinverse is zero and no change from $\mathbf{X}^{j}$ to $\mathbf{X}^{j+1}$ will occur. If the number of elements in both vectors is identical (i.e., the Jacobian is square), a unique solution to Eq. (7.10) results. Finally, if the number of elements in $\mathbf{X}$ exceeds the number of elements in $\mathbf{F}$ (i.e., the Jacobian contains more columns than rows), the problem is under-constrained and an infinite number of solutions to Eq. (7.10) exist. The solution is converged when the norm of $\delta\mathbf{X}$ is zero, to within some acceptable tolerance.

For all iterative processes, convergence is not guaranteed, even in the presence of a supposedly good initial guess. However, for the cases examined in this investigation,

a good initial guess originating from the reference trajectories generated by the FDM is sufficient to initialize a control scheme that allows the spacecraft to track the corresponding reference trajectory.

The first implementation of the turn-and-hold controller employs a least-squares fit such that the optimal values of ϕ and θ are selected to minimize the difference between the propagated path and the reference path. The second implementation employs a collocation-based two-point boundary value problem (TPBVP) solver that delivers three turns per segment such that the positions and velocities at the boundaries of the segment match the reference trajectory. A third implementation exploits multiple-shooting techniques for a system that employs more controls than targeting constraints.

7.3 Least-squares implementation

The least-squares method is a classical approach for over-determined, or over-constrained, systems by minimizing the sum of the squares of the residuals between some estimated values and its desired values. With two controls (ϕ and θ at the beginning of the segment) and at least six scalar constraints along the segment (positions and velocities at future times), a sail cannot supply sufficient control with a single, fixed attitude over the entire the segment to target a future state. However, by employing a least-squares solution, the sailcraft can get *close* to its target state at some future time along the path.

Given an initial guess for ϕ_i and θ_i, a trajectory is propagated along with a state-control transition matrix, $\mathbf{\Psi}(t_{i+1}, t_i)$, from Eq. (2.49), originating at time t_i until time t_{i+1}. The desired final state is compared to the propagated result to determine the error. The state-control transition matrix maps deviations in ϕ_i and θ_i to deviations

in the state vector at the end of the segment. This information is reflected in the following relationship,

$$\left\{ \begin{array}{c} \delta x_{i+1} \\ \vdots \\ \delta \dot{z}_{i+1} \end{array} \right\}_{6\times 1} = \mathbf{\Psi}_{6\times 2}(t_{i+1}, t_i) \left\{ \begin{array}{c} \delta\phi_i \\ \delta\theta_i \end{array} \right\}_{2\times 1} \tag{7.11}$$

where

$$\mathbf{\Psi}(t_{i+1}, t_i) = \left[\begin{array}{cc} \frac{\partial x_{i+1}}{\partial \phi_i} & \frac{\partial x_{i+1}}{\partial \theta_i} \\ \vdots & \vdots \\ \frac{\partial \dot{z}_{i+1}}{\partial \phi_i} & \frac{\partial \dot{z}_{i+1}}{\partial \theta_i} \end{array} \right] \tag{7.12}$$

and $\frac{\partial x_{i+1}}{\partial \phi_i}$ is the element of $\mathbf{\Psi}(t_{i+1}, t_i)$ that maps a deviation in ϕ_i to a deviation in x_{i+1}. The size of each matrix or vector is indicated by a subscript.

The formulation in Eq. (7.11) does not target acceleration states directly; therefore, it can sometimes be advantageous to target multiple intermediate positions and velocities at future times on the interior of an arc bounded by the initial state and the target state. Time between the intermediate target epochs does not need to be evenly spaced. Little extra computational cost is required with the incorporation of these interior states into the least-squares algorithm. The target state is labeled $t_{i+n/n}$, and is equivalent to t_{i+1} (the time of the next turn) in Fig. 7.4; the number of intermediate target states along the arc is labeled n. Note that the target state at the end of the arc, $t_{i+n/n}$, is included in the set of intermediate states for this least-squares algorithm. In the figure, the reference trajectory is represented with a black curve and the path resulting from a simulation incorporating the initial guesses for ϕ_i and θ_i is in red. The residuals between the simulated trajectory and the reference path at the "future" epochs are indicated with gray arrows. No residual exists at t_i because the controller cannot compensate for initial errors in the path, only future errors. A best-fit solution for ϕ_i and θ_i that is determined by incorporating these extra, intermediate targeting constraints can generally track the reference trajectory

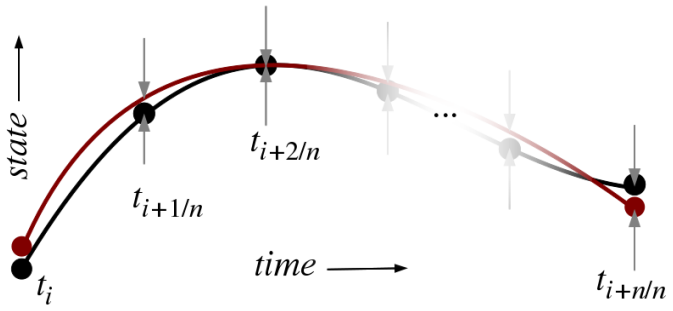


Fig. 7.4: A "turn-and-hold" segment (red) fit through n intermediate target points along a reference path (black). The attitude is held along the trajectory from t_i to $t_{i+n/n}$.

more closely at t_{i+1} (or $t_{i+n/n}$) and for an extended set of segments. The relationships in Eq. (7.11) are now written as

$$\left\{ \begin{array}{c} \delta x_{i+1/n} \\ \vdots \\ \delta \dot{z}_{i+n/n} \end{array} \right\}_{6n\times 1} = \mathbf{\Psi}_{6n\times 2}(t_{i+1}, t_i) \left\{ \begin{array}{c} \delta\phi_i \\ \delta\theta_i \end{array} \right\}_{2\times 1} \tag{7.13}$$

where

$$\mathbf{\Psi}(t_{i+1}, t_i) = \left[\begin{array}{cc} \frac{\partial x_{i+1/n}}{\partial \phi_i} & \frac{\partial x_{i+1/n}}{\partial \theta_i} \\ \vdots & \vdots \\ \frac{\partial \dot{z}_{i+n/n}}{\partial \phi_i} & \frac{\partial \dot{z}_{i+n/n}}{\partial \theta_i} \end{array} \right] \tag{7.14}$$

As mentioned, a solution to Eq. (7.13) for $\{\delta\phi_i \ \delta\theta_i\}^{\mathrm{T}}$ arises from a weighted least-squares implementation, that is,

$$\left\{ \begin{array}{c} \delta\phi_i \\ \delta\theta_i \end{array} \right\} = (\mathbf{\Psi}^{\mathrm{T}}\mathbf{W}\mathbf{\Psi})^{-1}\mathbf{\Psi}^{\mathrm{T}}\mathbf{W} \left\{ \begin{array}{c} \delta x_{i+1/n} \\ \vdots \\ \delta \dot{z}_{i+n/n} \end{array} \right\} \tag{7.15}$$

where $\mathbf{W}$ is a diagonal weighting matrix that balances the corrections between the position and velocity components and the mapping times in the numbered subscripts on $\mathbf{\Psi}$ have been dropped for convenience. The weighting matrices are typically the

identity matrix. Equation (7.15) is equivalent to the least-squares update associated with Eq. (7.10). The solution from Eq. (7.15) is used to update ϕ_i and θ_i, whereby a new path is simulated and compared to the reference trajectory in terms of $\left\{\delta x_{i+1/n} \ \ldots \ \delta z_{i+n/n}\right\}^{\mathrm{T}}$. The process is iterated until the values of $\{\delta\phi_i \ \delta\theta_i\}^{\mathrm{T}}$ are zero, to within some tolerance. Because the correction is based on a least-squares solution, the elements on the left side of either Eqs. (7.11) or (7.13) (or, alternatively, within the braces on the right side in Eq. (7.15)) may never converge to zero, but the sum of the squares of these elements is some minimum value. With a fixed attitude at t_i and multiple intermediate states, acceleration is targeted implicitly at $t_{i+n/n}$.

Selection of the number of intermediate target states, n, and the length of time between t_i and t_{i+1} does impact the resulting controlled trajectory. A controller that targets one or more states with only two controls, rarely, if ever, produces a trajectory that coincides with a reference path at the time of the scheduled turn at the end of each arc. However, if the controlled trajectory is sufficiently close to the reference path at the points when the angles are modified and the orientation is shifted, it is likely that the controlled solution will adequately track the reference path. Unfortunately, solutions from the least-squares approach often diverge for a given reference path because less control authority is available than is required to target future a state (or states). Therefore, a trial-and-error process is employed to uncover combinations of time between turns and the number of intermediate target epochs required to track a particular orbit. Given a particular combination of days between orientation changes and intermediate target epochs during each arc, the intervals (up to 36 months) that the reference is successfully tracked appear in Figs. 7.5 through 7.9 for the five reference paths from Section 5.2. For this least-squares simulation, the initial guesses for the latitude and longitude for a particular arc are computed via an average from trapezoidal integration, Eq. (7.6).

An example is useful for interpreting the results in Figs. 7.5 through 7.9. Consider the green box in Fig. 7.5 containing the number 4 and corresponding to 2 days between turns and 6 intermediate target epochs. Given the initial conditions from

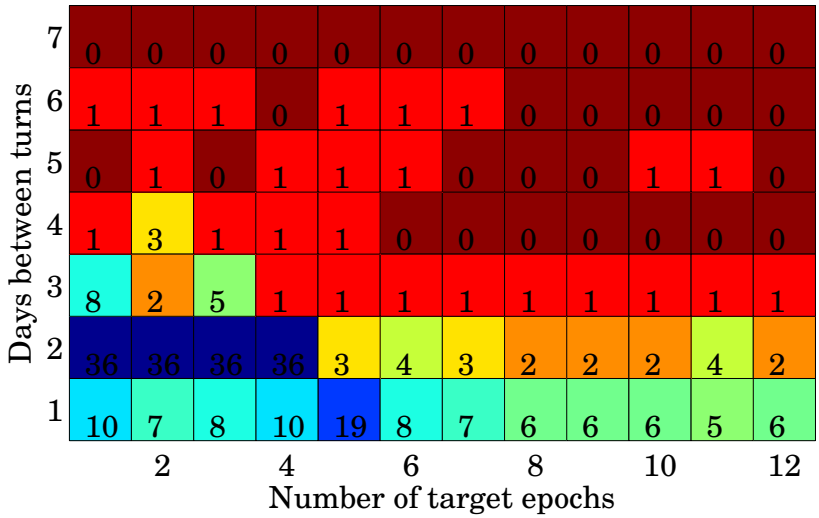


Fig. 7.5: Number of months the least-squares controller tracks the aqua reference path appearing in Fig. 5.6.

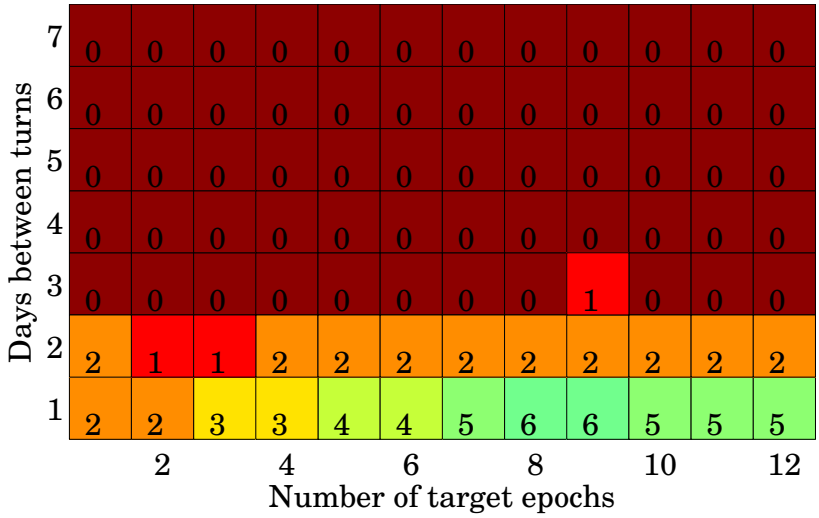


Fig. 7.6: Number of months the least-squares controller tracks the royal-blue reference path appearing in Fig. 5.6.

the aqua reference path, a simulation is structured to employ 2 days between orientation changes. Each 2-day arc employs 6 interior targets. Because each final target state coincides with a turn, the arcs in this least-squares application are composed of only one subarc. The number 4 indicates the number of months that the reference

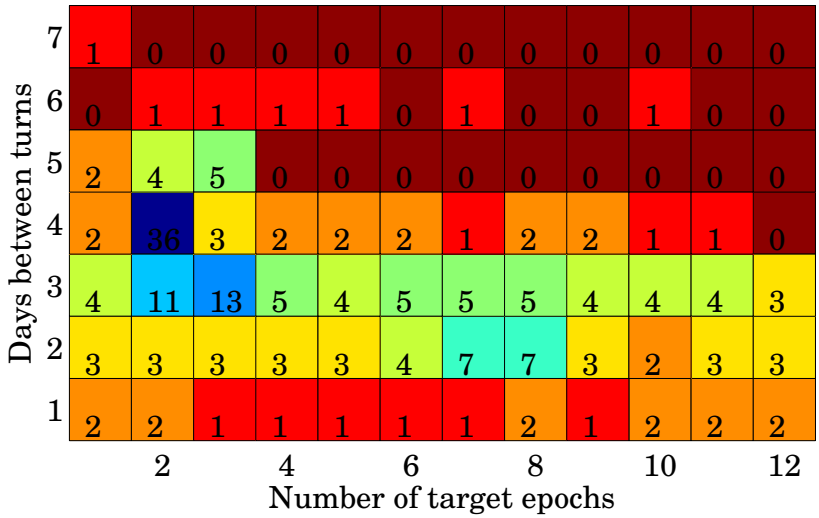


Fig. 7.7: Number of months the least-squares controller tracks the cyan reference path appearing in Fig. 5.6.

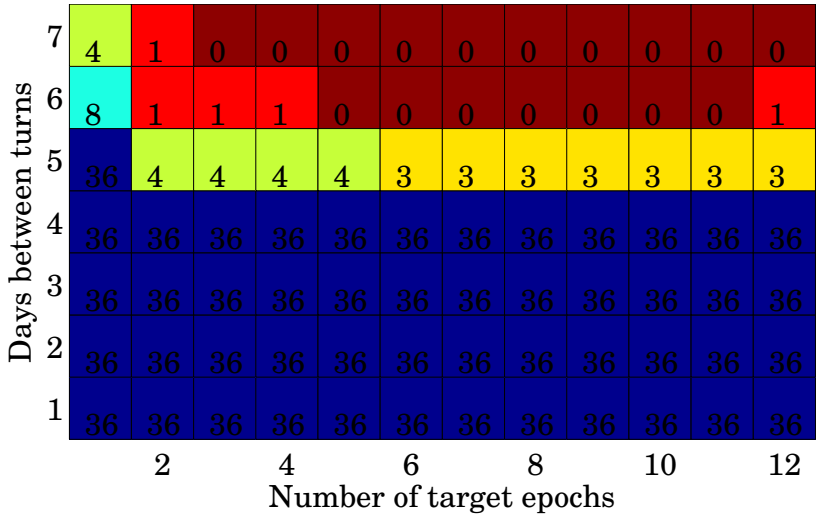


Fig. 7.8: Number of months the least-squares controller tracks the red-orange reference path appearing in Fig. 5.6.

trajectory is successfully tracked (in that the least-squares algorithm converges on a solution in the vicinity of the reference path) for this combination of arc length and intermediate target states. The color of each box in Figs. 7.5 through 7.9 corresponds to the number of months in that box; a jet scale is employed such that long tracking

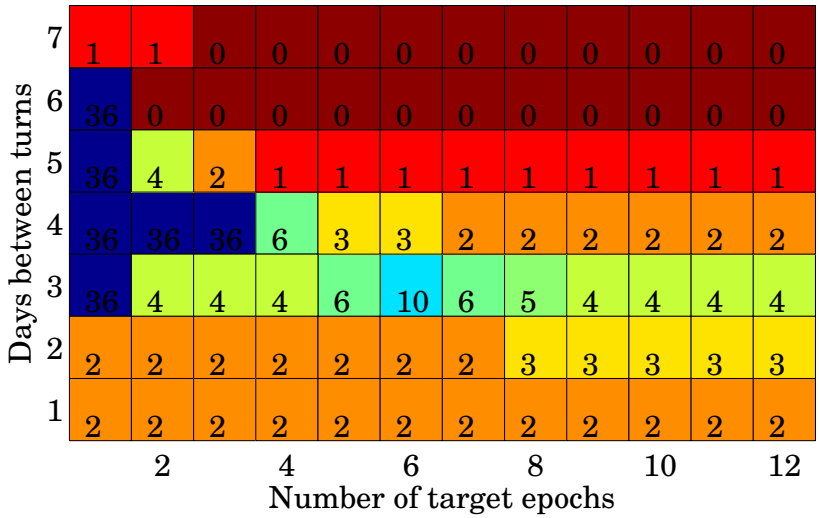


Fig. 7.9: Number of months the least-squares controller tracks the dark-blue reference path appearing in Fig. 5.6.

times are blue (e.g., 36 months) and short tracking times are red (e.g., 0 months). In these simulations, no errors are incorporated into the control scheme and all turns are presumed to occur instantaneously. Additionally, no path constraints, such as minimum elevation angle, are incorporated into the algorithm, and each intermediate target state along an arc is weighted equal to the other. Of course, the reference could already meet some or all of these constraints. Because the available control (2 attitude angles) is insufficient to achieve the over-constrained target, the controlled path will not match the reference path at the end of the arc, even in an ideal scenario. Incorporating errors may reduce the ability of the controller to track the reference path for a given combination of arc length and intermediate target epochs, but the formulation of this least-squares approach is not modified with the inclusion of these errors. Simply, one attitude per combination is determined that will best pass through the interior target states along an arc for the least-squares algorithm.

Perhaps surprisingly, some variations of this over-constrained, or under-controlled, least-squares implementation track the sample reference orbits for multiple months. Of particular focus are the combinations that track their reference paths for the full

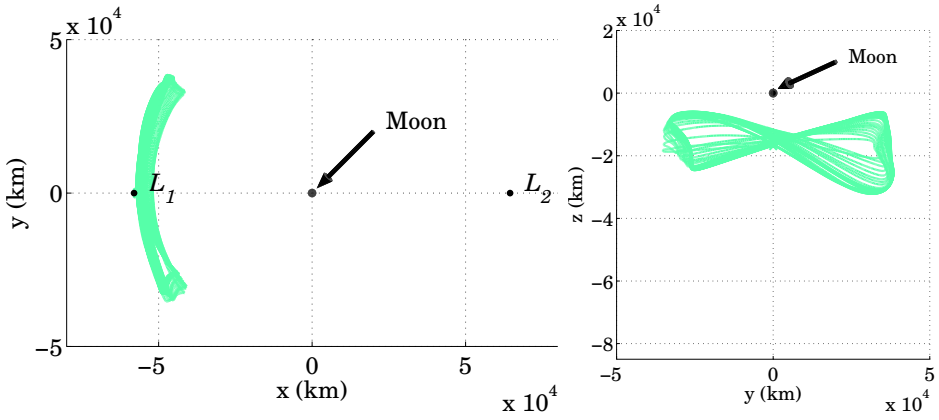
36-month time-frame or, at least, the greatest number of revolutions for a particular reference orbit. A cycle is defined as one revolution of the Sun about the Earth–Moon system, or a 29.5-day synodic month. Recall that the periods of the sample trajectories introduced in Section 5.2 are also one synodic month. Some successful combinations are immediately apparent in Figs. 7.5 through 7.9 (i.e., the dark-blue boxes).

The least-squares approach successfully tracks the aqua path from Section 5.2 (denoted Orbit #1 in Table 5.3) for 36 months with four combinations of days between turns (DBT) and intermediate target epochs, n, as apparent in Fig. 7.5. These four combinations are all based on a duration of 2 DBT, and require 1, 2, 3, or 4 interior target epoch along each respective 2-day arc. The minimum elevation angle and maximum excursion from the reference path for these four combinations are listed in Table 7.1. Recall that the reference paths are generated with a minimum

Table 7.1: Properties of the controlled aqua path

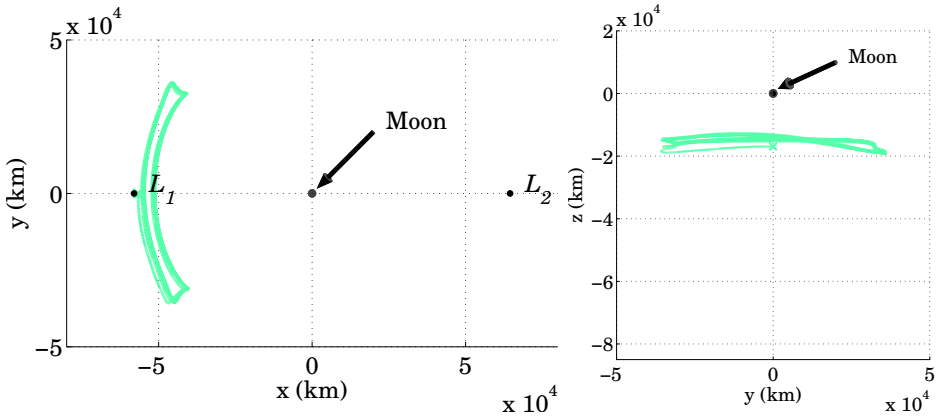
DBT$\times n$	2×1	2×2	2×3	2×4
Min Elev. Ang.	4.4°	11.1°	11.7°	11.1°
Max Excurs. (km)	13437	6055	5350	6203

elevation angle-constraint of 15°. Although the 2×1 combination does not diverge from the reference path, it does not track as closely as the other combinations. The top and side view from the perspective of an observer at L_2 of the trajectory developed with the 2×1 combination corresponding to the aqua reference path from Fig. 5.6 appears in Fig. 7.10. While the reference path is symmetric across the xz plane, this controlled orbit appears to twist about an imaginary axis in the y direction. The 2×3 combination more successfully replicates the reference path from Fig. 5.6, as demonstrated in Fig. 7.11.



(a) xy view (b) yz view

Fig. 7.10: Two views of the controlled path of the aqua orbit subject to a 2×1 least-squares flight-path control scheme.



(a) xy view (b) yz view

Fig. 7.11: Two views of the controlled path of the aqua orbit subject to a 2×3 least-squares flight-path control scheme.

The least-squares algorithm successfully tracks the royal-blue path in the vicinity of L_2 for, at most, 6 months before the path diverges from the reference trajectory such that the spacecraft is no longer recoverable, as apparent in Fig. 7.6. A plot of the controlled trajectory corresponding to the 1×8 combination from Fig. 7.6 appears in

Fig. 7.12. The trajectory originates at the "x" located under L_2 and terminates near this marker, and it is clear that terminal location of the path is not aligned with the rest of the trajectory.

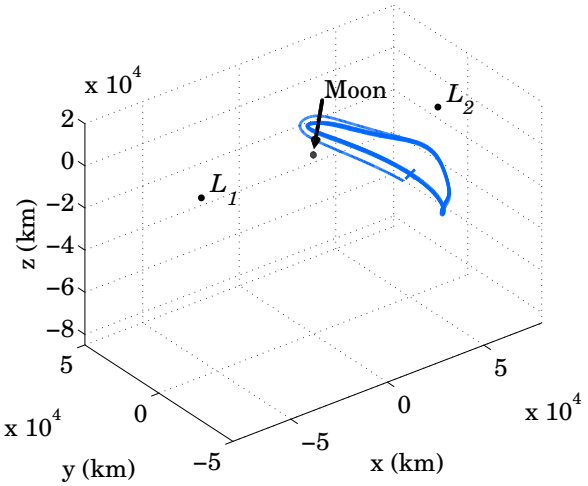


Fig. 7.12: The royal-blue path is tracked with the least-squares approach for a maximum of 6 months.

Only one combination of DBT and n intermediate target points (i.e., 4 DBT and $n = 2$) leads to a successfully tracked cyan orbit below the Moon for 36 months, as apparent in Fig. 7.7. Unfortunately, at one point along the orbit, the elevation is 4.8° *below* the horizon of the lunar south pole. The maximum excursion is also greater than 42000 km from the reference path. This path appears in Fig. 7.13.

In contrast to the other reference paths, the red-orange orbit from Fig. 5.6 is successfully tracked by most variations of the least-squares algorithm, as observed in Fig. 7.8. The maximum excursions and minimum elevation angles associated with the combinations that track the red-orange reference path for 36 months, that is, all combinations based on 1, 2, 3, or 4 days between turns as well as one case based on 5 days between turns and 1 target epoch, appear in Fig. 7.14. When the time between turns is 1 day, the algorithm tracks this reference orbit exceptionally well for all states reported by the integration algorithm (i.e., effectively a continuous solution). For example, the reference path is tracked to within 860 km and above 15.1° for the

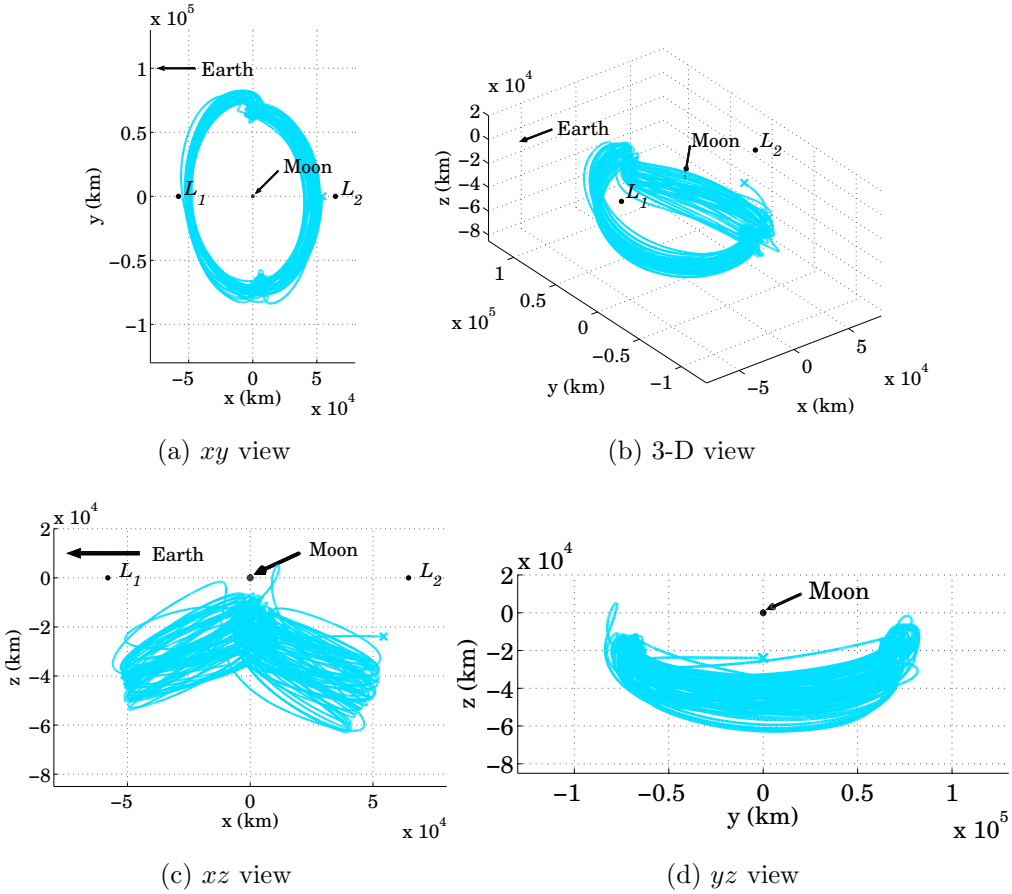
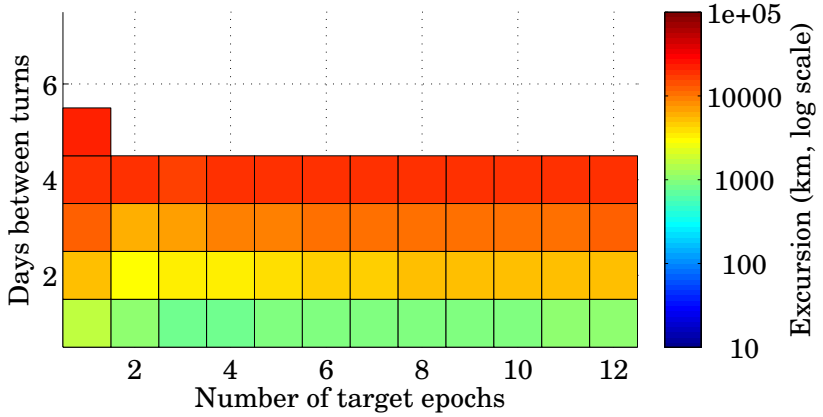


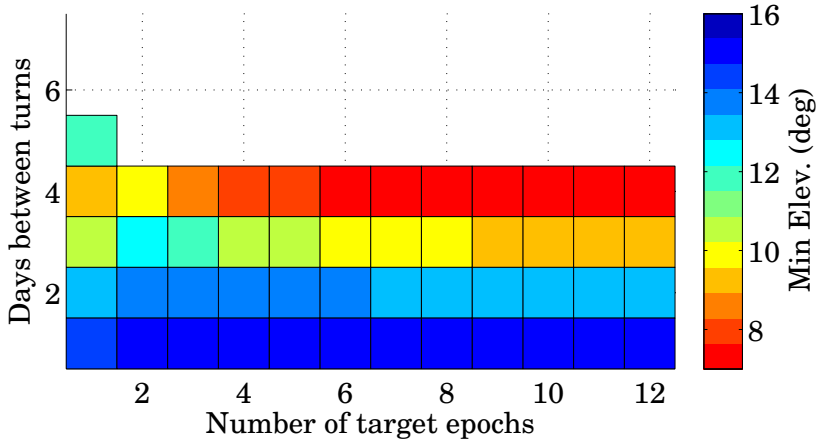
Fig. 7.13: The only variation of the least-squares approach not to diverge from the cyan path. This trajectory corresponds to a combination consisting of 4 days between turns and 2 intermediate target points.

1×3 and 1×4 cases. The worst tracking results for a controlled path that does not diverge for 36 orbits are observed in the 5×1 case (21393 km and 11.7°); the lowest elevation angle for a trajectory that did not diverge is 7.1° for the 4×12 case (with a maximum excursion of 18776 km). The paths of these orbits appear in Figs. 7.15(a) and 7.15(b), respectively.

Finally, the dark-blue path is successfully tracked by 6 combinations of days between turns and intermediate target epochs. The properties of these 6 orbits are listed in Table 7.2. A sample of the converged paths, the 6×1 case, appears in Fig. 7.16.



(a) Maximum excursions

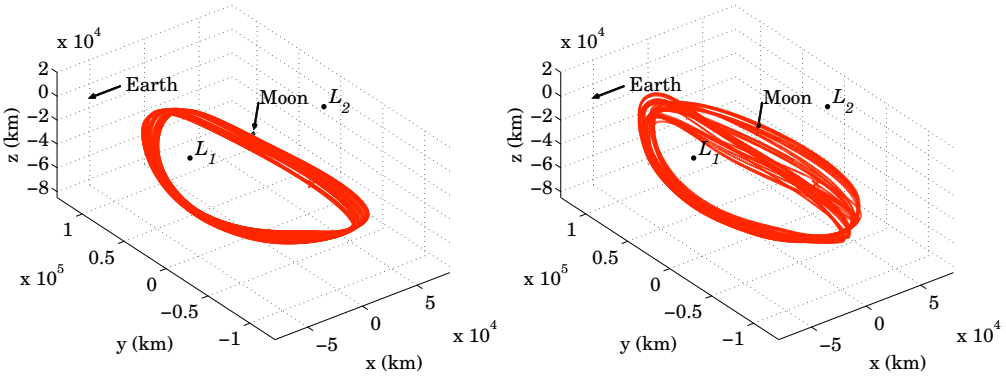


(b) Minimum elevation angles

Fig. 7.14: Maximum excursions and minimum elevation angles for the controlled paths of the least-squares algorithm that track the red-orange orbit from Fig. 5.6 for 36 months.

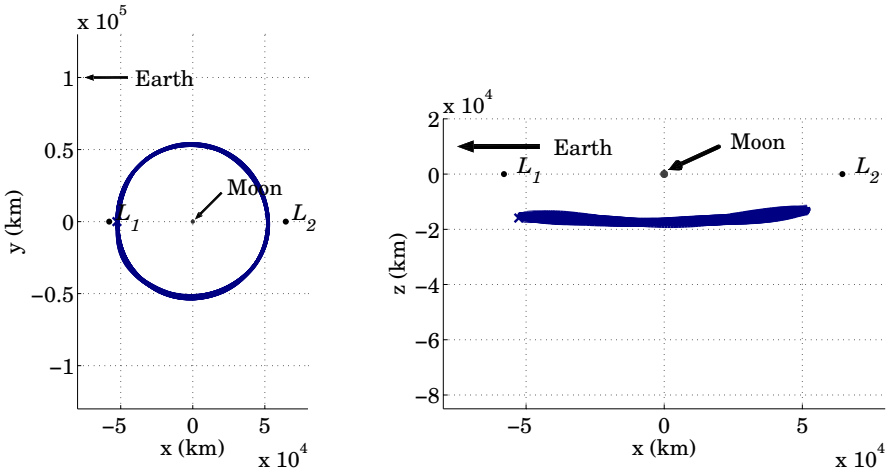
Table 7.2: Properties of the controlled dark-blue path

DBT$\times n$	6×1	5×1	4×1	4×2	4×3	3×1
Min Elev. Ang.	11.0°	12.0°	11.8°	12.1°	12.6°	9.7°
Max Excurs. (km)	9600	4366	3737	3329	4265	5706



(a) 5×1 case (b) 4×12 case

Fig. 7.15: Paths corresponding to (a) the maximum excursion and (b) the minimum elevation for all controlled trajectories corresponding to the red-orange reference path.



(a) xy view (b) xz view

Fig. 7.16: The 6×1 least-squares case tracks its reference path.

7.4 Three-Turn Two-Point Boundary Value Problem with BVP6C

The disadvantage in the least-squares formulation of the tracking controller involves the facts that (1) the number of scalar boundary conditions (six on either end of the arc, plus any intermediate target states) exceeds the number of differential equations (six) plus controls (two) and that (2) a solution for $\{\phi_i \ \theta_i\}^{\mathrm{T}}$ does not control the trajectory such that the target states at t_{i+1} are met exactly. A possible alternative is a strategy to target a position and velocity along the trajectory after exactly three turns, such that the number of controls is six and the two-point boundary value problem (TPBVP) is well-posed. Recall that t_i corresponds the time of a spacecraft orientation change, and t_{i+1}, t_{i+2}, etc. correspond to times of subsequent orientation changes. In this alternative strategy, turns implemented at t_i, t_{i+1}, and t_{i+2} (generally spaced days apart) supply sufficient control such that the positions and velocities along the compensated trajectory match the values on the reference trajectory at t_i and t_{i+3}, as illustrated in Fig. 7.17. Recall that an arc is defined as

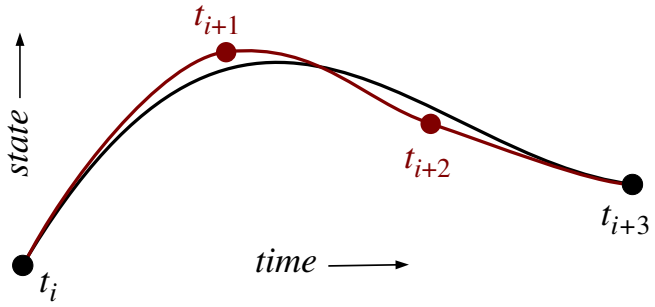


Fig. 7.17: A three-turn solution along an arc to target position and velocity at t_{i+1} given a position and velocity vectors at t_i.

the portion of the trajectory between the initial state and the target state and that turns define the boundaries of the subarcs (or segment). In the figure, the initial state is at t_i and the target state is at t_{i+1}. An arc along the reference trajectory, which supplies an initial guess for the TPBVP solver is represented in black, while the arc along the actual path that is associated with the three turns that solve the

equations of motion appears in red in the figure. In an error-free case, the controlled path aligns with the reference path at the boundaries of the arc, t_i and t_{i+3}.

Conveniently, MATLAB® supplies a suite of functions for solving TPBVPs based on collocation schemes. These three algorithms, BVP4C (most similar to the example illustrated by Fig. 3.7) [123], BVP5C [123], and BVP6C [127], algorithms differ primarily in the degree of their respective interpolating polynomials and the resulting accuracies. Both BVP4C and BVP5C are based on three- or four-stage Lobatto IIIa integration rules, respectively, and are fourth- or fifth-order accurate uniformly between the collocation nodes [123]. For BVP6C, a quintic interpolant is fit to a mono-implicit Runge-Kutta sixth-order formula for sixth-order uniform accuracy between the collocation nodes [127]. A mesh of 50 to 200 collocation nodes is typically employed when initializing any of these three algorithms for this application, and three methods employ mesh refinement to minimize the error with respect to the collocation scheme along the arc such that the mesh does not need to be uniformly spaced; node points may be added or eliminated as well.

Implementing any one of the three MATLAB® collocation functions requires an initial guess for the trajectory; a reference path supplied by the FDM is a convenient choice. Also required is an initial guess for the set of controls, that is, some initial set of sail angles (i.e., three pairs of latitude and longitude angles). Recall that the control profile for the reference trajectory is continuous and a fixed set of orientations is required for the turn-and-hold scheme. In the BVP4C, BVP5C, and BVP6C implementations, these sail angles are collected together as six unknown parameters (three turns, two angles) to be estimated by the algorithm. The attitude of the sailcraft is presumed to change at t_i, t_{i+1}, and t_{i+2} (from Fig. 7.17), and this set of six total angles represent the controls for the TPBVP. The input and output structures for each of the three BVP solvers is identical, making it simple for the user to switch between methods. For this analysis, BVP6C is employed since this algorithm possesses the best theoretical accuracy of the three methods. Fewer mesh points to represent the initial guess corresponding to the trajectory segment are required, and conver-

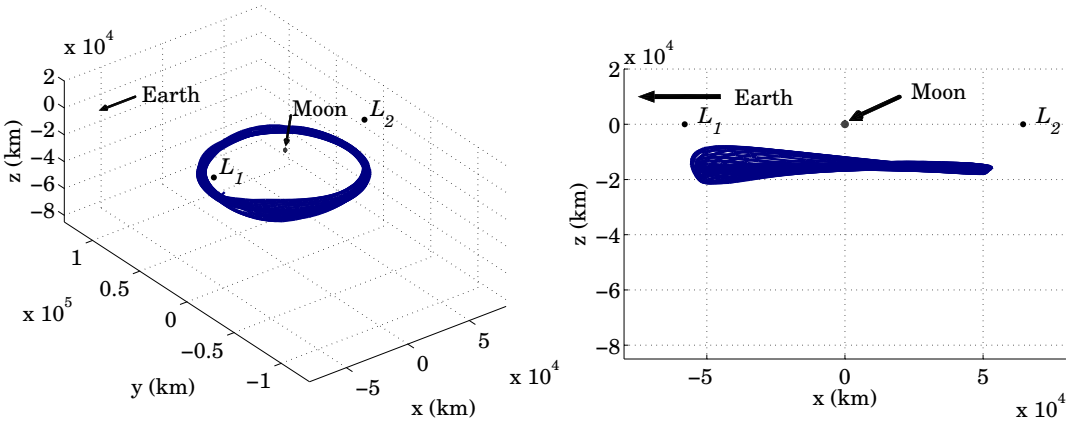
gence is typically faster than for the other two options. Nevertheless, a large number ($m \approx 151$) of mesh points within the arc bounded by $[t_i, t_{i+3}]$ are initially required, in practice, for convergence and to maintain the solution paths near the respective reference trajectories for this problem. Once a controlled trajectory is determined for the arc from t_i to t_{i+3}, the process is repeated for the next arc from t_{i+3} to t_{i+6}.

As designed, the orbit from the BVP6C algorithm matches the boundary conditions exactly. However, the path interior to t_i and t_{i+1} is only accurate to within the specified tolerances. The initial meshes for the BVP6C algorithm are "tuned" by the algorithm so that a solution converges to within the specified tolerances; in this tuning, the number of nodes and mesh spacing are adjusted through an automated mesh refinement [127]. For each sample orbit, seven different arc lengths are examined. In all cases, a 151-point initial mesh, with absolute and relative tolerances of 0.001 is attempted. When the absolute and relative tolerances are tightened, the BVP6C algorithm reports that it "cannot converge without exceeding the maximum number of allowable mesh points," and the simulation terminates. Based on the number of revolutions (i.e., synodic months) that are tracked by the BVP6C control scheme, as apparent in Table 7.3, it is clear that this method is not robust. Only the dark-blue

Table 7.3: Number of revolutions (i.e., months) the reference orbit is tracked based on a fixed arc length

		Arc length in days (3 turns per arc)						
Orbit No.	color	3	6	9	12	15	18	21
1	aqua	0.7	1.8	1.2	0.4	0	0	0.7
2	royal-blue	0	0	0.6	0	0.5	4.3	2.1
3	cyan	1.8	0	0.3	0	2.0	0.6	0.7
4	red-orange	0.2	0.6	0.3	0	3.5	0.6	0
5	dark-blue	6.3	1.2	5.5	21.6	0	36.0	36.3

reference trajectory is tracked using the BVP6C algorithm for the full 36 months of the test. For the 18-day arc-length option (corresponding to 6 days between turns), the minimum elevation angle at any point along the controlled dark-blue trajectory is 12.3° and the maximum excursion from the reference path is 10516 km. For the 21-day arc-length option (corresponding to 7 days between turns), the minimum elevation angle is 7.1° and maximum excursion is 29291 km. A plot of controlled path based on the 21-day arc-length option appears in Fig. 7.18. Recall, however, that



(a) 3-D view (b) xz view

Fig. 7.18: Path resulting from a 21-day arc length as applied to the dark-blue reference orbit

this dark-blue reference orbit is perhaps the "simplest" of the five options and that this scheme is a well-posed two-point boundary value problem: an amount of control is available, in the form of three turns to target six state elements, that is theoretically sufficient. The more "challenging" reference trajectories may require a more sophisticated control strategy.

The turns determined by the BVP6C algorithm are employed in an explicit integration scheme to assess the accuracy of the algorithm and to simulate an error-free flight condition. Incorporating the turns from the BVP6C algorithm into an explicit integration scheme results in a position and a velocity at the end of the first arc that are not aligned with the original reference path (note that the position and velocity at

the end of the arc *as reported by the BVP6C algorithm* match the target state along the reference path). Nevertheless, these positions and velocities from the integration are usually close to the target states. Not surprisingly, employing this new state as a boundary condition in the BVP6C algorithm for the subsequent arc often leads to divergence earlier than when using states from along the reference path. The number of months a particular orbit is tracked when the trajectory is propagated using an explicit integration scheme (MATLAB®'s ODE113 function with 1×10^{-13} absolute and relative tolerances) that incorporates the solutions from the BVP6C algorithm is listed for each arc length in Table 7.4. While most options in Table 7.4 diverge faster

Table 7.4: Number of months the reference orbit is tracked based BVP6C solutions that are re-propagated with ODE113

		Arc length in days (3 turns per arc)						
Orb. No.	color	3	6	9	12	15	18	21
1	aqua	0.5	1.0	1.2	0.4	0	0	0.7
2	royal-blue	0	0	0.6	0	0.5	2.4	0.7
3	cyan	0.2	0.2	0.3	0	2.0	0.6	0.7
4	red-orange	0.2	0.6	0.3	0	3.6	0.6	0
5	dark-blue	1.4	17.1	3.0	2.4	0	36.0	36.3

than their counterparts in Table 7.3, when the arc length is 6 days, the dark-blue reference is followed for 17 months when re-propagated with ODE113. This phenomenon illustrates the sensitivity of the problem. When the initial guesses for the attitude are calculated with the simple average in Eq. (7.7) instead of the integrated averages in Eq. (7.6), the results in Tables 7.3 and 7.4 are similar.

7.5 Multiple-shooting approach

A multiple-shooting approach supplies more controls than constraints for the solution of a TPBVP. In a least-squares implementation, the number of scalar controls is less than the number of constraints. Solutions delivered from a least-squares method do not reach the target states exactly. A well-posed two-point boundary value problem, where the number of controls equals the number of constraints at the end of the arc, is necessary, but not sufficient, for any three-turn control scheme to succeed using existing MATLAB functions. However, the accuracy of this scheme is limited when transferred to an explicit numerical propagation. It is possible to construct a three-turn target scheme using shooting techniques, but it is also possible to develop a guidance strategy whereby more controls than constraints do exist. A problem is formulated such that the number of controls is greater than the number of constraints, and, theoretically, an infinite number of solutions exist. In practice, three-turn schemes often possess convergence difficulties. However, for more than three turns, Eq. (7.10) is under-constrained, and convergence on viable trajectory control laws is more likely.

As mentioned in Section 3.1, shooting methods are also numerical techniques to solve two-point boundary problems. To develop a strategy that supplies more controls than are required to target a future state, a multiple-shooting scheme is employed. At an initial time, t_i, a state, $\mathbf{x}_i$, composed of position and velocity components is extracted from the reference trajectory. At some future time, t_{i+h}, another state, $\mathbf{x}_{i+h}$, is extracted from the reference trajectory.[2] This state is the target state. The complete path from t_i to t_{i+h} and is denoted as the arc. Additional, intermediate states are extracted from the trajectory at times t_{i+1}, t_{i+1}, $\ldots$, t_{i+h-1}. The vehicle is reoriented at these epochs as well as at t_i such that h turns are performed within an arc. In the intervals between these turn events, the vehicle orientation is held. Initial guesses for latitude, ϕ, and longitude, θ, angles are calculated at t_i to t_{i+h-1} using

[2]Note that the indices i and h are integers.

Eq. (7.7) and converted into unit vectors. The trajectory and a state transition matrix are numerically propagated from one epoch to the next using an explicit integration scheme, as apparent in Fig. 7.19. The state at t_{i+1} resulting from the integration is

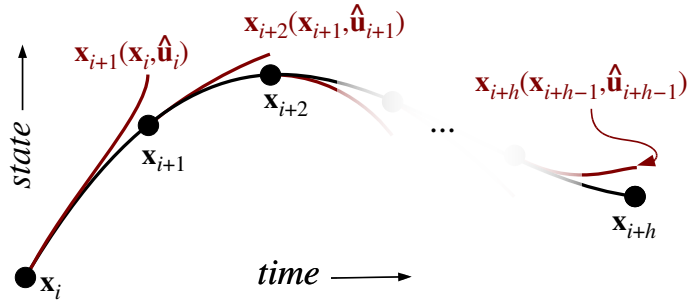


Fig. 7.19: Example of propagated subarcs extracted from a reference path. The arc is defined from t_i to t_h. The first subarc occurs between t_i and t_{i+1}.

labeled $\mathbf{x}_{i+1}(\mathbf{x}_i, \hat{\mathbf{u}}_i)$ because it is initialized with $\mathbf{x}_i$ and $\hat{\mathbf{u}}_i$, where $\hat{\mathbf{u}}$ is a unit vector describing the direction of the sail-face normal. These propagated trajectories are the subarcs and almost always diverge from their intended targets. When the subarcs are connected, they form the arc. In the figure, the black curve represents the reference path and the red curves reflect the actual path of the propagated trajectory with an initial set of fixed attitude angles, integrated from one epoch to the next for h epochs and using initial conditions extracted from the reference path. The difference between the state propagated from one epoch to the next and the state extracted from the reference path at that next epoch forms the basis for a multiple-shooting correction.

The aim of using multiple shooting for this application is to develop a continuous path from $\mathbf{x}_i$ to $\mathbf{x}_{i+h}$ by adjusting the control parameters, $\mathbf{u}_i$, $\mathbf{u}_{i+1}$, ..., $\mathbf{u}_{i+h}$,[3] and the intermediate states, $\mathbf{x}_{i+1}$, $\mathbf{x}_{i+2}$, ..., $\mathbf{x}_{i+h-1}$. Similar to the path constraints introduced for the FDM in Section 3.3, path constraints may also be incorporated into a multiple-shooting scheme. As discussed previously, the only required path constraint for a solar sail trajectory is the pitch-angle constraint, where $\alpha \leq \alpha_{\max}$ and $\alpha_{\max}$ is no greater

[3]As with the FDM approach for generating reference trajectories in Ch. 3.3, $\mathbf{u}$ is only unit length for initial guesses and converged solutions.

than 90°. This inequality constraint is again implemented via a slack variable, η_α. A second path constraint for elevation angle, where $E \geq E_{\min}$, is similarly implemented via a slack variable, η_E. These slack variables are determined in conjunction with the control variables and the intermediate states.[4] Other inequality path constraints may be similarly implemented. This investigation assumes that tracking a reference trajectory that already conforms to the specified path constraints is sufficient and minor incursions into constrained regions at times not coinciding with reorientation events is allowable. A vector comprising the control and slack variables, as well as the intermediate states, is written as

$$\mathbf{X} = \left\{ \begin{array}{c} \mathbf{u}_i \\ \eta_{\alpha,i} \\ \mathbf{x}_{i+1} \\ \eta_{E,i+1} \\ \mathbf{u}_{i+1} \\ \eta_{\alpha,i+1} \\ \vdots \\ \mathbf{x}_{i+h-1} \\ \eta_{E,i+h-1} \\ \mathbf{u}_{i+h-1} \\ \eta_{\alpha,i+h-1} \end{array} \right\} \tag{7.16}$$

Note that the control vector at the beginning of the arc, $\mathbf{u}_i$, may be adjusted, but the initial position and velocity, $\mathbf{x}_i$, are fixed. Consequently, the slack variable associated with the pitch-angle constraint is applied at t_i. Also, the control vectors in Eq. (7.16) are guaranteed to be unit length only when $\mathbf{X}$ represents an initial guess or a converged solution and not necessarily at any other iteration during the numerical solution process.

[4] In the current formulation, the path constraints only apply at times coincident with the intermediate states (the pitch-angle constraint is also applied at t_i). These constraints may be applied anywhere, including at other epochs between the intermediate states. However, additional path constraints anywhere along the orbit may impede convergence.

The elements in $\mathbf{X}$ are used to satisfy a vector equation such that all elements are theoretically zero for a converged solution. For this lunar south pole application, this vector is represented as

$$\mathbf{F}(\mathbf{X}) = \left\{ \begin{array}{c} \mathbf{u}_i^{\mathrm{T}}\mathbf{u}_i - 1 \\ \cos\alpha_{\max} - (\hat{\boldsymbol{\ell}}_i \cdot \mathbf{u}_i) + \eta_{\alpha,i}^2 \\ \mathbf{x}_{i+1}(\mathbf{x}_i, \mathbf{u}_i) - \mathbf{x}_{i+1} \\ \sin E_{\min} + \frac{z_{i+1}+R_m}{A_{i+1}} + \eta_{E,i+1}^2 \\ \vdots \\ \mathbf{u}_{i+h-1}^{\mathrm{T}}\mathbf{u}_{i+h-1} - 1 \\ \cos\alpha_{\max} - (\hat{\boldsymbol{\ell}}_{i+h-1} \cdot \mathbf{u}_{i+h-1}) + \eta_{\alpha,i+h-1}^2 \\ \mathbf{x}_{i+h}(\mathbf{x}_{i+h-1}, \mathbf{u}_{i+h-1}) - \mathbf{x}_{i+h} \end{array} \right\} = \mathbf{0} \tag{7.17}$$

The constraint $\mathbf{u}_i^{\mathrm{T}}\mathbf{u}_i - 1 = 0$ ensures that the control vector is unit length at the first epoch (similar constraints are employed for the remaining epochs).[5] The pitch-angle constraint for this multiple-shooting scheme is the same as developed in Eqs. (3.40) and (3.41). Similarly, an elevation-angle constraint, identical to Eqs. (3.38) and (3.41), is applied, but only to the intermediate states. Recall that the target state is extracted from the reference trajectory; therefore, the elevation-angle constraint is already satisfied at the end points. Because the algorithm is targeting the state at t_{i+h} and the state is fixed at t_i, an elevation-angle constraint at these end points would be either redundant or possibly inconsistent.

An alternative (or supplement) to the elevation-angle constraint is a distance restriction to limit the excursion from the reference path at points along the trajectory. To prevent singularities in the partial derivatives of the distance restriction with respect to position variables, the square of the distance is employed in formulating the constraint, that is,

$$d^2 \leq d_{\max}^2 \tag{7.18}$$

[5] A formulation that employs two angles to represent the attitude (obviating the unit length constraint) is not successful.

Similar to the other path constraints, this inequality constraint is imposed via a slack variable, η_d, resulting in an equality constraint,

$$d^2 - d_{\max}^2 + \eta_d^2 = 0 \tag{7.19}$$

Like the elevation-angle constraint, this distance constraint is only applied to the intermediate states.

The elements in Eq. (7.17) are not equal to zero initially. Using the differential corrections described by Eqs. (7.8) through (7.10), the states in Eq. (7.16) are updated. Propagations that employ information from these new states and updated attitudes are used to repeat the process. A series of iterations proceeds until a convergence tolerance is satisfied (or some stopping condition is triggered), and a continuous trajectory emerges from t_i to t_{i+h}, where the vehicle is reoriented at each epoch. The resulting path is illustrated in Fig. 7.20.

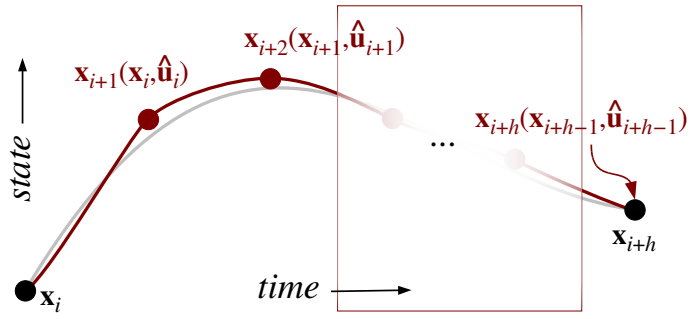


Fig. 7.20: A converged solution connecting $\mathbf{x}_i$ to $\mathbf{x}_{i+h}$.

The Jacobian for this multiple-shooting process relies on information from both the state transition matrix, $\mathbf{\Phi}$, and the state-control transition matrix, $\mathbf{\Psi}$ (introduced in Section 2.3). Dropping i for convenience, the Jacobian is evaluated from the form

$$\frac{\partial \mathbf{F}(\mathbf{X})}{\partial \mathbf{X}} = \begin{bmatrix}
2\mathbf{u}_0^{\mathrm{T}} & & & & & & & & & \\
\hat{\boldsymbol{\ell}}_0^{\mathrm{T}} & 2\eta_{\alpha,0} & & & & & & & & \\
\boldsymbol{\Psi}_{1,0} & 0 & -\mathbf{I} & & & & & & & \\
 & & \boldsymbol{\Gamma}_1 & 2\eta_{E,1} & & & & & & \\
 & & \mathbf{0} & 0 & 2\mathbf{u}_1^{\mathrm{T}} & & & & & \\
 & & \mathbf{0} & 0 & \hat{\boldsymbol{\ell}}_1^{\mathrm{T}} & 2\eta_{\alpha,1} & & & & \\
 & & \boldsymbol{\Phi}_{2,1} & 0 & \boldsymbol{\Psi}_{2,1} & 0 & -\mathbf{I} & & & \\
 & & & & & \boldsymbol{\Gamma}_2 & 2\eta_{E,2} & & & \\
 & & & & & & & \ddots & & \\
 & & & & & & & -\mathbf{I} & & \\
 & & & & & & & \boldsymbol{\Gamma}_{h-1} & 2\eta_{E,h-1} & \\
 & & & & & & & \mathbf{0} & 0 & 2\mathbf{u}_{h-1}^{\mathrm{T}} \\
 & & & & & & & \mathbf{0} & 0 & \hat{\boldsymbol{\ell}}_{h-1}^{\mathrm{T}} \\
 & & & & & & & \boldsymbol{\Phi}_{h,h-1} & 0 & \boldsymbol{\Psi}_{h,h-1}
\end{bmatrix} \begin{matrix} \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ \\ 2\eta_{\alpha,h-1} \\ 0 \end{matrix} \tag{7.20}$$

where $\mathbf{u}_0$ is the inertially defined three-element sail-face normal vector from Eq. (7.3) at time t_{i+0} (recall that this vector is guaranteed to be unit magnitude only at the initial guess and after convergence). The control parameters in Eqs. (7.16) through (7.20) are defined with respect to the inertial frame (i.e., Eq. (7.3)). The matrix $\mathbf{0}$ is a one-by-six array of zeros, $\mathbf{I}$ is the six-by-six identity matrix, and $\mathbf{\Phi}_{1,0}$ is the state transition matrix from t_{i+0} to t_{i+1}. The state-control transition matrix from t_{i+0} to t_{i+1}, $\mathbf{\Psi}_{1,0}$, is developed from the partial derivatives of the sail acceleration in Eq. (7.5) with respect to the components of the control vector, that is,

$$\frac{\partial \mathbf{a_s}(t)}{\partial \mathbf{u}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ \frac{\partial a_{x,1}}{\partial u_{X,0}} & \frac{\partial a_{x,1}}{\partial u_{Y,0}} & \frac{\partial a_{x,1}}{\partial u_{Z,0}} \\ \frac{\partial a_{y,1}}{\partial u_{X,0}} & \frac{\partial a_{y,1}}{\partial u_{Y,0}} & \frac{\partial a_{y,1}}{\partial u_{Z,0}} \\ \frac{\partial a_{z,1}}{\partial u_{X,0}} & \frac{\partial a_{z,1}}{\partial u_{Y,0}} & \frac{\partial a_{z,1}}{\partial u_{Z,0}} \end{bmatrix} \tag{7.21}$$

These partial derivatives are employed in the propagation of the variations of the state and control elements as $\mathbf{B}(t)$, per Eq. (2.48), to generate the state-control transition matrix, $\mathbf{\Psi}(t_1, t_0)$ from t_0 to t_1. Finally, the partial derivatives of the elevation-angle constraint with respect to the position coordinates are collected in the matrix $\mathbf{\Gamma}$. If the distance restriction is employed in lieu of the elevation-angle constraint, then the partial derivatives of the distance restriction with respect to the position components replaces $\mathbf{\Gamma}$ and η_d substitutes for η_E.

The number of turns along an arc determines the dimensions of the Jacobian matrix. The size of the Jacobian matrix in Eq. (7.20) is $(9(h-1)+8) \times (4+11(h-1))$. When $h > 3$, the Jacobian contains more columns than rows. In fact, increasing the number of subarcs, h, increases the control authority because the number of turns is increased. The difference between the number of columns and rows of the Jacobian is illustrated in Fig. 7.21. When the number of elements in $\mathbf{X}$ exceeds the number in $\mathbf{F}(\mathbf{X})$, $D\mathbf{F}(\mathbf{X})$ possesses more columns than rows and the pseudo-inverse of the Jacobian in Eq. (7.10) is a minimum-norm solution, identical to the formulation in

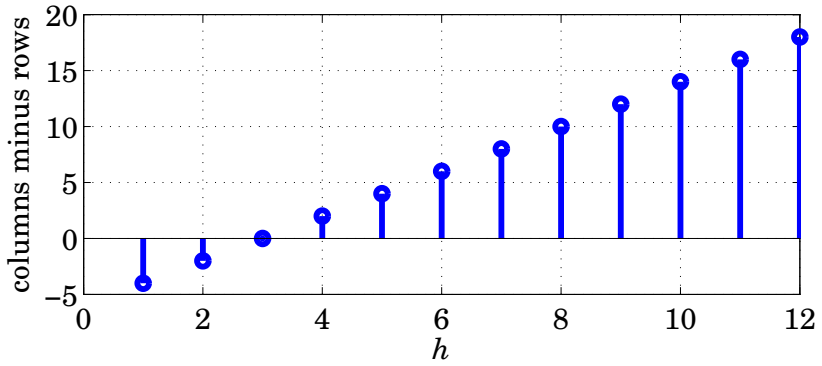


Fig. 7.21: The difference between the number of columns and rows of $D\mathbf{F}(\mathbf{X})$ as a function of h.

Eq. (3.22). Other forms of the pseudo-inverse exist; forms that employ QR decomposition or LU factorization often return a solution where the maximum number of elements in the solution vector possess zero value. An infinite number of solutions are available for Eq. (7.10), and the ability of the multiple-shooting algorithm to converge on a solution improves as more turns are incorporated into the formulation and the number of columns in the Jacobian matrix increases. A minimum-norm formulation supplies the smallest adjustment to the elements comprising the original guess for $\mathbf{X}$, thereby reducing the difference between the initial reference path and the resulting controlled trajectory.

7.5.1 Full-arc targeting, no errors

The first test of this approach is an application to an error-free scenario. Reference trajectories and initial guesses for the held attitudes are developed as described in Section 7.2. In this test, turns are assumed to be instantaneous and the spacecraft maintains its attitude throughout the hold period. Orbit determination and attitude control are assumed to be perfect as well. Path constraints limiting the maximum pitch angle to $\alpha_{\max} = 90^\circ$ and minimum elevation angle to $E_{\min} = 13^\circ$ are also incorporated into the multiple-shooting scheme.

The initial guess for the attitude is calculated using the two-point averages in Eqs. (7.7), and the latitude and longitude angles are converted into unit vectors. This two-point average is successful in initializing the multiple-shooting targeting scheme while averages from the trapezoidal integration in Eq. (7.6) are not successful, despite the proximity of the resulting initial guesses.

After the turn-and-hold sequence is determined, the initial spacecraft state is propagated from the beginning to the end of the arc, executing each turn from the sequence over the entire arc. This process is repeated for all subsequent arcs over the entire trajectory, typically 36 revolutions (recall that one revolution coincides with one synodic month). Because no errors are present in the design of the sequence and the subsequent simulated flight, the spacecraft matches the reference trajectory in 6 states at the end of the full arc. The number of days between turns (i.e., the subarc length) multiplied by the number of turns results in the arc length for the multiple-shooting scheme. These combinations, as well as the resulting arc lengths, are listed in Table 7.5. Recall that the time for each revolution is one synodic month,

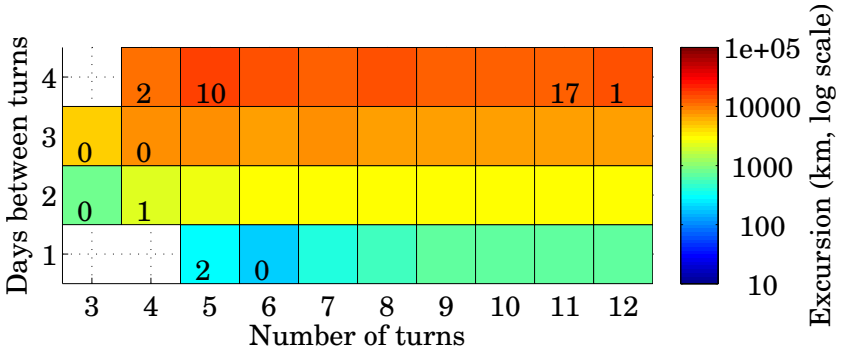
Table 7.5: Arc lengths (in days) for time between turns and h turns

Subarc	Number of turns, h									
length	3	4	5	6	7	8	9	10	11	12
4	12	16	20	24	28	32	36	40	44	48
3	9	12	15	18	21	24	27	30	33	36
2	6	8	10	12	14	16	18	20	22	24
1	3	4	5	6	7	8	9	10	11	12

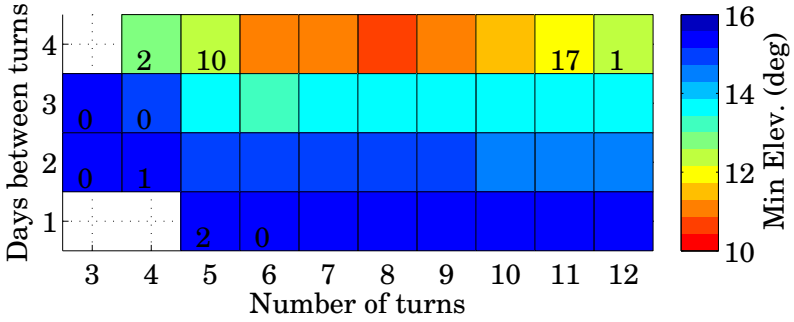
approximately 29.5 days. In general, longer arc lengths and longer times between turns allow for greater excursions from the reference path. In the presence of errors, these excursions can be detrimental to controlling the flight path. However, if the arc is short, the controlled path has less freedom to adjust and convergence may

be less likely. Conversely, more turns within an arc may increase the likelihood of convergence. From a practical, operations perspective, a lower frequency of turn-and-hold patterns, or longer time interval between turns, is desirable.

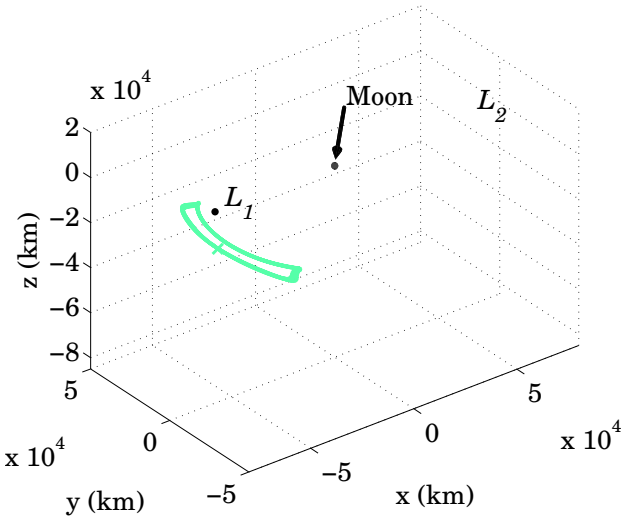
The five sample trajectories from Section 5.2 are employed to demonstrate this open-loop tracking scheme. The maximum distance from the reference trajectory and the minimum elevation angles along the path for this error-free multiple-shooting algorithm, as well as a sample trajectory corresponding to a specific combination of subarc length and number of turns per arc, appear in Figs. 7.22 through 7.26. The axes for the colored grids in each figure correspond to the subarc-length and turn combinations listed in Table 7.5. The colors in the boxes signify either the distance from the reference path (the upper grid, e.g., Fig. 7.22(a)) or the minimum elevation angle encountered along the controlled path (the lower grid, e.g., Fig. 7.22(b)). Note that blue boxes indicate either small excursions or large minimum elevation angles for that subarc-length and turn combination. For example, the path appearing in the Fig. 7.22(c) corresponds to a combination consisting of a 2-day subarc length and 7 turns per arc. This 2-by-7 combination yields a maximum excursion from the reference path of 2982 km (yellow), as demonstrated in Fig. 7.22(a), and a minimum elevation angle along the controlled path of 14.8° (dark blue), as demonstrated in Fig. 7.22(b). Numbers in the colored boxes indicate that the multiple-shooting strategy fails to converge on a solution after that many revolutions of the spacecraft; these combinations are not viable options. A "0" indicates that the algorithm successfully tracks the reference for only part of a cycle, but does not complete one full revolution; no box signifies that the algorithm is not able to converge on a solution for even the first arc based on that subarc-length and turns-per-arc combination (no subsequent arcs are attempted). Colored boxes without numbers signify that the algorithm converges for successive arcs over the full 36-month trial for the associated combination of subarc length and number of turns per arc. The algorithm converges successfully for successive arcs of the 2-by-7 combination associated with the sample aqua orbit for over 36 months, as apparent in Fig. 7.22(c). The controlled paths for



(a) Maximum excursion from the reference path.

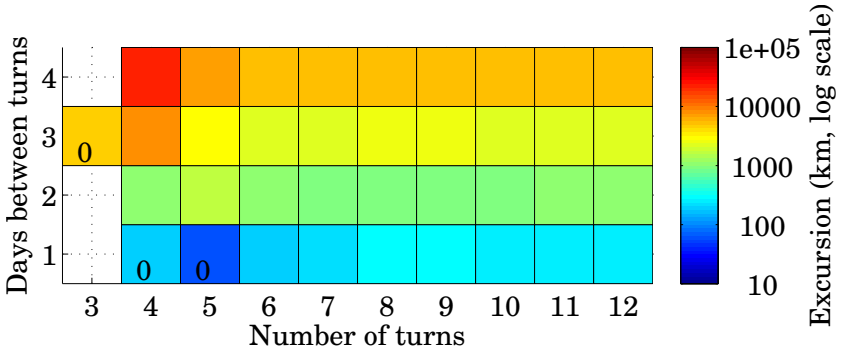


(b) Minimum elevation angle along the controlled path.

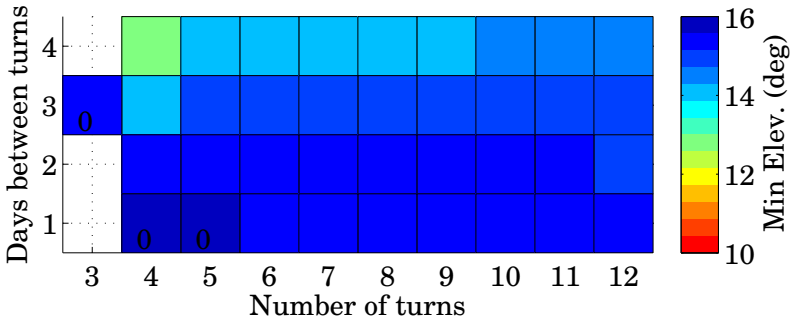


(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc)

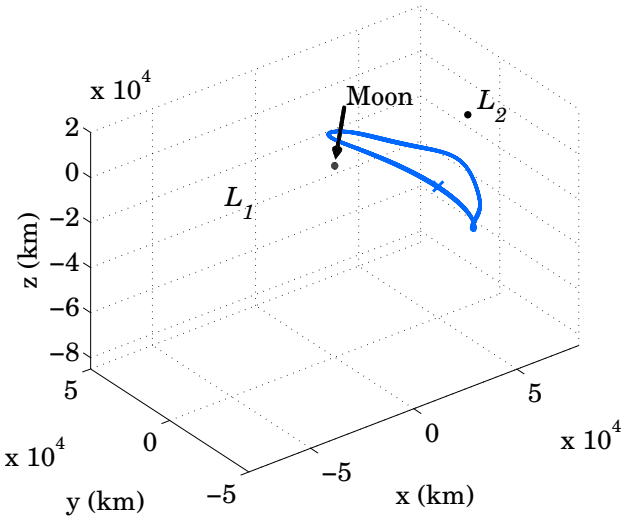
Fig. 7.22: Maximum excursions and minimum elevation angles for the aqua reference trajectory and a sample resulting path propagated for 36 months.



(a) Maximum excursion from the reference path.

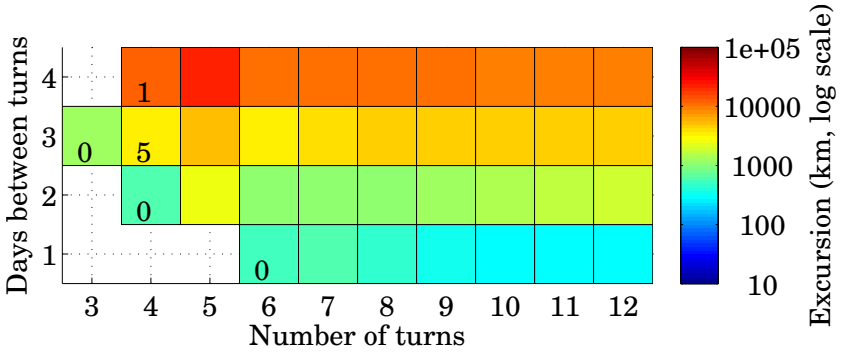


(b) Minimum elevation angle along the controlled path.

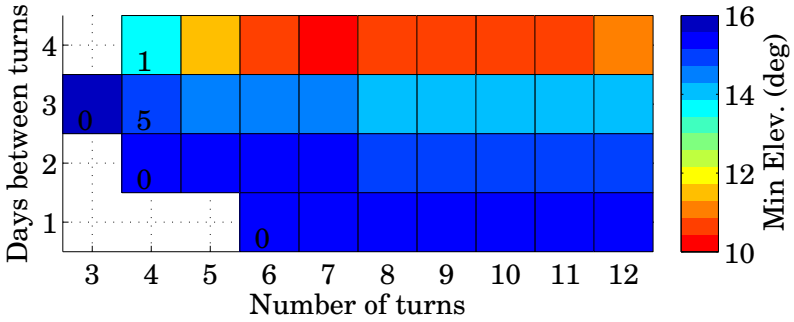


(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc)

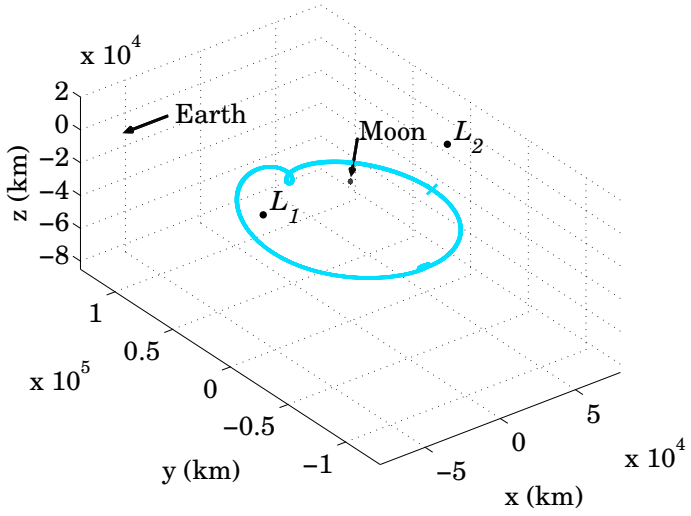
Fig. 7.23: Maximum excursions and minimum elevation angles for the royal-blue reference trajectory and a sample resulting path propagated for 36 months.



(a) Maximum excursion from the reference path.

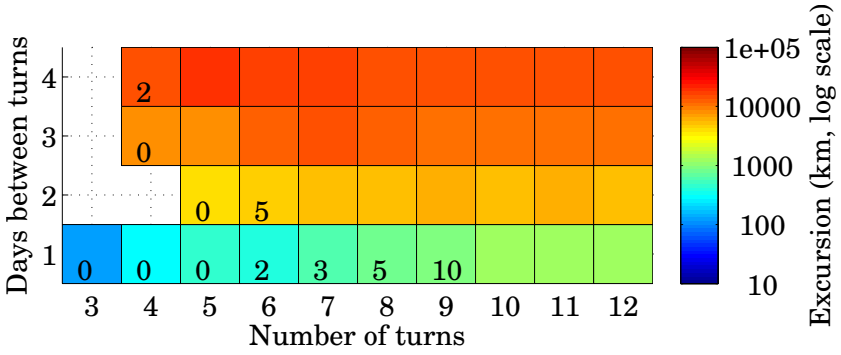


(b) Minimum elevation angle along the controlled path.

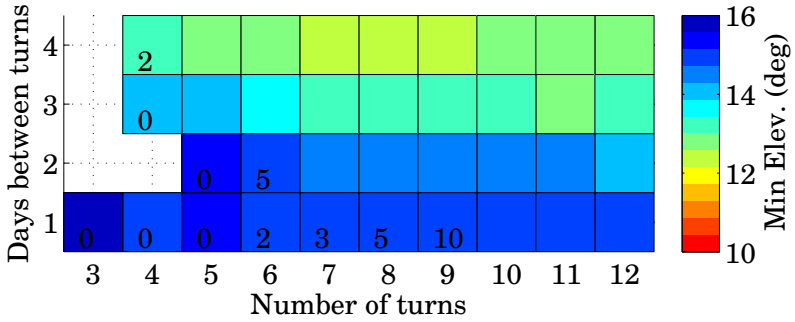


(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc)

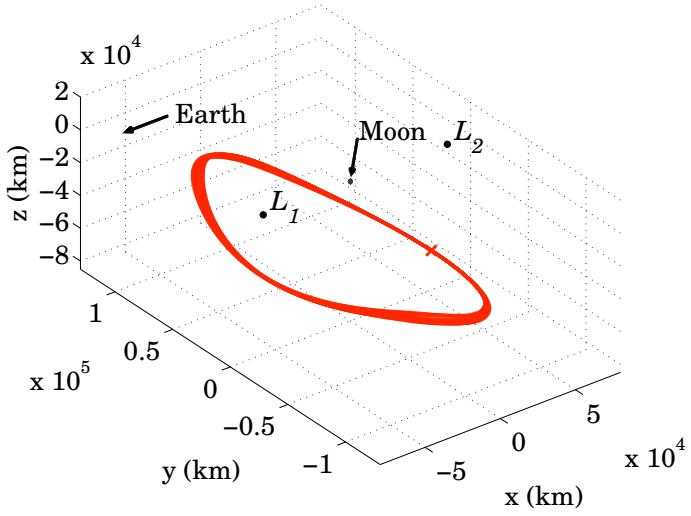
Fig. 7.24: Maximum excursions and minimum elevation angles for the cyan reference trajectory and a sample resulting path propagated for 36 months.



(a) Maximum excursion from the reference path.

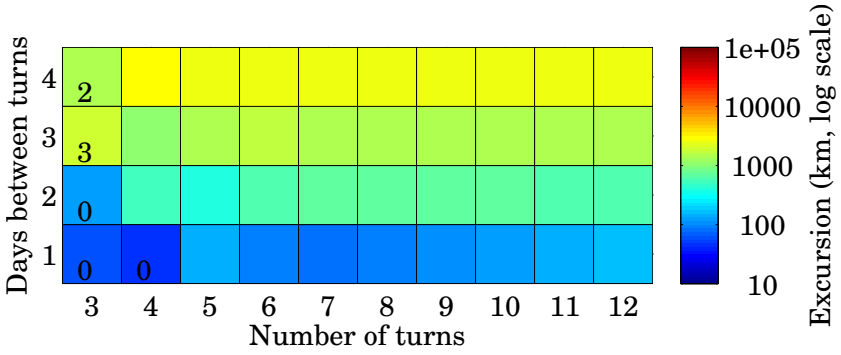


(b) Minimum elevation angle along the controlled path.

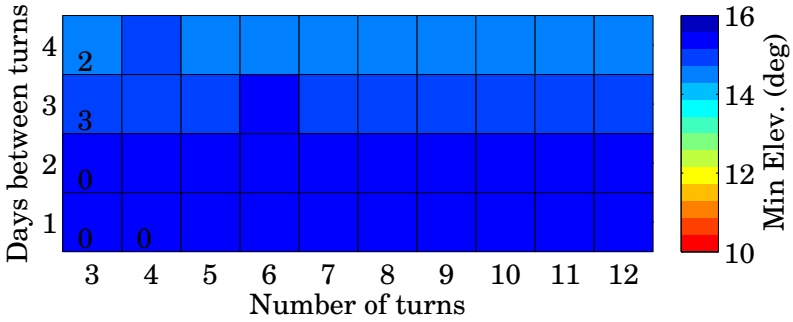


(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc)

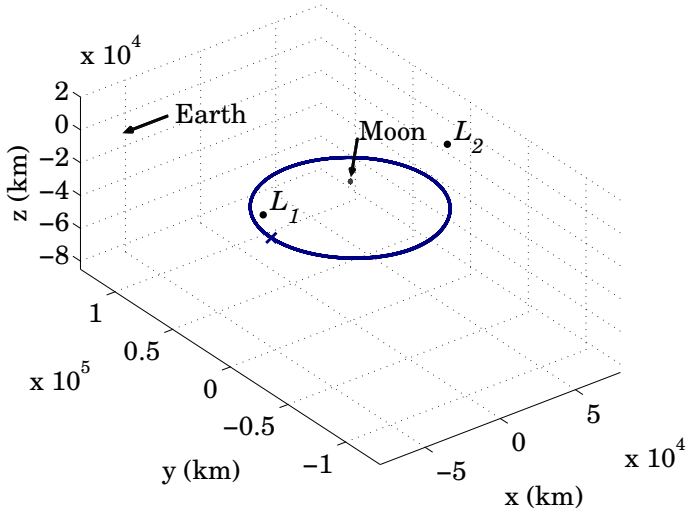
Fig. 7.25: Maximum excursions and minimum elevation angles for the red-orange reference trajectory and a sample resulting path propagated for 36 months.



(a) Maximum excursion from the reference path.



(b) Minimum elevation angle along the controlled path.



(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc)

Fig. 7.26: Maximum excursions and minimum elevation angles for the dark-blue reference trajectory and a sample resulting path propagated for 36 months.

the sample trajectories corresponding to the respective 2-by-7 combinations appear in Figs. 7.22 through 7.26. The orbits in the figures originate at the "x" and the spacecraft moves in the $-y$ direction initially. The path of the Sun originates from a position along the $-x$ axis and moves about the system clockwise in the xy plane. Note that although the elevation-angle constraint is set to $E_{\min} = 13^\circ$ across all figures, some combinations in the figures report elevation angles lower than this value. Recall that the constraint is only applied at the turn events. The path between the turns during the hold period is not constrained in any way. Recall that a well-posed two-point boundary value problem, where the number of controls (6: two angles and *three* turns) is the same as the number of target states (position and velocity at the target), supplies a unique solution to the targeting problem. However, this type of three-turn strategy is rarely successful and not suitable for implementation as a control law. The difficulties using BVP6C for a three-turn two-point BVP are discussed in Section 7.4. Clearly, an excess of control authority is necessary for targeting states along the reference path.

The maximum excursion from the reference path and the elevation-angle along the controlled path are two significant metrics for assessing the combinations employed in the error-free, open-loop, multiple-shooting control scheme. The extreme values of these two significant metrics across all of the combinations are listed in Table 7.6. Again, the elevation angle constraint is set to $E_{\min} = 13^\circ$ only at the turn events; the path along the subarc between turns is not constrained. The subarc length, g, and the number of turns, h, corresponding to these extreme values are listed in the table as well. The maximum excursions from the reference paths and minimum elevation angles along the controlled trajectories associated with combinations consisting of a 2-day subarc length and 7 turns per arcs are also listed in Table 7.6. The minimum elevation angles from the controlled orbits for the 2-by-7 combination are only slightly less than 15°, indicating that reference paths with elevation-angle constraints that include a few degrees of margin are sufficient for mission design. Even for the 2-by-7 combination associated with the red-orange orbit, the maximum excursion is

Table 7.6: Full-arc targeting: Elev. angles and excursions from reference paths

Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
Extreme values					
Max Excur. (km)	16670	20808	22055	17992	2885
g-by-h	4-by-5	4-by-4	4-by-5	4-by-6	4-by-5
Min Elev. Angle	10.8°	12.8°	10.4°	12.2°	14.6°
g-by-h	4-by-8	4-by-4	4-by-7	4-by-10	4-by-3
2 days between turns, 7 turns					
Max Excur. (km)	2982	947	1112	5066	652
Min Elev. Angle	14.8°	15.2°	15.2°	14.4°	15.3°

greater than 5000 km, but the minimum elevation angle is within one degree of the 15° elevation-angle constraint from the designed path. The extreme values of the maximum excursion and the minimum elevation angles occur when the number of days between turn events is greatest, not surprisingly. Additional path constraints between turns may reduce these extreme values. For the full-arc, error-free cases, all combinations corresponding to extreme values, except one (the 4-by-5 combination for the aqua orbit), converge successfully for the full 36 revolutions. That is, after each arc all conditions to match the reference path are met, and the algorithm proceeds to the next arc, and so on for 36 months; no additional control is necessary.

7.5.2 "Look-ahead" strategy with simulated errors

The next step in the evaluation of a multiple-shooting approach is the incorporation of errors in the orbit knowledge, control authority, and turn modeling. The development of trajectory state uncertainty for a solar sail spacecraft is the subject

of a full covariance analysis and depends on observation periods, data types, and other considered and estimated uncertainties. However, reasonable uncertainties in the position and velocity components at any point along the trajectory equal to 10 km (3σ) in position and 10 cm/sec (3σ) in velocity are assumed. For attitude control considerations, other researchers assume that a dead-band or pointing error of 1° (3σ) or less is possible [66, 141–143]. Finally, while sailcraft are generally large structures with flexible modes, turn rates of 0.02°/sec (72°/hr) and 0.057°/sec (205°/hr) are assumed by Wie [141] and Polites et al. [142], respectively. Even with these relatively fast slew rates, the effect of assuming instantaneous turns must be considered within the context of turns that occur over finite times. Therefore, a turn rate of 10°/hour is reasonable. Additionally, most (but not all) predicted turns are approximately 10° or less for the sample orbits.

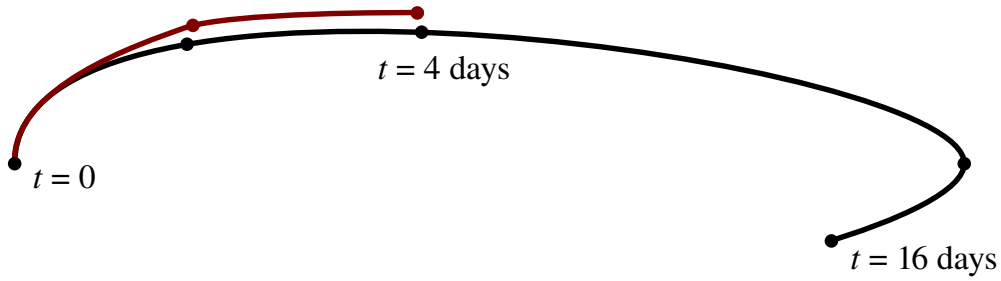
To simulate these errors, attitudes are determined for subarcs within the arc assuming instantaneous turns in a multiple-shooting scheme. The position and velocity of the spacecraft at the initial epoch are randomly perturbed by 10 km and 10 cm/sec (3σ), respectively, to incorporate errors in the spacecraft state and these perturbed states are employed when generating the turn sequence. After the turn sequence is produced, the attitude for each subarc is perturbed by a random error of 1° (3σ) to accommodate errors in the control. Larger errors, such as 3° (3σ), tend to overwhelm the control strategy. The unperturbed position and velocities are used to initiate the flight simulation that incorporates the error-induced control profile.[6] In this simulation, the turns are modeled as ramp functions over the time required to complete a reorientation using a slew rate of 10°/hour. Recall that the designed turn sequence assumes instantaneous turns. The error introduced by this discrepancy is the dominant error source if the actual turn rates are very slow (e.g., less than 5°/hour). Errors in the knowledge of the state do not affect the trajectory as much as the pointing

[6]Because imperfect state knowledge is used to create a turn profile, the profile is itself imperfect before errors in the control performance are incorporated. While state knowledge and control performance are considered separately in this investigation, their coupled effect is essentially a larger control error in the flight simulation than if either were considered alone.

errors or errors in the turn rates.[7] Because of these knowledge, control, and modeling errors, the actual trajectory does not return to the reference path and often diverges in as little as one or two arcs when executing all the turns over a full arc.

An adaptation of the multiple-shooting scheme, labeled the "look-ahead" strategy, is successful in tracking of the reference path in the presence of errors. In this approach, the spacecraft is propagated in the error configuration along one subarc. At the end of that subarc, a new set of attitudes is generated to target a point further along the reference path. After the next subarc, the process is repeated. An example appears in Fig. 7.27. Errors are introduced as described previously, a turn-and-hold sequence based on a full arc is generated, but only the first turn and subarc of the sequence are simulated in a flight trajectory. At the end of this first segment, $t = 2$

(a) First propagated 2-day subarc (red) and reference path (black) of a 14-day arc.



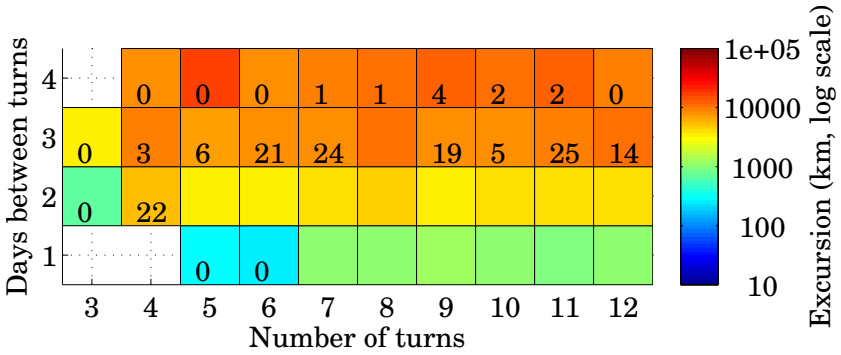
(b) First and second propagated 2-day subarcs (red) and reference path (black) up to day 16

Fig. 7.27: Progression of the look-ahead strategy.

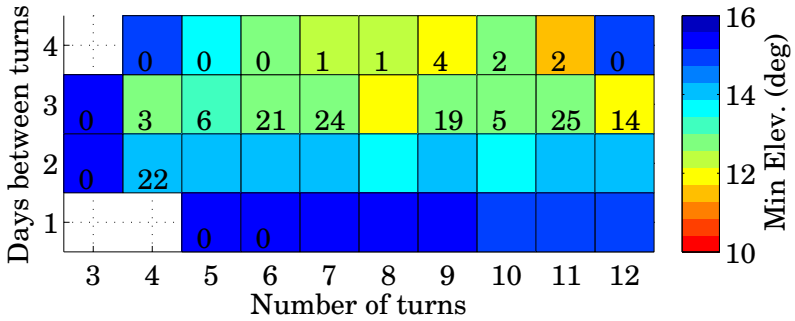
[7]A limited examination of the data comparing the turns generated with and without the state-knowledge errors of 10 km and 10 cm/s (3σ) for a particular arc in each of the five sample orbits resulted in differences in each turn comparable to an arcminute or less.

days in Fig. 7.27(a), the sailcraft is not on the reference trajectory. The sailcraft is not expected to return to the reference path because of the errors. A new set of turns is planned for 2-day increments over the next 14 days. The current state at day 2 (including state-knowledge error) and guesses for the attitudes, per Eq. (7.7), are employed in initializing a subsequent turn-and-hold sequence. The corrected second subarc and new 14-day arc appear in Fig. 7.27(b). Essentially, the corrector is allowed to "look" 14 days downstream when designing an attitude profile for the next two days. A slight variation on this look-ahead strategy is the use of two or more subarcs before reinitializing the attitude profile. This receding-horizon procedure is similar to both the target-point station-keeping strategy of Howell and Pernicka [193] and model-predictive control (without an optimized objective function) [194].

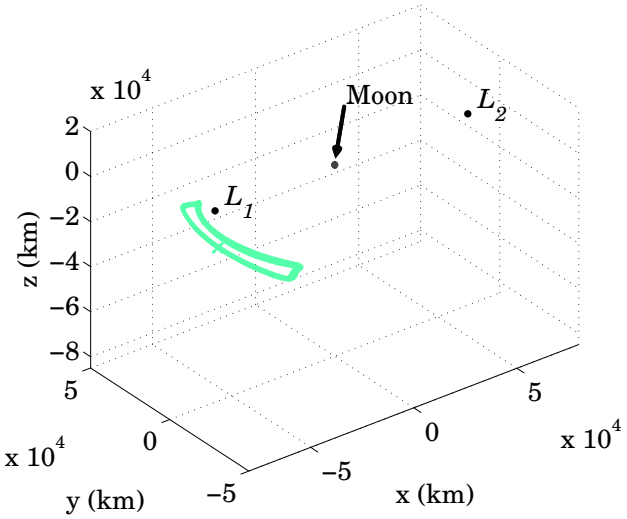
In the presence of simulated errors, the sailcraft tracks the reference path for most combinations of subarc length and number of turns listed in Table 7.5. The successful result is mostly likely due to the re-initialization of the turn-and-hold orientations after every subarc. The maximum distance from the reference trajectory and the minimum elevation angles along the path for this look-ahead approach, in the presence of the errors, as well as a sample trajectory corresponding to a specific combination of subarc length and number of turns per arc, appear in Figs. 7.28 through 7.32. The axes for the colored grids in each figure correspond to the subarc-length and turn combinations listed in Table 7.5. The colors in the boxes signify either the distance from the reference path (the upper grid, e.g., Fig. 7.28(a)) or the minimum elevation angle encountered along the controlled path (the lower grid, e.g., Fig. 7.28(b)). Blue boxes indicate either small excursions or large minimum elevation angles for that subarc-length and turn combination. For example, the path appearing in the Fig. 7.28(c) corresponds to a combination consisting of a 2-day subarc length and 7 turns per arc. This 2-by-7 combination yields a maximum excursion from the reference path of 4003 km (yellow), as demonstrated in Fig. 7.28(a), and a minimum elevation angle along the controlled path of 14.0° (blue), as demonstrated in Fig. 7.28(b). Numbers in the colored boxes indicate that the look-ahead strategy fails to converge on a solution



(a) Maximum excursion from the reference path.

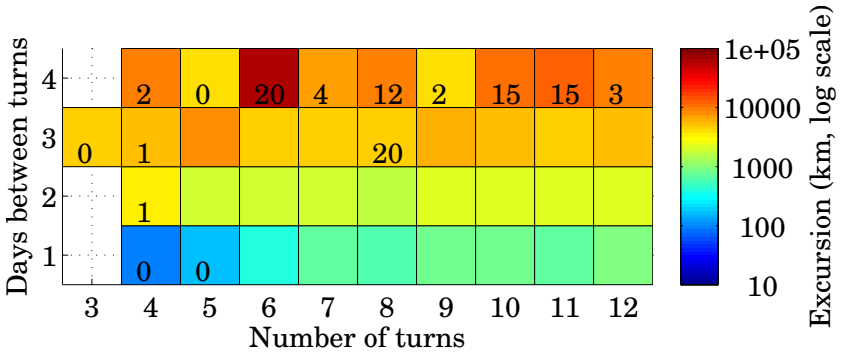


(b) Minimum elevation angle along the controlled path.

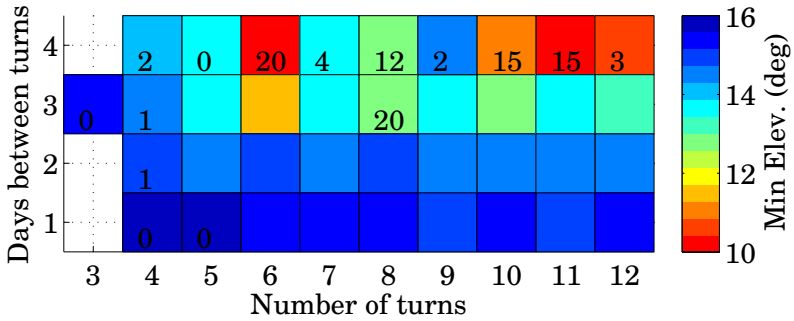


(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc). Only the first turn is performed, then the procedure is re-initialized.

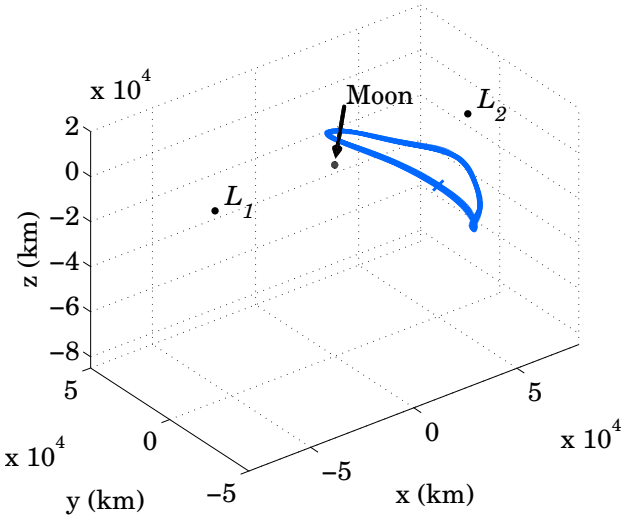
Fig. 7.28: Maximum excursions and minimum elevation angles for the aqua reference trajectory and a sample resulting path propagated for 36 months.



(a) Maximum excursion from the reference path.

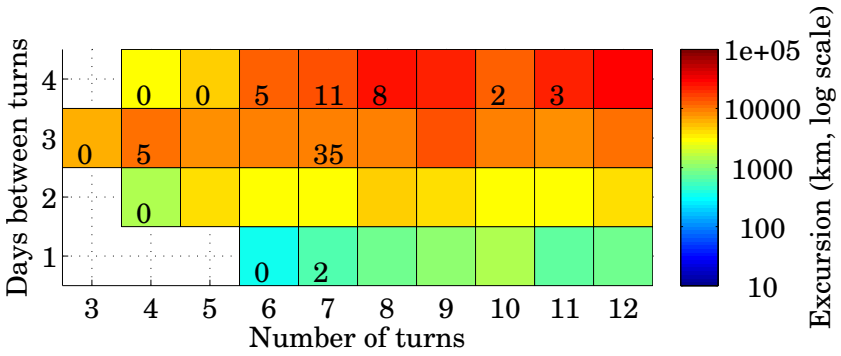


(b) Minimum elevation angle along the controlled path.

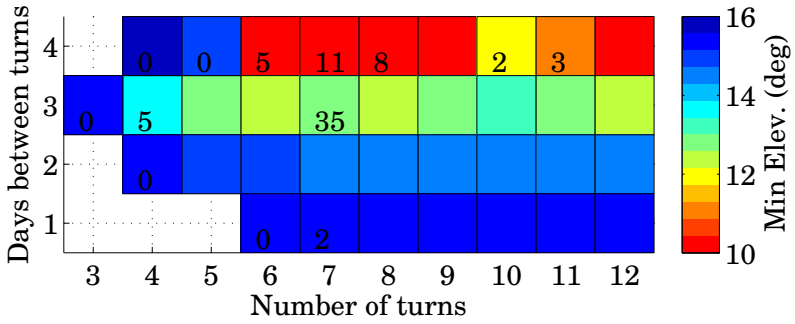


(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc). Only the first turn is performed, then the procedure is re-initialized.

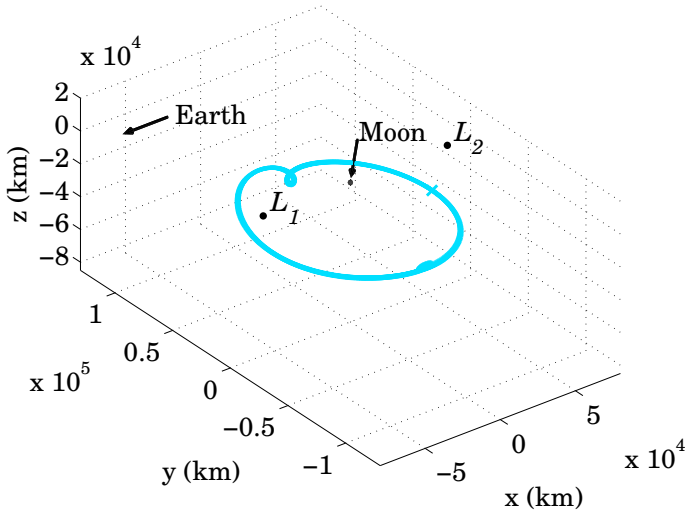
Fig. 7.29: Maximum excursions and minimum elevation angles for the royal-blue reference trajectory and a sample resulting path propagated for 36 months.



(a) Maximum excursion from the reference path.

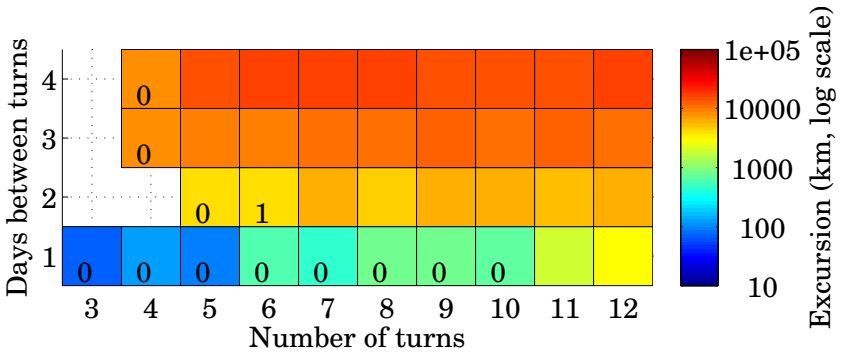


(b) Minimum elevation angle along the controlled path.

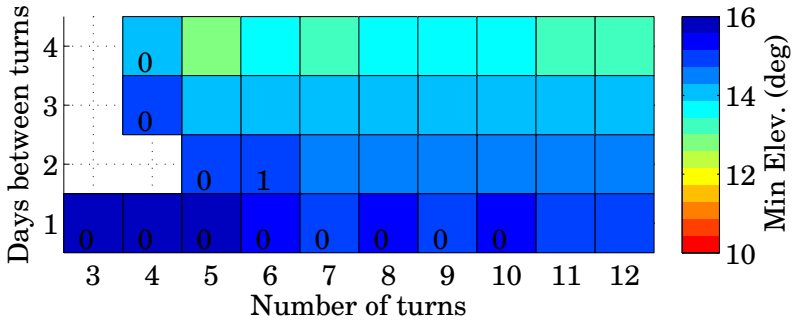


(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc). Only the first turn is performed, then the procedure is re-initialized.

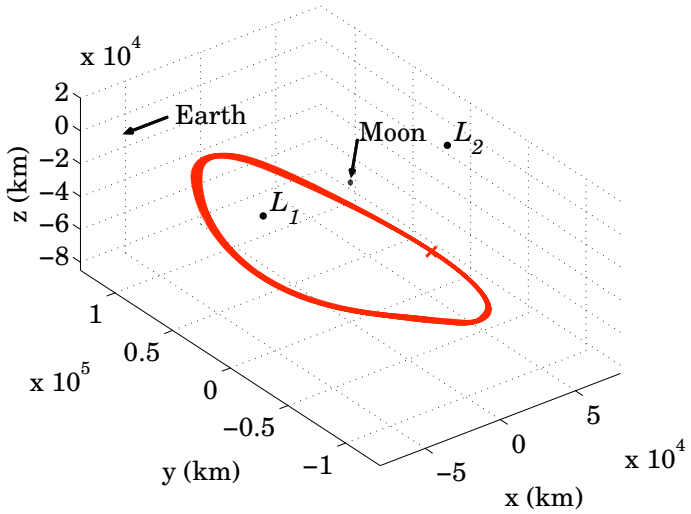
Fig. 7.30: Maximum excursions and minimum elevation angles for the cyan reference trajectory and a sample resulting path propagated for 36 months.



(a) Maximum excursion from the reference path.

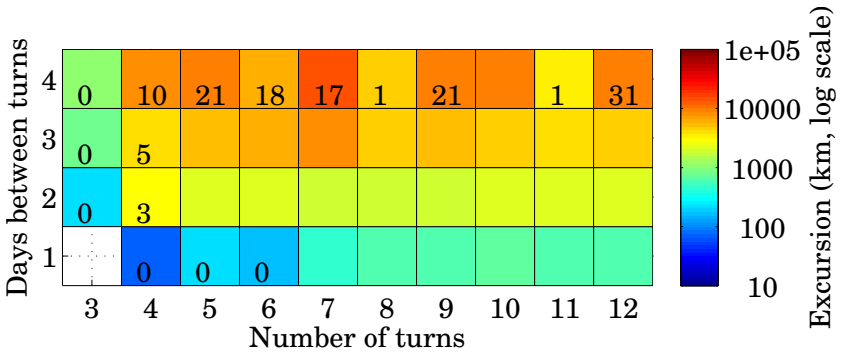


(b) Minimum elevation angle along the controlled path.

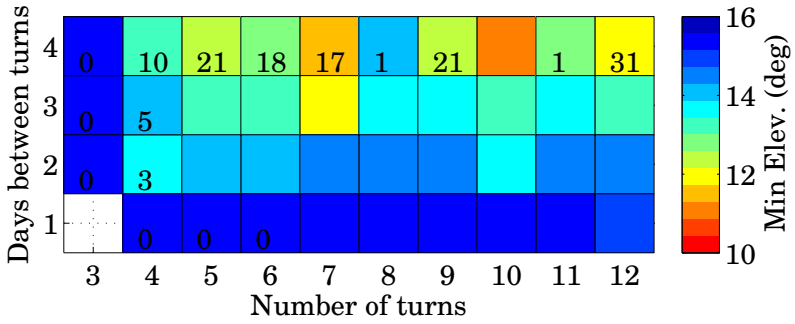


(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc). Only the first turn is performed, then the procedure is re-initialized.

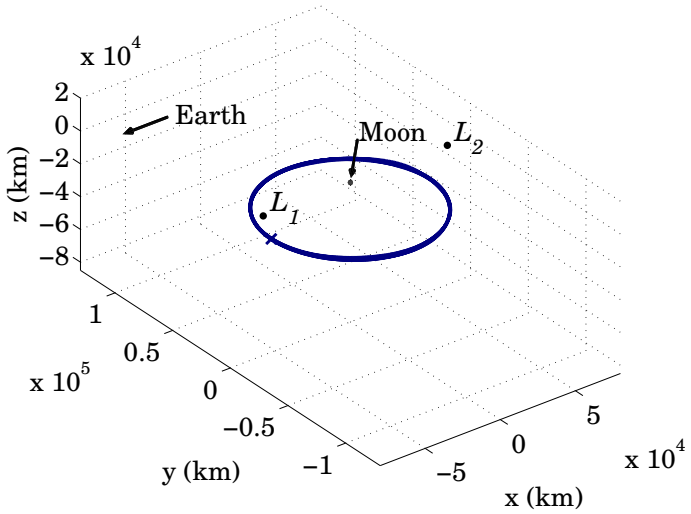
Fig. 7.31: Maximum excursions and minimum elevation angles for the red-orange reference trajectory and a sample resulting path propagated for 36 months.



(a) Maximum excursion from the reference path.



(b) Minimum elevation angle along the controlled path.



(c) Path corresponding to a 2-day subarc length and 7 turns per arc (i.e., 14 days per arc). Only the first turn is performed, then the procedure is re-initialized.

Fig. 7.32: Maximum excursions and minimum elevation angles for the dark-blue reference trajectory and a sample resulting path propagated for 36 months.

after that many revolutions of the spacecraft. A "0" indicates that the algorithm successfully tracks the reference for only part of a cycle, but does not complete one full revolution; no box signifies that the algorithm is not able to converge on a solution for even the first arc based on that subarc-length and turns-per-arc combination (no subsequent arcs are attempted). Note, however, that these numbers are based on a trajectory initiated when $t = 0$ and the Sun, Earth, and Moon are aligned on the x axis. If the simulation were initiated at some other point along the orbit, the number of revolutions (i.e., synodic months) before the algorithm fails may be different, and other combinations may fail as well. Colored boxes without numbers signify that the algorithm converges for successive arcs over the full 36-month trial for the associated combination of subarc length and number of turns per arc. The algorithm successfully converges for successive arcs of the 2-by-7 combination associated with the sample aqua orbit for over 36 months, as apparent in Fig. 7.28(c). The controlled paths for the sample trajectories corresponding to the respective 2-by-7 combinations appear in Figs. 7.22 through 7.26. The orbits in the figures originate at the "x" and the spacecraft initially moves in the $-y$ direction. The path of the Sun originates from a position along the $-x$ axis and moves about the system clockwise in the xy plane. Note that although the elevation-angle constraint is set to $E_{\min} = 13^\circ$ across all figures, some combinations in the figures report elevations lower than this value. Recall that the constraint is only applied at the turn events. The path between the turns during the hold period is not constrained in any way.

Maximum excursion from the reference path and elevation-angle along the controlled path are two significant metrics for assessing the combinations employed in the error-free multiple-shooting control scheme. The extreme values of these two significant metrics along each trajectory across all of the combinations are listed in Table 7.7, along with the number of completed revolutions before the look-ahead scheme is no longer able to obtain a solution. The elevation angle constraint is set to $E_{\min} = 13^\circ$ only at the turn events; the path along the subarc between turns is not constrained and the extreme values in the table are calculated from trajectory

Table 7.7: Look-ahead targeting: Excursions from the reference paths and elevation angles along controlled paths, based on a configuration that incorporates 1° errors in the control

Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
Fig.	7.28	7.29	7.30	7.31	7.32
Extreme values					
Max Excur. (km)	15725	58140	30485	17201	15106
g-by-h	4-by-5	4-by-6	4-by-12	4-by-8	4-by-7
No. of revs.	0	20	36	36	17
Min Elev. Angle	11.5°	−87.4°	8.1°	12.9°	11.1°
g-by-h	4-by-11	4-by-6	4-by-9	4-by-5	4-by-10
No. of revs.	2	15	36	36	36
2-day subarcs, 7 turns per arc in designed sequence					
Max Excur. (km)	4003	1882	2972	5887	1983
Min Elev. Angle	14.0°	14.5°	14.6°	14.5°	14.5°

data reported by the numerical propagator. The subarc length, g, and the number of turns, h, corresponding to these extreme values are listed in the table as well. The maximum excursions from the reference paths and minimum elevation angles along the controlled trajectories associated with combinations consisting of a 2-day subarc length and 7 turns per arcs are also listed in Table 7.7. Note the extreme minimum elevation angle associated with the royal-blue reference path. In the simulation (not depicted), the controlled path leaves the vicinity of L_2 and loops around the Moon before rejoining the reference path during the fourteenth revolution. The multiple-shooting algorithm is unable to converge on a solution for this royal-blue trajectory after revolution 15.

Understandably, the excursions are greater (and the minimum elevation angles are lower) for the trajectories that incorporate multiple error sources. Additionally, even with look-ahead targeting, not all combinations of arc length and number of turns are successful in tracking the reference path. The look-ahead algorithm fails to converge reliably for arcs with an insufficient number of turns (the abscissa of the tiled plots in Figs. 7.28 through 7.32). It also sometimes fails to converge reliably for sequences that incorporated long subarcs. A larger control error of 3° (3σ) significantly reduces the ability of the look-ahead method to control the flight-path. Results from those cases appear in Appendix D. Clearly, attitude control and modeling drive the performance of solar sail spacecraft trajectory design. Although certain combinations of subarc length and number of turns are successful in both the error-free, full-arc approach and the look-ahead implementations that incorporate errors, a more robust control law is generally required.

7.5.3 Adaptive control with simulated errors

An alternative strategy to track a reference trajectory is examined. To generate a sequence, assume that the subarc length and the number of turns employed are adaptable. Similar to the look-ahead scheme, only the first turn-and-hold subarc from a particular solution is employed. Thereafter, the turn sequence is updated based on new state information.

For this preliminary investigation, it is convenient to schedule turns after an interval numbers of days, preferably separated by multiple days. It also is desirable to maintain the actual path close to the reference orbit over the first subarc. These two objectives are sometimes in conflict, as demonstrated previously. For this flight-path control strategy, a given subarc length and a given number of turns are initially employed to generate a turn sequence. If the multiple-shooting algorithm fails to converge, the number of turns is increased such that an integer number of days between turns is preserved. The number of turns, and subsequently, the arc length, is increased

to some maximum value. If no attempts for a given subarc length converge, the subarc length (i.e., the time between turns) is decreased. A sample strategy appears in Fig 7.33. Initially, a 12-day arc consisting of three 4-day subarcs is attempted. If

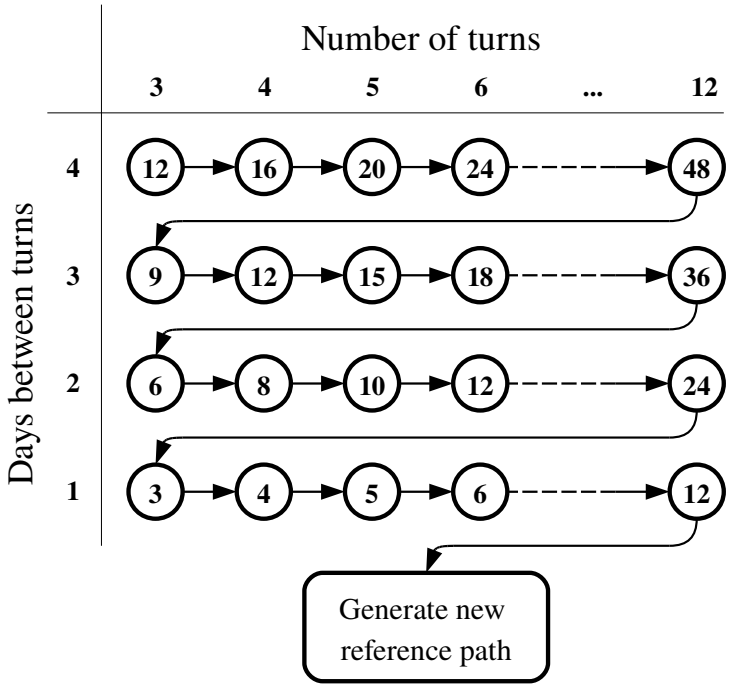


Fig. 7.33: Map of adaptive strategy. The circled number indicates the arc length.

that combination does not result in convergence of the algorithm to an acceptable trajectory, another turn is added. To add a turn and maintain the same time interval between turns, the length of the arc is increased until a limit for the number of turns is reached (12 turns in this example). Then, the subarc length decreases to three days. Consistent with the previous implementation of the look-ahead method, a new turn profile is generated after each subarc.

Two different strategies are examined for the adaptive tracking scheme. The first employs an elevation-angle constraint of 13°. Because the simulated flight of the spacecraft could diverge from the reference path, but remain within the elevation-angle constraint, a second strategy employs an excursion constraint at the turn events

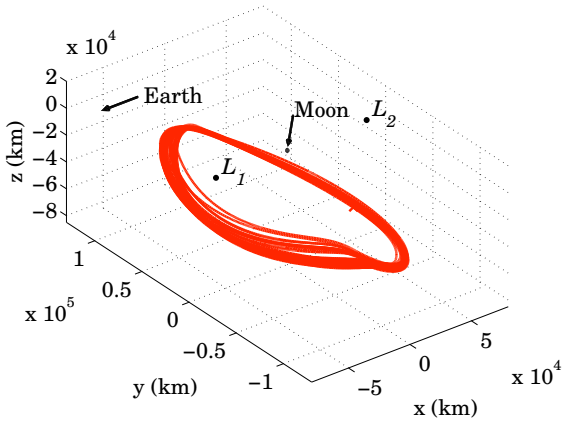
when designing the turn sequence. In both strategies, the random errors in the knowledge of the state are 10 km in position and 10 cm/sec in velocity (3σ) at the beginning of the arc, with a 3σ random turn execution error of 1°. The simulations for both strategies also incorporate a 10°/hr turn that is not included when designing the turn sequence, which assumes instantaneous turns. The results for these two strategies for each of the sample orbits are summarized in Table 7.8. Note that for these cases, the average subarc length, g, is between 3 and 4 days; additionally, the average number of turns, h, used to generate a turn sequence is generally below 6 turns. Because the initial guess for the time between turns that is employed to examine this strategy is large ($g = 4$ days) and the path constraints are only enforced at the time of the turn, the controlled trajectory diverges from the reference path for

Table 7.8: Adaptive targeting: 4-day window, 1° control error (3σ) with state knowledge of 10 km in position and 10 cm/sec in velocity (3σ) and 10°/hr slew rates

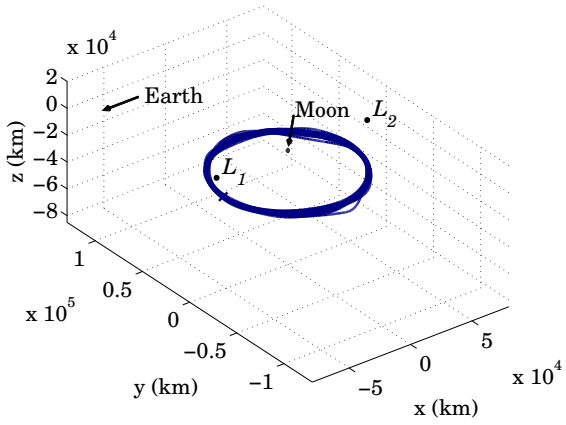
Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
13° elevation-angle constraint					
Max Excur. (km)	10179	8666	17216	23471	18609
Min Elev. Angle	13.2°	13.5°	11.5°	12.9°	12.1°
No. of revs.	0	3	3	36	36
Ave. g (days)	3.67	4	4	4	4
Ave. h (turns)	4.33	3.39	3.63	3.65	3.21
5000 km excursion constraint					
Max Excur. (km)	5316	5957	4259	7482	7419
Min Elev. Angle	12.4°	13.4°	14.5°	14.2°	12.1°
No. of revs.	1	2	0	3	36
Ave. g (days)	3.15	3.82	3.83	3.23	3.99
Ave. h (turns)	4.08	4.14	5.5	3.97	3.15

most of the scenarios listed in Table 7.8. Only the red-orange and dark-blue orbits maintain their paths near the respective references for the full 36 months with the elevation-angle constraint, and only the dark-blue orbit tracks its reference path for the complete trajectory with the excursion constraint. The resulting paths appear in Fig. 7.34.

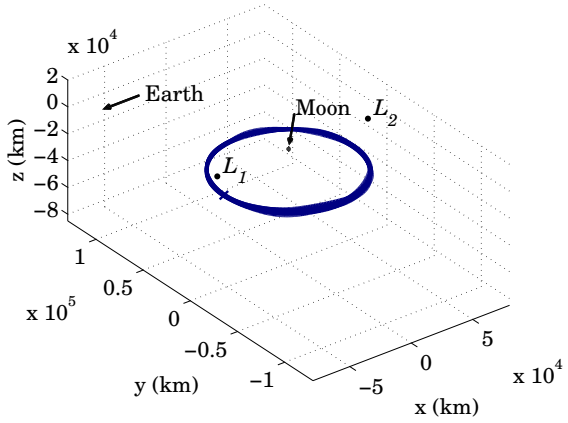
Using the same error configurations, the adaptive scheme is initiated with 2 days between turns. Under this approach, all reference paths are tracked, regardless of whether the elevation-angle constraint or the excursion constraint is employed. The



(a) Elev. const., Orbit #4



(b) Elev. const., Orbit #5



(c) Excur. const., Orbit #5

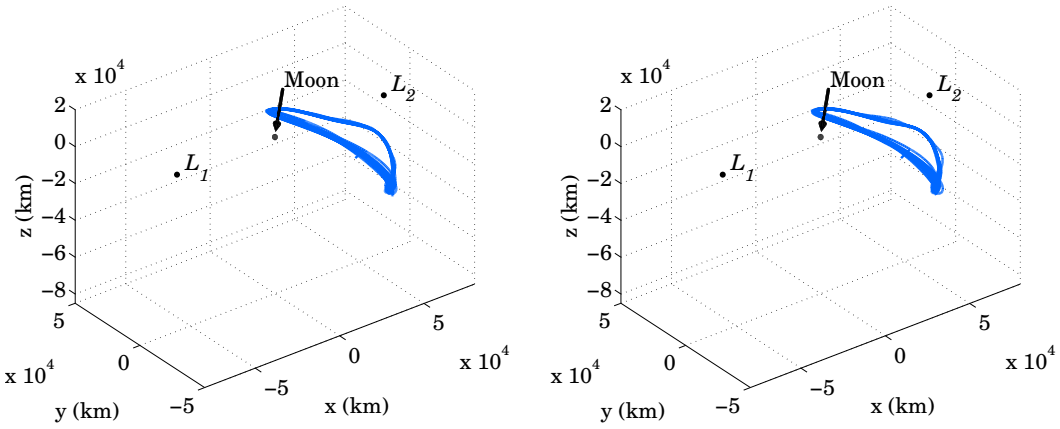
Fig. 7.34: Two successful trajectories emerge using the adaptive tracking scheme, initiated with 4 days between turns.

results from employing a 2-day window are listed in Table 7.9. With minor exceptions along the royal-blue orbit under L_2 (plotted in Fig. 7.35), all of the orbits from Table 7.9 are visibly similar to their respective reference trajectories.

Table 7.9: Adaptive targeting: 2-day window, 1° control error (3σ) with state knowledge of 10 km in position and 10 cm/sec in velocity (3σ) and 10°/hr slew rates

Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
13° elevation-angle constraint					
Max Excur. (km)	10140	5831	3324	9073	2373
Min Elev. Angle	14.3°	14.8°	15.2°	14.5°	14.5°
No. of revs.	36	36	36	36	36
Ave. b (days)	2	2	2	2	2
Ave. h (turns)	3.42	3.45	3.47	4.29	3.24
5000 km excursion constraint					
Max Excur. (km)	4798	4824	3270	5041	2992
Min Elev. Angle	13.2°	14.5°	15.1°	14.5°	13.5°
No. of revs.	36	36	36	36	36
Ave. b (days)	2	2	2	2	2
Ave. h (turns)	3.42	3.48	3.43	4.23	3.2

As discussed previously, control errors affect the ability of a flight-path control algorithm to track a reference trajectory successfully. Results from (1) adaptive tracking schemes that incorporate no orbit determination errors, no control errors, and no turn modeling errors along with (2) similar schemes that incorporate 3σ random errors of 10 km in position and 10 cm/s in velocity, as well as 10°/hr slew rates (instantaneous turns are assumed in the design of the sequence) and 3° control errors (as opposed to the 1° control errors examined previously), appear in Appendix D. Even in the error-free cases, the adaptive scheme based on the 4-day window exhibits difficulty in



(a) Elev. const., Orbit #2

(b) Excur. const., Orbit #2

Fig. 7.35: Successful trajectories from the adaptive tracking scheme, initiated with 4 days between turns, for the royal-blue orbit (#2) corresponding to a 13° elevation-angle constraint and a 5000 km excursion constraint.

tracking some reference trajectories, while all reference trajectories are tracked successfully for the full 36 months when the adaptive scheme is based on a 2-day window. Interestingly, some orbits can be tracked in the presence of large control errors (3°, 3σ), even in the presence of the standard state knowledge and turn modeling errors. These results indicate that path constraints between the reorientation events (i.e., the times of the turns) may be required if longer times between turns are desired, especially if pointing errors are large.

Any turn-and-hold scheme applied to these reference trajectories is essentially a hybrid control system where a discrete control law is applied to a continuous system. In practice, an upper limit on the time between turns will exist, independent of any number of path constraints. In general, the adaptive scheme is more successful when an elevation-angle constraint is incorporated into the turn sequence design as an alternative to an excursion constraint. More trials involving various sizes of the excursion constraint may improve the performance of the algorithm. Alternatively, an approach that adds path constraints on the velocity vector (such as a cone constraint

or a magnitude constraint) may improve the ability of the adaptive scheme to track a reference path. The adaptive tracking scheme, in the presence of large control errors, can successfully track a sailcraft trajectory and warrants further investigation.

8. CONCLUSIONS

The goal of this investigation is to determine what level of technology is required to displace a single solar sail spacecraft sufficiently such that it solves the lunar south pole (LSP) coverage problem. To address this problem, a broad range of potential trajectory options are considered. Because of the coupling of a solar sail trajectory to the vehicle attitude, flight-path control issues arise for whatever trajectory option is selected. Control laws for solar sail spacecraft are developed in response to this challenge.

Fortunately, a variety of numerical techniques for solving boundary value problems are available for uncovering trajectory options and to aid in the understanding of motion within the Earth–Moon system. One such technique that is robust, quick, and conceptually simple is the finite-difference method (FDM). This approach is central to a survey of the design space for solar sail trajectories in the circular restricted Earth–Moon system. Through this survey, millions of trajectory options are revealed that can accommodate a solar sail relay spacecraft for communications with the LSP. Clearly, technological advances are required to enable a solar sail to serve as a relay. The FDM is also adaptable to transition solutions in the Earth–Moon system to higher-fidelity representations, such as an ephemeris-based model.

Regardless of the fidelity of any resulting trajectory or the level of sail technology, the motion of the vehicle must be controlled along a representative trajectory. Again, the use of numerical methods for solving boundary value problems is employed, this time for flight-path control. Motion along a solar sail trajectory is extremely sensitive to errors. Strategies that incorporate multiple turns over a give arc to achieve a desired state at the end of that arc are generally preferable to those that enlist only one orientation throughout the arc.

8.1 Survey of the Earth–Moon design space for solar sail spacecraft

A plethora of solutions emerges for the LSP coverage problem when various combinations of sail characteristic acceleration, initial guesses for the path and control profile, and path constraints are examined. Solar sail spacecraft supply a solution to the LSP coverage problem that requires only one space vehicle. By employing survey techniques, an understanding of the design space emerges. For a minimum elevation-angle constraint of 15°, a characteristic acceleration of 1.6 mm/s^2 is required to maintain trajectories below L_1 and the Moon, and a characteristic acceleration of 1.3 mm/s^2 is required to orbit below L_2. Solutions exist that both (1) require turning torques that are assumed to be possible for future solar sail spacecraft and (2) do not require excessively large pitch angles with respect to the Sun. Furthermore, if a stationary antenna with a fixed field of view is employed at a facility at the LSP, trajectory options that locate the sailcraft within a 70° swath width exist located below the Lagrange points. These results indicate that advances in current solar sail technology are necessary for a sailcraft to serve as a relay for an outpost at the LSP.

Any numerical method may be used as the engine in a search of the design space. A FDM scheme is particularly well suited for this task in the Earth–Moon system because the algorithm does not require a good initial guess for the trajectory and control profile, nor is it particularly complicated to develop. Furthermore, processor times for this algorithm are small, and a survey of tens of millions of initial guesses in the Earth–Moon system is completed in a few days using multiple computing platforms.

8.2 Numerical methods to generate displaced trajectories in dynamically challenging regimes

Finite-difference methods are often overlooked as options for trajectory development because of their lower accuracy in comparison to other numerical tools. However, FDM schemes are useful tools for generating trajectories in poorly understood dy-

namical regimes, such as the mechanics of flying a solar sail in a multi-body system. Perhaps the greatest advantage of FDM schemes is that they are simple to understand and implement, notably in the presence of path constraints; the fundamental relationships in the FDM algorithm are velocities and accelerations immediately available from the equations of motion and velocities and accelerations easily deduced from central-difference approximations. As a result, FDM schemes are able to return results quickly in a regime where little intuition concerning the trajectory and sail-angle history exists. The FDM approach developed for this analysis is easily adaptable to other regimes, such as an ephemeris-based, inertial system. Likewise, FDM schemes need not be applied to just solar sail problems, but should be applicable to a large variety of mission design problems.

Higher-fidelity methods, such as collocation and shooting methods, are useful on their own in generating trajectories and supplying insight into many categories of problems. However, these methods may be complicated or require an accurate initial guess. Results from FDM schemes may be employed to initialize these higher-fidelity methods to refine the path or to furnish a reference path for flight-path control. All three categories of numerical algorithms are useful tools for trajectory designers.

8.3 Flight-path control for solar-sail trajectories

In this investigation, turn-and-hold flight-path control schemes are developed based on a variety of shooting strategies for solving two-point boundary value problems. Solar sail spacecraft possess fewer degrees of control than degrees of freedom. To target a future position and velocity along a reference path at some future epoch requires either (1) selecting a single orientation that best accomplishes the targeting or (2) determining a set of orientations that result a trajectory that tracks the reference path. In both strategies, the time between turns is essentially an additional control parameter.

In the first approach, a least-squares update is exploited to track the reference path for a three-year mission. No errors are examined in this strategy because the controlled path is not guaranteed to converge with the reference path. In addition to examining different hold times between turns, a variety of target points along the reference path are employed in the targeting of a future state. The utility of this scheme appears to depend on the selection of the reference path; some paths are more amenable than others and these orbits are successfully tracked for the full three years. However, this course is not recommended as a primary strategy for solar sail flight path control because its use often leads to a trajectory that diverges from the desired path into a region in which the vehicle is unrecoverable.

In the second approach, both a collocation-based scheme and a multiple-shooting technique are employed to target the vehicle to a position and velocity at some future epoch by executing three or more turns along an arc ending at the target state. The collocation scheme uses MATLAB®'s BVP6C algorithm and only permits three turns per arc. This technique is successful for tracking only one reference path for the three-year mission and is not recommended.

The multiple-shooting approach is generally the most robust technique for targeting future states because it supplies an excess of control authority. Between 3 and 12 turns per arc are examined. Again, the number of days between turns is varied. In actual flight, errors in the attitude modeling, knowledge, and control, as well as orbit determination uncertainty, will persist, and two strategies are developed to accommodate these issues. The "look-ahead" strategy capably tracks the reference path for most hold-time and number-of-turn scenarios for each sample trajectory. Similar to other strategies, success depends on the reference path itself and selecting the right combination of hold times and turns. The adaptive strategy is also successful but requires shorter maximum durations between turns for a favorable outcome.

All strategies require constant monitoring of the vehicle. Motion along a sailcraft trajectory is extremely sensitive to uncertainties and errors in the attitude and state

knowledge as well as the control. Efforts must be made to ensure that these errors and uncertainties are minimal prior to any sailcraft flight in a LSP coverage mission.

8.4 Recommendations for future work

In the course of this study, several areas for future investigation are identified. These areas are broadly divided into (1) space vehicle dynamics, navigation, and control and (2) applications of numerical tools.

8.4.1 Dynamics, navigation, and control of a solar sail vehicle

Natural extension of this investigation of solar sail trajectories are the improved strategies for modeling their dynamics, navigation studies to support mission planning, and additional approaches for flight-path control.

Improved model fidelity After adding the ephemerides, it is believed that improving the solar sail reflectivity model will provide the next greatest improvement in the accuracy of the equations of motion. Effects of an optical sail model are discussed in Appendix A. Improvements to the flight-path control algorithm include modeling the ramp-up effects of the turns in a turn-and-hold control scheme.

Navigation A thorough investigation into the ground operations of a solar sail spacecraft would add insight into the feasibility of the relay mission. Orbit determination covariance studies and maneuver planning could limit the viability of such a mission. In that case, on-board autonomous navigation and turn design may be required.

Flight-path control An additional turn-and-hold control schemes that should be investigated is a multiple step, optimal maneuver strategy developed by Dwivedi [195,196] and used by Gordon [197], Howell and Pernicka [193], and A. McInnes

[118]. Linear quadratic regulators supply control laws for a continuously reorienting sailcraft. Likewise, polynomials can be fit to continuous reference profiles; coefficients can be adjusted to ensure that the flown trajectory remains close to the reference path.

8.4.2 Applications of numerical tools

The numerical methods enlisted in this research may be used to determine trajectories based on alternative propulsion devices to solve the LSP coverage problem or in other complicated dynamical systems. Additional approaches may reveal new options for the LSP application.

Alternatives to solar sails Solar sails have only been recently demonstrated in space. As mentioned, one avenue for their evolution is via a hybrid sail/low-thrust system. These hybrid devices have been proposed for both lunar and terrestrial polesitter missions. Low-thrust propulsion alone can also solve the polesitter problem. Another, untested, concept for non-Keplerian orbits are electric sails. The FDM algorithm, survey methods, and control techniques used in this investigation are all suitable for studying the feasibility of these propulsion devices at the LSP or other mission categories, such as the terrestrial poles, transfer trajectories from the Earth to other bodies, navigating between different multi-body systems (i.e., Sun–Earth to Earth–Moon), or even the non-Keplerian orbits in the vicinity of moons in orbit about a gas giant.

Alternative methods to produce reference trajectories No parameters are optimized in this investigation. A multi-objective function that minimizes characteristic acceleration or the specific transverse torques, subject to elevation and altitude constraints, may help uncover trajectories that do not require large improvements to the current state of sail technology. A genetic algorithm could be employed in the search for these optimal orbits.

8.5 Epilogue

By 1609, Kepler developed his first two laws of planetary motion, ensuring that elliptical paths about a central body would be known as "Keplerian" orbits. Solar sail trajectories are often labeled "non-Keplerian" because they do not follow elliptical paths. Displaced orbits, such as the solar sail trajectories in this investigation, are often denoted as highly non-Keplerian. Given the dream Kepler shared with Galileo in 1610 of "sails adapted to the breezes of heaven," perhaps it is time to reconsider the definition of "non-Keplerian."

LIST OF REFERENCES

LIST OF REFERENCES

[1] G. Vulpetti, L. Johnson, and G. L. Matloff, *Solar Sails: A Novel Approach to Interplanetary Travel*. New York: Springer, 2008.

[2] R. L. Forward, "Statite: A spacecraft that does not orbit," *Journal of Spacecraft and Rockets*, vol. 28, pp. 606–611, September–October 1991.

[3] C. R. McInnes, "Solar sail trajectories at the lunar L_2 Lagrange point," *Journal of Spacecraft and Rockets*, vol. 30, pp. 782–784, November–December 1993.

[4] C. R. McInnes, A. J. C. McDonald, J. F. L. Simmons, and E. W. MacDonald, "Solar sail parking in restricted three-body systems," *Journal of Guidance, Control, and Dynamics*, vol. 17, pp. 399–406, March–April 1994.

[5] C. R. McInnes and J. F. L. Simmons, "Solar sail halo orbits II: Geocentric case," *Journal of Spacecraft and Rockets*, vol. 29, pp. 472–479, July–August 1992.

[6] C. R. McInnes, *Solar Sailing: Technology, Dynamics and Mission Applications*. Space Science and Technology, New York: Springer-Praxis, 1999.

[7] E. Morrow, D. Lubin, and D. Scheeres, "Solar sail orbit operations at asteroids," *Journal of Spacecraft and Rockets*, vol. 38, pp. 279–286, March–April 2001.

[8] M. Macdonald and C. R. McInnes, "Analytical control laws for planet-centered solar sailing," *Journal of Guidance, Control, and Dynamics*, vol. 28, pp. 1038–1048, September–October 2005.

[9] H. Baoyin and C. R. McInnes, "Solar sail orbits at artificial Sun–Earth libration points," *Journal of Guidance, Control, and Dynamics*, vol. 28, pp. 1328–1330, November–December 2005.

[10] J. Bookless and C. R. McInnes, "Dynamics and control of displaced periodic orbits using solar-sail propulsion," *Journal of Guidance, Control, and Dynamics*, vol. 29, pp. 527–537, May–June 2006.

[11] A. Farrés and À. Jorba, "Station keeping close to unstable equilibrium points with a solar sail," in *AAS/AIAA Astrodynamics Specialist Conference*, (Mackinac Island, Michigan), August 2007. Paper No. AAS 07-347.

[12] A. Farrés and À. Jorba, "Solar sail surfing along families of equilibrium points," *Acta Astronautica*, vol. 63, pp. 249–257, July–August 2008.

[13] A. Farrés and À. Jorba, "Dynamical system approach for the station keeping of a solar sail," *Journal of the Astronautical Sciences*, vol. 58, pp. 199–230, April–June 2008.

[14] J. Simo and C. R. McInnes, "Solar sail orbits at the Earth–Moon libration points," *Communications in Nonlinear Science and Numerical Simulation*, vol. 14, pp. 4191–4196, December 2009.

[15] J. Simo and C. R. McInnes, "Asymptotic analysis of displaced lunar orbits," *Journal of Guidance, Control, and Dynamics*, vol. 32, pp. 1666–1671, September–October 2009.

[16] J. Simo and C. R. McInnes, "Designing displaced lunar orbits using low-thrust propulsion," *Journal of Guidance, Control, and Dynamics*, vol. 33, pp. 259–265, January–February 2010.

[17] J. Simo and C. R. McInnes, "On the stability of approximate displaced lunar orbits," in *AAS/AIAA Spaceflight Mechanics Meeting*, (San Diego, California), February 2010. Paper No. AAS 10-181.

[18] J. Simo and C. R. McInnes, "Displaced solar sail orbits: Dynamics and applications," in *AAS/AIAA Spaceflight Mechanics Meeting*, (San Diego, California), February 2010. Paper No. AAS 10-222.

[19] J. Simo and C. R. McInnes, "Displaced periodic orbits with low-thrust propulsion in the Earth–Moon system," in *AAS/AIAA Spaceflight Mechanics Meeting*, (Savannah, Georgia), February 2009. Paper No. AAS 09-153.

[20] "The Vision for Space Exploration," Tech. Rep. NP-2004-01-334-HQ, National Aeronautics and Space Administration, February 2004.

[21] "NASA Space Communication and Navigation Architecture Recommendations for 2005–2030." Space Communication Architecture Working Group (SCAWG), National Aeronautics and Space Administration, Washington, DC, May 2006.

[22] J. Schier, J. Rush, P. Vrotsos, and W. Williams, "Space communication architecture supporting exploration and science: Plans & studies for 2010-2030," in *1st Space Exploration Conference: Continuing the Voyage of Discovery*, (Orlando, Florida), January 2005. Paper No. AIAA-2005-2517.

[23] J. Gal-Edd and C. Fatig, "Lunar robotic relay the first phase of building the lunar ground network," in *Aerospace Conference, 2007 IEEE*, pp. 1–7, March 2007.

[24] M. Flanegan, J. Gal-Edd, L. Anderson, J. Warner, T. Ely, B. Shah, A. Vaisnys, and J. Schier, "NASA lunar communication and navigation architecture," in *SpaceOps 2008 Conference*, (Heidelberg, Germany), May 2008. Paper No. AIAA-2008-3589.

[25] T. A. Ely and E. Lieb, "Constellations of elliptical inclined lunar orbits providing polar and global coverage," *Advances in the Astronautical Sciences*, vol. 123, pp. 1447–1462, August 2005. Paper No. AAS 05-343.

[26] T. A. Ely, "Stable constellations of frozen elliptical inclined lunar orbits," *Journal of the Astronautical Sciences*, vol. 53, pp. 301–316, July–September 2005.

[27] K. C. Howell, D. J. Grebow, and Z. P. Olikara, "Design using Gauss' perturbing equations with applications to lunar south pole coverage," *Advances in the Astronautical Sciences*, vol. 127, pp. 641–670, February 2007. Paper No. AAS 07-143.

[28] D. J. Grebow, M. T. Ozimek, K. C. Howell, and D. C. Folta, "Multibody orbit architectures for lunar south pole coverage," *Journal of Spacecraft and Rockets*, vol. 45, pp. 344–358, March–April 2008.

[29] J. L. West, "The lunar polesitter," in *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, August 2008. Paper No. AIAA-2008-7073.

[30] M. T. Ozimek, D. J. Grebow, and K. C. Howell, "Design of solar sail trajectories with applications to lunar south pole coverage," *Journal of Guidance, Control, and Dynamics*, vol. 32, pp. 1884–1897, November–December 2009.

[31] D. J. Grebow, M. T. Ozimek, and K. C. Howell, "Design of optimal low-thrust lunar pole-sitter missions," *Advances in the Astronautical Sciences*, vol. 134, pp. 741–760, February 2009. Paper No. AAS 09-148.

[32] D. J. Grebow, M. T. Ozimek, and K. C. Howell, "Advanced modeling of optimal low-thrust pole-sitter trajectories," *Acta Astronautica*, vol. 67, pp. 991–1001, October–November 2010.

[33] M. T. Ozimek, D. J. Grebow, and K. C. Howell, "A collocation approach for computing solar sail lunar pole-sitter orbits," in *AAS/AIAA Astrodynamics Specialists Conference*, (Pittsburgh, Pennsylvania), August 2009. Paper No. AAS 09-378.

[34] G. G. Wawrzyniak and K. C. Howell, "Generating solar sail trajectories in the Earth–Moon system using augmented finite-difference methods," *International Journal of Aerospace Engineering*, vol. 2011, 2011. Article ID 4776197, 13 pages.

[35] G. G. Wawrzyniak and K. C. Howell, "Investigating the design space for solar sail trajectories in the Earth–Moon system," *The Open Aerospace Engineering Journal*, 2011. In Press, Corrected Proof.

[36] J. Kepler and E. Rosen (translator), *Kepler's Conversation with Galileo's Sidereal Messenger*. No. 5 in The Sources of Science, New York: Johnson Reprint Corportation, 1965.

[37] J. Kepler, *De Cometis Libelli Tres*. Augsburg, Germany: A. Apergeri, 1619.

[38] P. Lebedew, "The pressure of light on gases," *Astrophysical Journal*, vol. 31, pp. 385–393, June.

[39] J. C. Maxwell, *A Treatise on Elecricity and Magnetism*, vol. 2. London: MacMillian and Co., 1873.

[40] A. Bartoli, "Il calorico raggiante e il secondo principio di termodynamica," *Nuovo Cimento*, vol. 15, pp. 196–202, 1884.

[41] P. Lebedev, "Untersuchungen uber die druckkr afte des lichtes," *Annalen der Physik*, vol. 311, no. 11, pp. 433–458, 1901.

[42] E. F. Nichols and G. F. Hull, "The pressure due to radiation," *The Astrophysical Journal*, vol. 17, pp. 315–351, June 1903.

[43] C. R. McInnes and J. F. L. Simmons, "Solar sail halo orbits I: Heliocentric case," *Journal of Spacecraft and Rockets*, vol. 29, pp. 466–471, July–August 1992.

[44] C. L. Yen, "Solar sail geostorm warning mission design," *Advances in the Astronautical Sciences*, vol. 119, pp. 69–82, February 2004. Paper No. AAS 04-107.

[45] C. G. Sauer, "The L1 diamond affair," *Advances in the Astronautical Sciences*, vol. 119, pp. 2791–2808, February 2004. Paper No. AAS 04-278.

[46] M. Lisano, D. Lawrence, and S. Piggott, "Solar sail transfer trajectory design and stationkeeping control for missions to the sub-L1 equilibrium region," *Advances in the Astronautical Sciences*, vol. 120, pp. 1837–1854, 2005.

[47] D. Lawrence and S. Piggott, "Solar sailing trajectory control for sub-L1 stationkeeping," in *AIAA Guidance, Navigation, and Control Conference and Exhibit*, (Providence, Rhode Island), August 2004. Paper No. AIAA-2004-5014.

[48] T. J. Waters and C. R. McInnes, "Invariant manifolds and orbit control in the solar sail three-body problem," *Journal of Guidance, Control, and Dynamics*, vol. 31, pp. 554–562, May–June 2008.

[49] C. J. Sauer, "Optimal solar-sail interplanetary trajectories," in *AAS/AIAA Astrodynamics Conference*, August 1976. Paper No. AIAA-1976-792.

[50] M. Otten and C. R. McInnes, "Near minimum-time trajectories for solar sails," *Journal of Guidance, Control, and Dynamics*, vol. 24, pp. 632–634, March–April 2001.

[51] G. Mengali and A. A. Quarta, "Solar sail trajectories with piecewise-constant steering laws," *Aerospace Science and Technology*, vol. 13, pp. 431–441, December 2009.

[52] B. Dachwald, G. Mengali, A. A. Quarta, and M. Macdonald, "Parametric model and optimal control of solar sails with optical degredation," *Journal of Guidance, Control, and Dynamics*, vol. 29, pp. 1170–1178, September–October 2006.

[53] G. Mengali and A. A. Quarta, "Optimal three-dimensional interplanetary rendezvous using nonideal solar sail," *Journal of Guidance, Control, and Dynamics*, vol. 28, pp. 173–177, January–February 2005.

[54] G. Mengali and A. A. Quarta, "Optimal control laws for axially symmetric solar sails," *Journal of Spacecraft and Rockets*, vol. 42, pp. 1130–1133, November–December 2005.

[55] Z. Maorui, W. Guanglun, and S. Yong, "Solar sail trajectory optimization for Earth-Mars mission," in *Control Conference (CCC), 2010 29th Chinese*, pp. 139–143, July 2010.

[56] A. L. Herman, *Improved Collocation Methods with Application to Direct Trajectory Optimization*. Urbana, Illinois: Ph.D. Dissertation, University of Illinois at Urbana-Champaign, September 1995.

[57] D. J. Grebow, M. T. Ozimek, and K. C. Howell, "Advanced modeling of optimal low-thrust pole-sitter trajectories," in *60^{th} International Astronautical Congress*, (Daejeon, Republic of Korea), October 2009. Paper No. IAC-09-C1.5.4.

[58] D. L. Shirley, "The Mariner 10 mission to Venus and Mercury," *Acta Astronautica*, vol. 53, pp. 375–385, August–November 2003.

[59] D. J. O'Shaughnessy, J. V. McAdams, K. E. Williams, and B. R. Page, "Fire Sail: MESSENGER's use of solar radiation pressure for accurate Mercury flybys," *Advances in the Astronautical Sciences*, vol. 133, pp. 61–76, 2009. Paper No. AAS 09-014.

[60] G. Colombo, "The stabilization of an artificial satellite at the inferior conjunction point of the Earth–Moon system," Tech. Rep. 80, Smithsonian Astrophysical Observatory, November 1961.

[61] R. H. Laprade, J. A. Miller, and S. J. Worley, "Satellite stationkeeping by solar radiation pressure," *Space/Aeronautics*, vol. 47, pp. 114–117, April 1967.

[62] W. L. Black, M. C. Crocker, and E. H. Swenson, "Stationkeeping a 24-hour satellite using solar radiation pressure," *Journal of Spacecraft*, vol. 5, pp. 335–337, March 1968.

[63] R. W. Farquhar, "The control and use of libration point satellites," Tech. Rep. TR-R-346, National Aeronautics and Space Administration, February 1970.

[64] M. C. Crocker, "Attitude control of a sun-pointing spinning spacecraft by means of solar radiation pressure," *Journal of Spacecraft and Rockets*, vol. 7, pp. 757–759, March 1970.

[65] V. J. Modi and K. Kumar, "Attitude control of satellites using the solar radiation pressure," *Journal of Spacecraft and Rockets*, vol. 9, pp. 711–713, September 1972.

[66] B. W. Stuck, "Solar pressure three-axis attitude control," *Journal of Guidance and Control*, vol. 3, pp. 132–139, February 1980.

[67] O. Mori, Y. Tsuda, H. Sawada, R. Funase, T. Yamamoto, T. Saiki, K. Yonekura, H. Hoshino, H. Minamino, T. Endo, J. Kawaguchi, and IKAROS Demonstration Team, "Worlds first demonstration of solar power sailing by IKAROS," in *2nd International Symposium on Solar Sailing, New York City College of Technology, City University of New York*, (Brooklyn, New York), July 2010.

[68] M. Nakamiya, Y. Tsuda, and Y. Kawakatsu, "A study of the guidance method for the small solar power sail demonstrator, IKAROS," in *22nd International Symposium on Space Flight Dynamics*, (São José dos Campos, Brazil), February 2011.

[69] T. Yamaguchi, H. Ikeda, Y. Mimasu, Y. Tsuda, H. Takeuchi, and M. Yoshikawa, "Solar sail force modeling for spinning solar sail using the radiometric tracking data," in *22nd International Symposium on Space Flight Dynamics*, (São José dos Campos, Brazil), February 2011.

[70] A. Mann, "Solar sails pick up speed." *Nature*. [Website, accessed 14 June 2011], http://www.nature.com/news/2011/110203/full/news.2011.68.html, February 2011.

[71] "NASA - NanoSail-D Homepage." [Website, accessed 13 June 2011], http://www.nasa.gov/mission_pages/smallsats/nanosaild.html, 2011.

[72] L. Johnson, M. Whorton, A. Heaton, R. Pinson, G. Laue, and C. Adams, "NanoSail-D: A solar sail demonstration mission," *Acta Astronautica*, vol. 68, no. 5–6, pp. 571–575, 2011. Special Issue: Aosta 2009 Symposium.

[73] "NANOSAIL-D Satellite details 2011-999A NORAD 90027." [Website, accessed 13 June 2011], http://www.n2yo.com/satellite/?s=90027, 2011.

[74] L. Johnson, R. Young, E. Montgomery, and D. Alhorn, "Status of solar sail technology within NASA," *Advances in Space Research*, 2010. In Press, Corrected Proof. DOI: 10.1016/j.asr.2010.12.011.

[75] J. C. Mankins, "Technology readiness levels: A white paper." Advanced Concepts Office, Office of Space Access and Technology, NASA, April 1995.

[76] M. Meyer, L. Johnson, B. Palaszewski, D. Goebel, H. White, and D. Coote, "NASA In-Space Propulsion Systems Technology Area Roadmap – TA02." National Aeronautics and Space Administration, Washington, DC, November 2010.

[77] L. D. Freidman, "LightSail: A new way and a new chance to fly on light." [Website, accessed 12 November 2010], http://www.planetary.org/programs/projects/solar_sailing/tpr_lightsail.html, 2011.

[78] M. Macdonald and C. R. McInnes, "Solar sail science mission applications and advancement," *Advances in Space Research*, 2011. In Press, Corrected Proof. DOI: 10.1016/j.asr.2011.03.018.

[79] D. Lichodziejewski, B. Derbes, D. Sleight, and T. Mann, "Vacuum deployment and testing of a 20m solar sail system," in *47th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, (Newport, Rhode Island), May 2006. Paper No. AIAA-2006-1705.

[80] Japan Aerospace Exploration Agency, "JAXA — small solar power sail demonstrator 'IKAROS' confirmation of photon acceleration." [Website, accessed 23 June 2011], http://www.jaxa.jp/press/2010/07/20100709_ikaros_e.html, July 2010.

[81] C. Adams. Technical communications, 6 August 2008.

[82] V. Szebehely, *Theory of Orbits: the Restricted Problem of Three Bodies*. New York: Academic Press, 1967.

[83] D. T. Greenwood, *Principles of Dynamics*. Upper Saddle River, New Jersey: Prentice Hall, 2nd ed., 1988.

[84] A. E. Roy, *Orbital Motion*. Bristol, United Kingdom: Institute of Physics, 4th ed., 2004.

[85] D. Vallado, *Fundamentals of Astrodynamics and Applications*. Hawthorne, California: Microcosm Press, 2007.

[86] A. E. Roy and M. W. Ovenden, "On the occurrence of commensurable mean motions in the solar system: II. The mirror theorem," *Monthly Notices of the Royal Astronomical Society*, vol. 115, no. 3, pp. 296–309, 1955.

[87] A. Farrés and À. Jorba, "Periodic and quasi-periodic motions of a solar sail close to SL_1 in the Earth–Sun system," *Celestial Mechanics and Dynamical Astronomy*, vol. 107, pp. 233–253, June 2010.

[88] J. M. Longuski, J. J. Guzmán, and J. E. Prussing, *Optimization of Space Trajectories*. West Lafayette, Indiana: Purdue University, 2011. Unpublished manuscript.

[89] S. D. Conte and C. de Boor, *Elementary Numerical Analysis*. New York: McGraw-Hill, 1980.

[90] J. Stoer and R. Bulirsch, *Introduction to Numerical Analysis*. New York: Springer-Verlag, 3rd ed., 2002.

[91] K. C. Howell and H. J. Pernicka, "Numerical determination of Lissajous trajectories in the restricted three-body problem," *Celestial Mechanics*, vol. 41, no. 1–4, pp. 107–124, 1988.

[92] H. B. Keller, *Numerical Methods for Two-Point Boundary-Value Problems*. Waltham, Massachusettes: Blaisdell Publishing, 1968.

[93] D. R. Kincaid and E. W. Cheney, *Numerical Analysis: Mathematics of Scientific Computing*. Providence, Rhode Island: American Mathematical Society, 2002.

[94] G. G. Wawrzyniak and K. C. Howell, "Numerical methods to generate solar sail trajectories," in *2nd International Symposium on Solar Sailing, New York City College of Technology, City University of New York*, (Brooklyn, New York), pp. 195–200, July 2010.

[95] G. G. Wawrzyniak and K. C. Howell, "Numerical techniques for generating and refining solar sail trajectories," *Advances in Space Research*, 2011. In Press, Corrected Proof.

[96] E. W. Brown, *An Introductory Treatise on Lunar Theory*. London: Cambridge University Press, 1896.

[97] D. L. Richardson, "Analytic construction of periodic orbits about the collinear points," *Celestial Mechanics and Dynamical Astronomy*, vol. 22, pp. 241–253, October 1980.

[98] S. H. Strogatz, *Nonlinear Dynamics And Chaos: With Applications To Physics, Biology, Chemistry, And Engineering*. Cambridge, Massachusetts: Westview Press, 2001.

[99] D. Izzo, "Global optimization and space pruning for spacecraft trajectory design," in *Spacecraft Trajectory Optimization* (B. A. Conway, ed.), (New York), Cambridge University Press, 2010.

[100] M. Pontani and B. A. Conway, "Swarming theory applied to space trajectory optimization," in *Spacecraft Trajectory Optimization* (B. A. Conway, ed.), (New York), Cambridge University Press, 2010.

[101] R. Wilson and K. C. Howell, "Trajectory design in the Sun-Earth-Moon system using multiple lunar gravity assists," *Journal of Spacecraft and Rockets*, vol. 35, pp. 191–198, March–April 1998.

[102] B. Marchand, K. C. Howell, and R. S. Wilson, "An improved corrections process for constrained trajectory design in the n-body problem," *Journal of Spacecraft and Rockets*, vol. 44, pp. 884–897, July–August 2007.

[103] O. von Stryk and R. Bulirsch, "Direct and indirect methods for trajectory optimization," *Annals of Operations Research*, vol. 37, pp. 357–373, December 1992.

[104] J. T. Betts, "Survey of numerical methods for trajectory optimization," *Journal of Guidance, Control, and Dynamics*, vol. 22, pp. 193–207, March–April 1998.

[105] A. V. Rao, "A survey of numerical methods for trajectory optimization," *Advances in the Astronautical Sciences*, vol. 135, pp. 497–538, 2009. Paper No. AAS 09-334.

[106] C. R. Hargraves and S. W. Paris, "Direct trajectory optimization using nonlinear programming and collocation," *Journal of Guidance, Control, and Dynamics*, vol. 10, pp. 338–342, July–August 1987.

[107] P. J. Enright and B. A. Conway, "Optimal finite-thrust spacecraft trajectories using collocation and nonlinear programming," *Journal of Guidance, Control, and Dynamics*, vol. 14, pp. 981–985, September–October 1991.

[108] P. J. Enright and B. A. Conway, "Discrete approximations to optimal trajectories using direct transcription and nonlinear programming," *Journal of Guidance, Control, and Dynamics*, vol. 12, pp. 994–1002, July–August 1992.

[109] A. L. Herman and B. A. Conway, "Direct optimization using collocation based on high-order Gauss-Lobatto quadrature rules," *Journal of Guidance, Control, and Dynamics*, vol. 19, pp. 592–599, May–June 1996.

[110] R. D. Russell and L. F. Shampine, "A collocation method for boundary value problems," *Numerical Methods*, vol. 19, pp. 1–28, February 1972.

[111] E. Hairer, S. Nørsett, and G. Wanner, *Solving Ordinary Differential Equations I: Nonstiff problems*. New York: Springer-Verlag, 1987.

[112] E. J. Doedel, "Finite difference collocation methods for nonlinear two point boundary value problems," *SIAM Journal on Numerical Analysis*, vol. 16, pp. 173–185, April 1979.

[113] S. Tang and B. A. Conway, "Optimization of low-thrust interplanetary trajectories using collocation and nonlinear programming," *Journal of Guidance, Control, and Dynamics*, vol. 18, pp. 599–604, May–June 1995.

[114] R. G. Melton, "Comparison of direct optimization methods applied to solar sail problems," in *AIAA/AAS Astrodynamics Specialists Conference and Exhibit*, (Monterey, California), August 2002. Paper No. AIAA 2002-4728.

[115] D. G. Hull, "Conversion of optimal control problems into parameter optimization problems," in *AIAA/AAS Guidance, Navigation, and Control Conference*, (San Diego, California), July 1996. Paper No. AIAA 1996-3812.

[116] G. G. Wawrzyniak and K. C. Howell, "Accessing the design space for solar sails in the Earth–Moon system," *Advances in the Astronautical Sciences*, vol. 135, pp. 767–786, 2009. Paper No. AAS 09-348.

[117] J. S. Nuss, *The Use of Solar Sails in the Circular Restricted Problem of Three Bodies*. West Lafayette, Indiana: M.S. Thesis, Purdue University, May 1998.

[118] A. I. S. McInnes, *Strategies for Solar Sail Mission Design in the Circular Restricted Three-Body Problem*. West Lafayette, Indiana: M.S. Thesis, Purdue University, August 2000.

[119] T. J. Waters and C. R. McInnes, "Periodic orbits above the ecliptic in the solar-sail restricted three-body problem," *Journal of Guidance, Control, and Dynamics*, vol. 30, pp. 687–693, May–June 2007.

[120] V. Szebehely, *Theory of Orbits: The Restricted Problem of Three Bodies*. New York: Academic Press, 1967.

[121] S. J. Julier, J. K. Uhlmann, and H. F. Durrant-Whyte, "A new approach for filtering nonlinear systems," in *American Control Conference*, (Seattle, Washington), June 1995. Paper No. 5058100.

[122] P. Sengupta, S. R. Vadali, and K. T. Alfriend, "Second-order state transition for relative motion near perturbed, ellptic orbits," *Celestial Mechanics and Dynamical Astronomy*, vol. 97, pp. 101–129, January 2007.

[123] The MathWorks™, Inc., *MATLAB®, Version 7.11.0 (R2010b)*. August 2010. *www.mathworks.com*.

[124] I. M. Ross and F. Fahroo, "User's manual for DIDO 2002: A MATLAB package for dynamic optimization," Tech. Rep. AA-02-002, Department of Aeronautics and Astronautics, Naval Postgraduate School Technical Report, Naval Postgraduate School, Monterey, California, June 2002.

[125] A. V. Rao, D. A. Benson, G. T. Huntington, C. Francolin, C. L. Darby, and M. A. M. A. Patterson, *User's Manual for GPOPS: A MATLAB Package for Dynamic Optimization Using the Gauss Pseudospectral Method*. University of Florida, Naval Postgraduate School, Monterey, California, August 2008.

[126] A. V. Rao, D. A. Benson, C. L. Darby, M. A. Patterson, C. Francolin, and G. T. Huntington, "Algorithm 902: GPOPS, a MATLAB software for solving multiple-phase optimal control problems using the Gauss pseudospectral method," *ACM Transactions on Mathematical Software*, vol. 37, pp. 1–39, April–June 2010.

[127] N. Hale and D. R. Moore, "A sixth-order extension to the MATLAB package bvp4c of J. Kierzenka and L. Shampine," Tech. Rep. 08/04, Oxford University Computing Laboratory, Numerical Analysis Group, Oxford, April 2008.

[128] N. Nassiri, N. Mehdizadeh, and M. Jalali, "Interplanetary flight using solar sails," in *RAST 2005. 2nd International Conference on Recent Advances in Space Technologies*, (Istanbul, Turkey), pp. 330–334, June 2005.

[129] M. T. Ozimek, D. J. Grebow, and K. C. Howell, "A collocation approach for computing solar sail lunar pole-sitter orbits," *The Open Aerospace Engineering Journal*, vol. 3, pp. 65–75, 2010.

[130] M. T. Ozimek, *Low-Thrust Trajectory Design and Optimization of Lunar South Pole Coverage Missions*. West Lafayette, Indiana: Ph.D. Dissertation, Purdue University, May 2010.

[131] D. J. Grebow, *Trajectory Design in the Earth-Moon System and Lunar South Pole Coverage*. West Lafayette, Indiana: Ph.D. Dissertation, Purdue University, May 2010.

[132] J. T. Betts, *Practical Methods for Optimal Control Using Nonlinear Programming*. Philadelphia, Pennsylvania: SIAM, 2001.

[133] C. R. Ortiz Longo and S. L. Rickman, "Method for the calculation of spacecraft umbra and penumbra shadow terminator points," NASA Technical Paper 3547, Lyndon B. Johnson Space Center, Houston, Texas, April 1995.

[134] D. E. Salane, "Adaptive routines for forming Jacobians numerically," Tech. Rep. SAND86-1319, Sandia National Laboratories, Albuquerque, New Mexico, 1986.

[135] S. A. Forth and M. Edvall, *User Guide for MAD - A Matlab Automatic Differentiation Package, TOMLAB/MAD, Version 1.4 The Forward Mode*. Shrivenham, United Kingdom: Engineering Systems Department, Defence College of Management & Technology, Cranfield University, June 2007.

[136] L. F. Shampine, "Accurate numerical derivatives in MATLAB," *ACM Transactions on Mathematical Software*, vol. 33, August 2007.

[137] Y. Tsuda, O. Mori, R. Funase, H. Sawada, T. Yamamoto, T. Saiki, T. Endo, and J. Kawaguchi, "Flight status of IKAROS deep space solar sail demonstrator," in *61st International Astronautical Congress*, (Prague, Czech Republic), September 2010. Paper No. IAC-10-A3.6.8.

[138] B. Wie, "Solar sail attitude dynamics and control, part 1," *Journal of Guidance, Control, and Dynamics*, vol. 27, pp. 526–535, July–August 2004.

[139] R. Burton, V. Coverstone, J. Hargens-Rysanek, K. Ertmer, T. Botter, G. Benavides, B. Woo, D. Carroll, P. Gierow, G. Farmer, and J. Cardin, "Ultra-Sail - ultra-lightweight solar sail concept," in *AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, (Tucson, Arizona), July 2005.

[140] R. Funase, O. Mori, Y. Tsuda, Y. Shirasawa, T. Saiki, Y. Mimasu, and J. Kawaguchi, "Attitude control of IKAROS solar sail spacecraft and its flight results," in *61st International Astronautical Congress*, (Prague, Czech Republic), September 2010. Paper No. IAC-10-C1.4.3.

[141] B. Wie, "Solar sail attitude dynamics and control, part 2," *Journal of Guidance, Control, and Dynamics*, vol. 27, pp. 536–544, July–August 2004.

[142] M. Polites, J. Kalmanson, and D. Mangus, "Solar sail attitude control using small reaction wheels and magnetic torquers," *Proceedings of the Institution of Mechanical Engineers, Part G: Journal of Aerospace Engineering*, vol. 222, no. 1, pp. 53–62, 2008.

[143] H. W. Price, J. Ayon, C. Garner, G. Klose, E. Mettler, and G. Sprague, "Design for a solar sail demonstration mission," in *Space Technology and Applications International Forum (STAIF-2001)*, (Albuquerque, New Mexico), February 2001.

[144] B. Wie and D. Murphy, "Solar-sail attitude control design for a sail flight validation mission," *Journal of Spacecraft and Rockets*, vol. 44, pp. 809–821, July–August 2007.

[145] C. Scholz, D. Romagnoli, and B. Dachwald, "Performance analysis of an attitude control system for solar sails using sliding masses," in *2nd International Symposium on Solar Sailing, New York City College of Technology, City University of New York*, (Brooklyn, New York), pp. 171–176, July 2010.

[146] M. Nehrenz, A. Diaz, T. Svitek, and C. Biddy, "Initial design and simulation of the attitude determination and control system for LightSail-1," in *2nd International Symposium on Solar Sailing, New York City College of Technology, City University of New York*, (Brooklyn, New York), pp. 135–140, July 2010.

[147] A. Pukniel, V. Coverstone, J. Warner, and R. Burton, "A preliminary study of the dynamics and control of the CubeSail spacecraft," *Advances in the Astronautical Sciences*, vol. 135, pp. 1929–1948, August 2009.

[148] A. Pukniel, V. Coverstone, R. Burton, and D. Carroll, "Attitude control of the CubeSail solar sailing spacecraft in low Earth orbit," in *2nd International Symposium on Solar Sailing, New York City College of Technology, City University of New York*, (Brooklyn, New York), pp. 121–126, July 2010.

[149] B. Campbell, *An Analysis of Thrust of a Realistic Solar Sail with Focus on a Flight Validation Mission in a Geocentric Orbit.* Washington, DC: Ph.D. Dissertation, The George Washington University, May 2010.

[150] M. Leipold and M. Götz, "Hybrid photonic/electric propulsion," Tech. Rep. SOL4-TR-KTH-001, Kayser–Threde, GmBH, Munich, Germany, 2002.

[151] J. Kawaguchi, "A Power Sailer Mission for a Jovian Orbiter and Trojan Asteroid Flybys," in *35th COSPAR Scientific Assembly* (J.-P. Paillé, ed.), vol. 35, p. 1655, 2004.

[152] G. Mengali, A. Quarta, C. Circi, and B. Dachwald, "Refined solar sail force model with mission applications," *Journal of Guidance, Control, and Dynamics*, vol. 20, pp. 512–520, March–April 2007.

[153] S. Baig and C. R. McInnes, "Artificial three-body equilibria for hybrid low-thrust propulsion," *Journal of Guidance, Control, and Dynamics*, vol. 31, pp. 1644–1655, November–December 2008.

[154] M. Ceriotti and C. R. McInnes, "Systems design of a hybrid sail pole-sitter," *Advances in Space Research*, 2011. In Press, Corrected Proof.

[155] R. McKay, M. Macdonald, J. Biggs, and C. R. McInnes, "Survey of highly non-Keplerian orbits with low-thrust propulsion," *Journal of Guidance, Control, and Dynamics*, vol. 34, pp. 645–666, May–June 2011.

[156] G. G. Wawrzyniak and K. C. Howell, "Trajectory control for a solar sail spacecraft in an offset lunar orbit," in *61st International Astronautical Congress*, (Prague, Czech Republic), September 2010. Paper No. IAC-10.C1.2.3.

[157] B. Dachwald, M. Macdonald, C. R. McInnes, G. Mengali, and A. A. Quarta, "Impact of optical degradation on solar sail mission performance," *Journal of Spacecraft and Rockets*, vol. 44, pp. 740–749, July–August 2007.

[158] E. T. Whittaker, *A Treatise On the Analytical Dynamics of Particles and Rigid Bodies: With an Introduction to the Problem of Three Bodies*. Cambridge University Press, 4 ed., 1937. First edition published 1904.

[159] J. Barrow-Green, *Poincaré and the Three Body Problem*, vol. 11 of *History of Mathematics*. Providence, Rhode Island: American Mathematical Society, 1997.

[160] S. I. Newton, *PhilosophæNaturalis Principia Mathematica I–III*. London: The Royal Society, 1687. *Mathematical Principles of Natural Philosopy*, translated by Andrew Motte, 1729. Revised and edited by Florian Cajori, Univeristy of California Press, 1934.

[161] F. R. Moulton, D. Buchanan, T. Buck, F. Griffin, W. Longley, and W. MacMillan, *Periodic Orbits*. Washington, DC: Carnegie Insitution of Washington, 1920.

[162] R. W. Farquhar and A. A. Kamel, "Quasi-periodic orbits about the translunar libration point," *Celestial Mechanics*, vol. 7, pp. 458–473, 1970.

[163] D. Richardson and N. Cary, "A uniformly valid solution for motion about the interior libration point of the perturbed elliptic-restricted problem," *Advances in the Astronautical Sciences*, vol. 33, July 1975. Paper No. AAS 75-021.

[164] J. V. Breakwell and J. V. Brown, "The 'halo' family of 3-dimensional periodic orbits in the Earth–Moon restricted 3-body problem," *Celestial Mechanics and Dynamical Astronomy*, vol. 20, no. 4, pp. 389–404, 1979.

[165] K. C. Howell and J. V. Breakwell, "Almost rectilinear halo orbits," *Celestial Mechanics and Dynamical Astronomy*, vol. 32, no. 1, pp. 29–52, 1984.

[166] K. C. Howell, "Three-dimensional, periodic, halo orbits," *Celestial Mechanics and Dynamical Astronomy*, vol. 32, no. 1, pp. 53–71, 1984.

[167] D. J. Grebow, *Generating Periodic Orbits in the Circular Restricted Three-Body Problem with Applications to Lunar South Pole Coverage*. West Lafayette, Indiana: M.S. Thesis, Purdue University, May 2006.

[168] K. C. Howell, "Families of orbits in the vicinity of the collinear libration points," *Journal of the Astronautical Sciences*, vol. 49, pp. 107–125, January–March 2001.

[169] R. H. Farquhar, "Introduction," *Journal of the Astronautical Sciences*, vol. 49, pp. 1–9, January–March 2001. Special Issue: Libration-Point Missions.

[170] M. Michalodimitrakis, "The circular restricted four-body problem," *Astrophysics and Space Science*, vol. 75, pp. 289–305, 1981.

[171] D. J. Scheeres, "The restricted Hill four-body problem with applications to the Earth-Moon-Sun system," *Celestial Mechanics and Dynamical Astronomy*, vol. 70, pp. 75–98, February 1998.

[172] R. S. Wilson, *Trajectory Design in the Sun–Earth–Moon Four Body Problem*. West Lafayette, Indiana: Ph.D. Dissertation, Purdue University, May 1998.

[173] M. A. Andreu, *The Quasi-Bicircular Problem*. Barcelona, Spain: Ph.D. Dissertation, Universitat de Barcelona, October 1998.

[174] J. J. Guzmán, *Spacecraft Trajectory Design in the Context of a Coherent Restricted Four-Body Problem*. West Lafayette, Indiana: Ph.D. Dissertation, Purdue University, May 2001.

[175] J. J. Guzmán and K. C. Howell, "Periodic orbits in the vicinity of L1 and L2 in the coherent restricted four-body problem," in *2001 Flight Mechanics Symposium*, pp. 25–39, June 2001.

[176] F. Gabern and À. Jorba, "Restricted four and five body problems in the solar system," in *Proceedings of the International Conference on Libration Point Orbits and Applications*, June 2002.

[177] E. S. G. Leandro, "On the contral configurations of the planar restricted four-body problem," *Journal of Differential Equations*, vol. 226, pp. 323–351, July 2006.

[178] T. J. Kalvouridis, M. Arribas, and A. Elipe, "Parametric evolution of periodic orbits in the restricted four-body problem with radiation pressure," *Planetary and Space Science*, vol. 55, pp. 475–493, March 2007.

[179] A. Farrés and À. Jorba, "Solar sail surfing along families of equilibrium points," in *58th International Astronautics Congress*, (Hyderabad, India), October 2007.

[180] J. Simo and C. R. McInnes, "Solar sail trajectories at the Earth–Moon Lagrange points," in *59th International Astronautical Congress*, (Glasgow, Scotland), 2008. Paper No. IAC-08.C1.3.13.

[181] G. G. Wawrzyniak and K. C. Howell, "The solar sail lunar relay system: An application of solar sails in the Earth–Moon system," in *59th International Astronautical Congress*, (Glasgow, Scotland), September 2008. Paper No. IAC-08.C1.3.14.

[182] J. J. Guzmán, *Spacecraft Trajectory Design in the Context of a Coherent Restricted Four-Body Problem*. West Lafayette, Indiana: Ph.D. Dissertation, Purdue University, May 2001.

[183] S.-S. Huang, "Very restricted four-body problem," Tech. Rep. TN D-501, National Aeronautics and Space Administration, September 1960.

[184] J. M. A. Danby, *Fundamentals of Celestial Mechanics*. Richmond, Virginia: Willmann-Bell, Inc., 2 ed., 1988.

[185] M. Macdonald and C. R. McInnes, "Solar sail mission applications and future advancement," in *2nd International Symposium on Solar Sailing, New York City College of Technology, City University of New York*, (Brooklyn, New York), pp. 1–26, July 2010.

[186] K. C. Howell, B. T. Barden, and M. W. Lo, "Application of dynamical systems theory to trajectory design for a libration point mission," *Journal of Astronautical Sciences*, vol. 45, pp. 161–178, April–June 1997.

[187] E. M. Standish, "JPL planetary and lunar ephemerides, DE405/LE405," Interoffice Memorandum IOM 312.F - 98 - 048, Jet Propulsion Laboratory, Pasadena, California, 1998.

[188] W. M. Folkner, E. M. Standish, J. G. Williams, and D. H. Boggs, "Planetary and lunar ephemerides DE 418," Interoffice Memorandum IOM 343R-07-005, Jet Propulsion Laboratory, Pasadena, California, 2007.

[189] W. M. Folkner, J. G. Williams, and D. H. Boggs, "The planetary and lunar ephemerides DE 421," Interoffice Memorandum IPN Progress Report 42-178, Jet Propulsion Laboratory, Pasadena, California, August 2009.

[190] Navigation and Ancillary Information Facility, "NAIF Mice toolkit hypertext documentation." [Website, accessed 2 July 2011], http://naif.jpl.nasa.gov/pub/naif/toolkit_docs/MATLAB/, June 2011.

[191] L. Rios-Reyes and D. J. Scheeres, "Solar-sail navigation: Estimation of force, moments, and optical parameters," *Journal of Guidance, Control, and Dynamics*, vol. 30, pp. 660–668, May–June 2007.

[192] H. Akima, "A new method of interpolation and smooth curve fitting based on local procedures," *Journal of the Association for Computing Machinery*, vol. 17, pp. 589–602, October 1970.

[193] K. C. Howell and H. J. Pernicka, "Stationkeeping method for libration point trajectories," *Journal of Guidance, Control, and Dynamics*, vol. 16, no. 1, pp. 151–159, 1993.

[194] G. C. Goodwin, S. F. Graebe, and M. E. Salgado, *Control System Design*. Upper Saddle River, New Jersey: Prentice Hall, 2001.

[195] N. P. Dwivedi, "Maneuver strategies for multiplanet missions," *AIAA Journal*, vol. 11, pp. 1229–1230, September 1973.

[196] N. P. Dwivedi, "Deterministic optimal maneuver strategy for multi-target missions," *Journal of Optimization Theory and Applications*, vol. 17, pp. 133–153, October 1975.

[197] S. C. Gordon, *Orbit Determination Error Analysis and Station-Keeping for Libration Point Trajectories*. West Lafayette, Indiana: Ph.D. Dissertation, Purdue University, December 1991.

[198] B. Derbes, D. L. Lichodziejewski, and G. Veal, "A 'yank and yaw' control system for solar sails," *Advances in the Astronautical Sciences*, vol. 119, pp. 2893–2908, February 2004.

[199] R. M. Georgevic, "The solar radiation pressure force and torques model," *Journal of the Astronautical Sciences*, vol. 20, pp. 257–274, March–April 1973.

[200] J. L. Wright, *Space Sailing*. Philadelphia, Pennsylvania: Gordon and Breach Science Publishers, 1992.

[201] G. G. Wawrzyniak, D. T. Baird, E. J. Graat, T. P. McElrath, B. M. Portock, and M. M. Watkins, "Mars Exploration Rovers orbit determination system modeling," *Journal of the Astronautical Sciences*, vol. 54, pp. 175–197, April–June 2006.

APPENDICES

A. Limitations of an idealized solar sail model

The sail model in this investigation is a perfectly reflecting, flat solar sail. Billowing is not incorporated in this force model, and the maximum pitch angle, $\alpha_{\max}$ is constrained to less than 90° to ensure that the sail-face normal is always directed away from the Sun. In reality, sail luffing (i.e., flapping) is assumed to occur at high pitch angles and the correspondence between an ideal model and a more realistic model diverges [198] effectively reducing the maximum available pitch angle. Higher-fidelity models include optical models [6], parametric models that incorporate billowing in addition to optical effects [6,191], and realistic models based on finite-element analysis that incorporates optical properties and manufacturing flaws [149].

The essence of an optical model (and consequently other higher-fidelity models) is that not all energy is reflected specularly from the sail [199]. In the energy balance,

$$1 = a + \rho s + \rho(1 - s) \tag{A.1}$$

some energy is absorbed (represented by a, the coefficient of absorptivity) and some is reflected (ρ, the coefficient of reflectivity). Of the reflected energy, some is reflected specularly (s, the specular fraction) and the rest diffusely $(1 - s)$. With specular reflection, the angle of reflection is equal to the angle of incidence (e.g., a mirror). Diffuse reflection assumes that light is scattered in all directions from the reflecting surface. It is assumed that no energy is transmitted though the sail. An objective in solar sail design is to maximize specular reflectivity of the sail film material while minimizing the mass of the material itself.

An illustration of the optical force model for a flat plate appears in Fig. A.1. In the figure, sunlight impinges on the surface imparting an incident force, $\mathbf{f}_\ell$, that is,

$$\mathbf{f}_\ell = PA\left(\cos^2\alpha\,\hat{\mathbf{u}} + \cos\alpha\sin\alpha\,\hat{\mathbf{t}}\right) \tag{A.2}$$

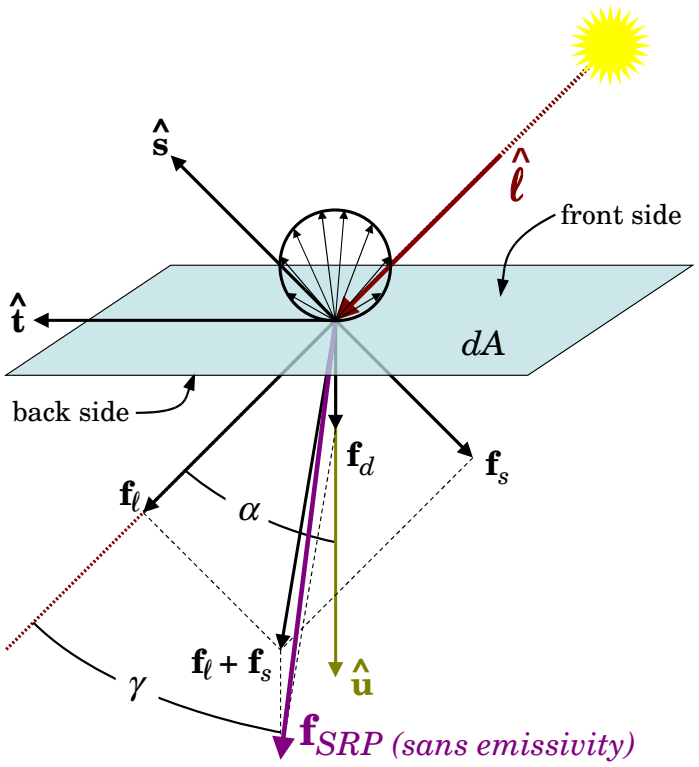


Fig. A.1: A differential element of a solar sail subject to various components of an optical force model. Forces from emitted radiation are omitted for clarity.

where P is the pressure due to solar radiation at some distance from the Sun, A is the area of the sail, $\hat{\mathbf{t}}$ is the transverse direction in the plane formed by the sunlight direction, $\hat{\boldsymbol{\ell}}$, and the sail-face normal, $\hat{\mathbf{u}}$. From Eq. 1.4, the specular force, $\mathbf{f}_s$ is opposite to the direction of specular reflection, $\hat{\mathbf{s}}$, also in the plane formed by $\hat{\boldsymbol{\ell}}$ and $\hat{\mathbf{u}}$. In a non-ideal model, $f_s < f_\ell$, because the specular fraction and the coefficient of reflectivity are both less than 1. The specular force is

$$\mathbf{f}_s = \rho s P A \left(\cos^2\alpha\,\hat{\mathbf{u}} - \cos\alpha\sin\alpha\,\hat{\mathbf{t}}\right) \tag{A.3}$$

Some reflected energy is reflected diffusely and the resulting force, $\mathbf{f}_d$, is directed along the sail-face normal, $\hat{\mathbf{u}}$.

$$\mathbf{f}_d = D_f \rho (1-s) P A \cos\alpha\,\hat{\mathbf{u}} \tag{A.4}$$

where D_f is the front-side diffuse-power fraction. Diffuse reflection is represented by the circle on top the differential area in Fig. A.1, and any forces from diffuse reflection in the $\hat{\mathbf{t}}$ or $\hat{\mathbf{u}} \times \hat{\mathbf{t}}$ directions cancel. The diffuse-power fraction compensates for this cancellation and attenuates the diffuse force.

Absorbed radiation is re-radiated, or emitted, as heat primarily via the back side of the sail. It is assumed to radiate in a diffuse pattern. The net force from the emitted radiation is modeled in the $-\hat{\mathbf{u}}$ direction, mitigating the net SRP force. The force from emitted radiation is derived from a thermal balance and is

$$\mathbf{f}_e = PA(1-\rho)\frac{\epsilon_f D_f - \epsilon_b D_b}{\epsilon_f + \epsilon_b}\cos\alpha\,\hat{\mathbf{u}} \tag{A.5}$$

where ϵ_f and ϵ_b are the front- and back-side surface emissivities and D_b is the back-side diffuse-power fraction.

The forces in Eqs. (A.2) through (A.5) are added together resulting in a net optical force. This net force is

$$\mathbf{f}_{SRP} = \mathbf{f}_\ell + \mathbf{f}_s + \mathbf{f}_d + \mathbf{f}_e \tag{A.6}$$

$$= f_u\hat{\mathbf{u}} + f_t\hat{\mathbf{t}} \tag{A.7}$$

In component form,

$$f_u = PA\left\{(1+\rho s)\cos^2\alpha + D_f(1-s)\rho\cos\alpha + (1-\rho)\frac{\epsilon_f D_f - \epsilon_b D_b}{\epsilon_f + \epsilon_b}\cos\alpha\right\} \tag{A.8}$$

$$f_t = PA(1-\rho s)\cos\alpha\sin\alpha \tag{A.9}$$

From Eqs. (A.8) and (A.9), it is apparent that if ρ and s are both equal to 1, the resulting solar radiation pressure (SRP) force will be aligned with $\hat{\mathbf{u}}$ and the cone angle, γ, which is the angle between $\mathbf{f}_{SRP}$ and $\hat{\boldsymbol{\ell}}$, will equal the pitch angle, α. The result is the ideal solar sail model from Eq. (2.8). However, because of absorption, thermal emission, and diffuse reflection, thus, the cone angle is less than or equal to the pitch angle (i.e., $\gamma \leq \alpha$).

Standard values for the parameters in an optical sail force model are supplied by Wright [200] and are based on a JPL study from the late 1970s for a sail mission to

Halley's Comet. These values are listed in Table A.1 for an ideal sail and the non-ideal sail. For comparison to a "traditional" spacecraft, the equivalent parameters for the cruise stage of the Mars Exploration Rovers (MER), composed of solar panels and a spacecraft bus, is also listed [201]. Note that the reflectivity of a solar sail spacecraft is

Table A.1: Optical parameters for an ideal sail, the non-ideal sail, and the MER cruise stage (MER C/S)

	ρ	s	ϵ_f	ϵ_b	D_f	D_b
Ideal sail	1	1	0	0	2/3	2/3
Non-ideal sail	0.88	0.94	0.05	0.55	0.79	0.55
MER C/S	0.26	0.58	N/A	N/A	2/3	N/A

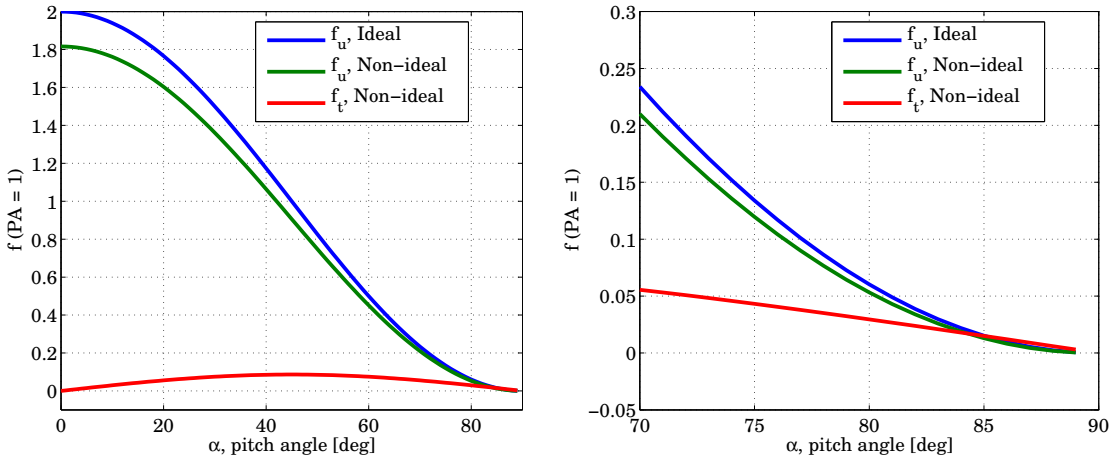
much greater than a traditional spacecraft and that the areal density is much greater as well. This larger value of areal density accounts for much of the difference in solar radiation pressure acceleration between a solar sail and a traditional spacecraft. The normal and transverse force components for an ideal and non-ideal sail as a function of pitch angle appear in Fig. A.2. The product of the pressure and the sail area, PA, from Eqs. (A.8) and (A.9) is equal to 1 in the figure. Note that because of the optical terms in Eqs. (A.3) through (A.5), the force on the sail at $\alpha = 0^\circ$ is approximately 90% of the force from an ideal sail.

Based on the values in Table A.1, it is reasonable to question the validity of the ideal sail model that is summarized by a characteristic acceleration from Eq. (1.6). An ideal sail reflects only specularly. In all non-ideal models, the resulting acceleration from a solar sail is not perfectly parallel to the sail-face normal but, instead, is increasingly offset from the sail-face normal as the sail is pitched further from the sunlight direction [6]. The difference between the pitch angle and the cone angle is

$$\alpha - \gamma = \arctan\left(\frac{f_t}{f_u}\right) \tag{A.10}$$

and increases as a function of pitch angle for the non-ideal model, as demonstrated in Fig. A.3, because $f_t \geq 0$. The cone angle as a function of pitch angle appears in Fig. A.4. For example, a pitch angle of $\alpha = 60°$ corresponds to a cone angle of 60° for an ideal sail and 50.6° for a non-ideal sail. It is apparent that the cone angle, which may be considered an *effective* pitch angle, achieves a maximum value of 55.5° when the true pitch angle is 72.6° for the non-ideal model. For high pitch angles, the reflective SRP forces contribute a smaller portion of the total force in the non-ideal model as compared to an ideal model. The result is that the incident SRP force, though mitigated by the cosine terms in Eq. (A.2), dominates the overall SRP force, both of which are now primarily in the transverse direction (as apparent in Fig. A.2).

It is appropriate to compare the net force on the sail from the ideal and non-ideal models as a function of cone angle, γ, to indicate the true direction of the control force supplied by the sail. Recall that $\gamma = \alpha$ for the ideal sail, but $\gamma \leq \alpha$ for the non-ideal sail. This comparison provides an assessment of the reflection efficiency, η, necessary to reflect the characteristic acceleration, a_c, accurately from Eq. (1.6). Recall that the characteristic acceleration is also a function of the spacecraft mass.



(a) For all pitch angles

(b) Zoom for high pitch angles

Fig. A.2: Force components verse α for an ideal and non-ideal sail

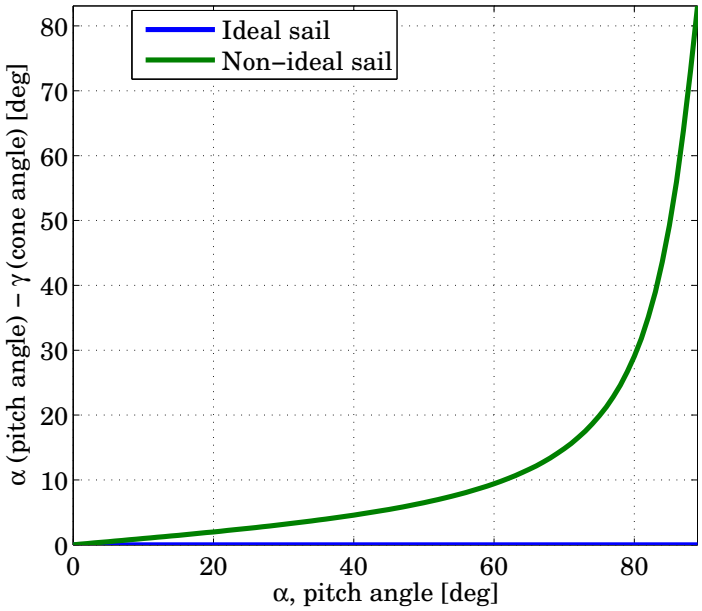


Fig. A.3: The difference between α and γ as a function of α for an ideal and non-ideal sail.

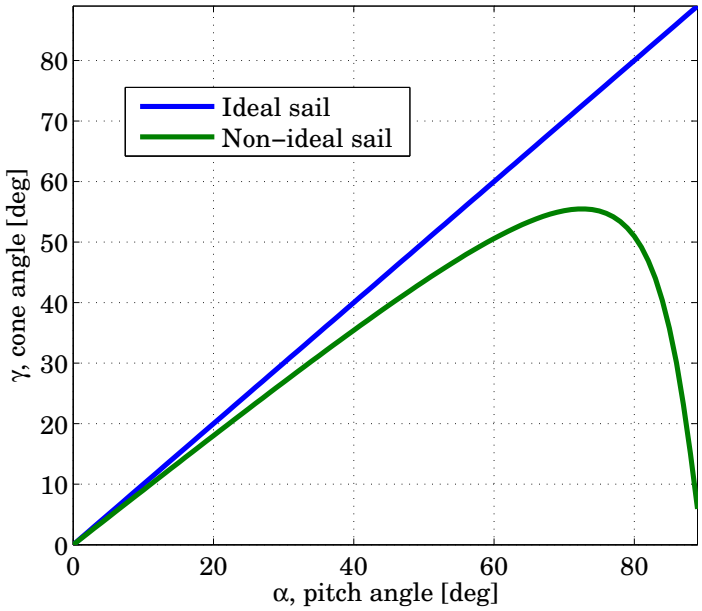
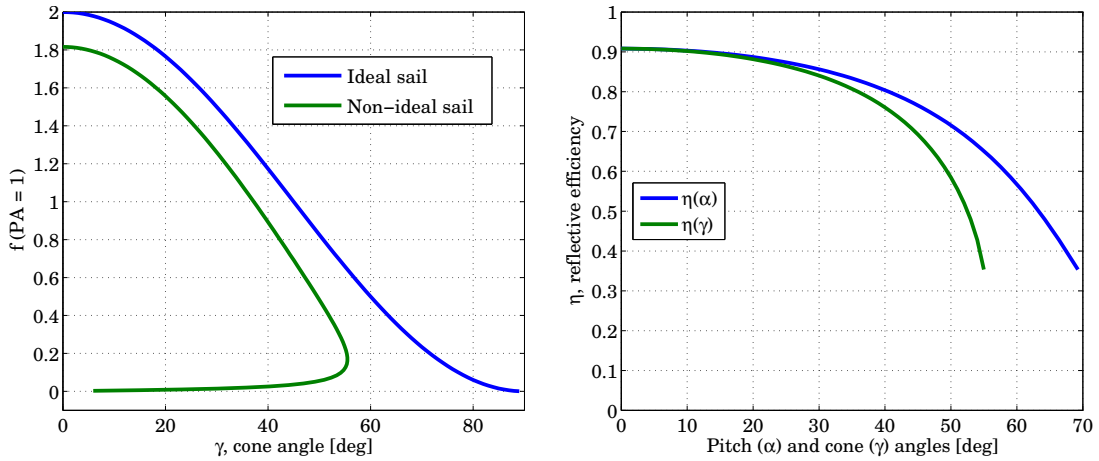


Fig. A.4: Cone angle, γ, as a function of pitch angle, α.



(a) Sail force

(b) Reflection efficiency

Fig. A.5: Sail force and associated reflection efficiency as a function of cone angle.

By dividing the force from the non-ideal sail as a function of its respective cone angle by that of the ideal sail as a function of its respective cone angle, a measure of the reflection efficiency is calculated. The forces as a function of their respective cone angles appear in Fig. A.5(a), and the reflection efficiency necessary to compute the associated characteristic acceleration appears in Fig. A.5(b). The reflection efficiency as a function of the corresponding pitch angle, α, also appears in Fig. A.5(b). For example, if the pitch angle is 60°, the corresponding reflection efficiency is 56.7%; the cone angle associated with these α and η values is $\gamma = 50.6°$. Based on Figs. A.4 and A.5, an assumption of an ideal sail is appropriate for low pitch angles.

If the reflection efficiency was independent of any other parameter, including sail orientation, an assumption of a constant characteristic acceleration a_c, would be appropriate. However, it must be noted that a_c neither fully captures nor accurately represents the reflective property of a realistic solar sail over all orientations. That said, if the expected range of pitch or cone angles is small over the duration of a

mission, an effective a_c appropriate to that range provides a sufficient initial approximation of reflectivity for the overall solar sail mission design problem.

Fully incorporating other realistic solar sail properties, such as billowing and manufacturing defects, further mitigates the sail efficiency [149]. Not only the magnitude of the force, but the effective direction of the force with respect to the sail face normal, depend on the pitch angle in the case of a non-ideal sail. Therefore, it is prudent to adapt the methods in this study to non-ideal sails in future research efforts. Nevertheless, this analysis employs an ideal sail to lend insight into the technology level that is required to solve the LSP coverage problem and into the use of numerical methods for generating and controlling solar sail trajectories.

B. Sail orientation transformations

A series of rotations is employed to transform a vector from a sailcraft body-fixed frame to the inertial frame [83]. These rotations aid in expressing the angular velocity vector of the body-fixed frame with respect to the inertial frame, ${}^{I}\boldsymbol{\omega}^{S}$. A variety of rotation sequences are enlisted to describe these rotations, however, it is useful to develop ${}^{I}\boldsymbol{\omega}^{S}$ based on existing orientation angles, such as α, δ, and Ωt (pitch, clock, and sunlight angle, respectively). This analysis assumes that the Sun's rays are parallel at 1 AU and that the Sun moves in a circle about the Earth–Moon barycenter as well as in the Earth–Moon orbit plane.

In any sequence of transformations, the intermediate coordinate frames are clearly defined. The first step in transforming from the inertial frame to a body-fixed frame is an initial transformation from the inertial frame, I, to a solar frame, F, where $\hat{\mathbf{x}}'$ is aligned with $\hat{\boldsymbol{\ell}}$, as evident in Fig. B.1. Because the rotating frame R is moving counterclockwise about a common $\hat{\mathbf{Z}}$ axis with respect to the inertial frame and the

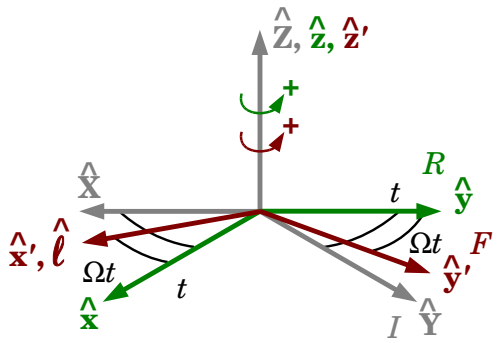


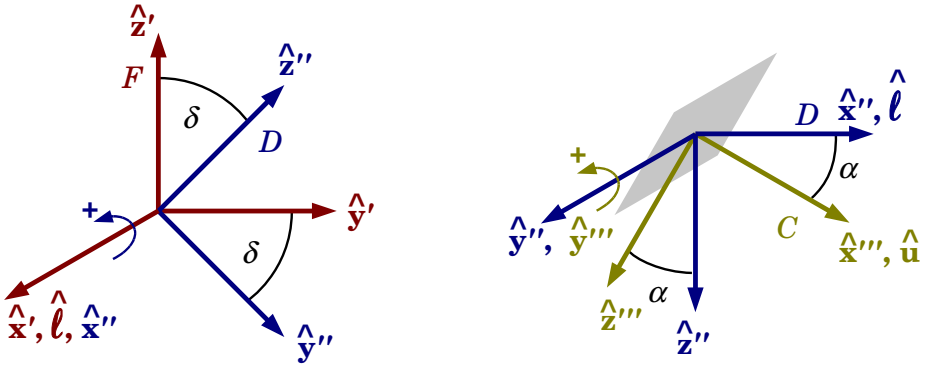
Fig. B.1: Rotations from the inertial frame, I, to the solar frame, F, via the Earth–Moon rotating frame, R.

Sun is moving clockwise about the same $\hat{\mathbf{z}}$ axis at a rate of Ω, the first two rotations are consolidated into a single rotation about the $\hat{\mathbf{Z}}$ axis such that

$$\begin{Bmatrix} \hat{\mathbf{x}}' \\ \hat{\mathbf{y}}' \\ \hat{\mathbf{z}}' \end{Bmatrix} = {}^{F}\mathbf{T}_3^I(t-\Omega t) \begin{Bmatrix} \hat{\mathbf{X}} \\ \hat{\mathbf{Y}} \\ \hat{\mathbf{Z}} \end{Bmatrix} \tag{B.1}$$

$${}^{F}\mathbf{T}_3^I\,(t-\Omega t) = \begin{bmatrix} \cos(t-\Omega t) & \sin(t-\Omega t) & 0 \\ -\sin(t-\Omega t) & \cos(t-\Omega t) & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{B.2}$$

where ${}^{F}\mathbf{T}_3^I(t-\Omega t)$ is a rotation about the $\hat{\mathbf{z}}$ (or $\hat{\mathbf{z}}'$) axis, also known as the "3-axis," from the inertial frame I, to the solar frame, F, by the angle $(t-\Omega t)$. The next set of rotations transforms the axes from the solar frame, F, to a frame oriented consistent with the sail axis, but not rotating with, the sailcraft, denoted the C frame. The coordinate frame is rotated through the clock angle, δ, about the $\hat{\mathbf{x}}'$ axis, then by the pitch angle, α, about the $\hat{\mathbf{y}}''$ axis, as defined in Fig. B.2. The associated



(a) F frame to D frame (b) D frame to C frame

Fig. B.2: Rotation from the solar frame, F, to the sailcraft frame, C, via an intermediate frame, D.

transformation equations are

$$\left\{ \begin{array}{c} \hat{\mathbf{x}}''' \\ \hat{\mathbf{y}}''' \\ \hat{\mathbf{z}}''' \end{array} \right\} = {}^{C}\mathbf{T}_2^D(-\alpha)\,{}^{D}\mathbf{T}_1^F(-\delta) \left\{ \begin{array}{c} \hat{\mathbf{x}}' \\ \hat{\mathbf{y}}' \\ \hat{\mathbf{z}}' \end{array} \right\} \tag{B.3}$$

$${}^{C}\mathbf{T}_2^D\,(-\alpha)\,{}^{D}\mathbf{T}_1^F\,(-\delta) = \left[\begin{array}{ccc} \cos\alpha & \sin\alpha\sin\delta & \cos\delta\sin\alpha \\ 0 & \cos\delta & -\sin\delta \\ -\sin\alpha & \cos\alpha\sin\delta & \cos\alpha\cos\delta \end{array} \right] \tag{B.4}$$

The matrix in Eq. (B.4) can be combined with the matrix in Eq. (B.2).

A final rotation is required to transform from the C frame to a body-fixed frame, S, via ψ about the sail-face normal, $\hat{\mathbf{u}}$, as indicated in Fig. B.3. Note that $\dot{\psi}$ is the relative rotation rate and not the spin rate [83]. If the spin rate is fixed, the relative rotation angle is a function of the spin rate and other angular terms. The associated transformation equations are

$$\left\{ \begin{array}{c} \hat{\mathbf{x}}^{iv} \\ \hat{\mathbf{y}}^{iv} \\ \hat{\mathbf{z}}^{iv} \end{array} \right\} = {}^{S}\mathbf{T}_1^C\,(\psi) \left\{ \begin{array}{c} \hat{\mathbf{x}}''' \\ \hat{\mathbf{y}}''' \\ \hat{\mathbf{z}}''' \end{array} \right\} \tag{B.5}$$

$${}^{S}\mathbf{T}_1^C\,(\psi) = \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & \cos\psi & \sin\psi \\ 0 & -\sin\psi & \cos\psi \end{array} \right] \tag{B.6}$$

A full rotation from the inertial frame to the body frame is then

$${}^{S}\mathbf{T}^I = {}^{S}\mathbf{T}_1^C(\psi)\,{}^{C}\mathbf{T}_2^D(-\alpha)\,{}^{D}\mathbf{T}_1^F(-\delta)\,{}^{F}\mathbf{T}_3^R(-\Omega t)\,{}^{R}\mathbf{T}_3^I(t) \tag{B.7}$$

The above rotations are employed to formulate the angular velocity vector.

An angular velocity vector and an angular acceleration vector relating the motion in the body frame to motion in the inertial frame are recovered from the transforma-

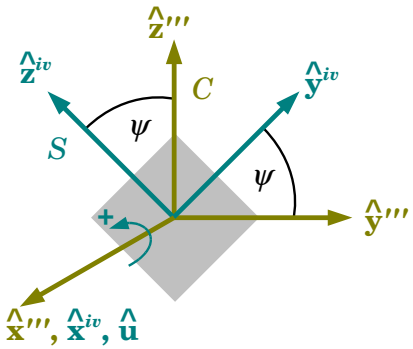


Fig. B.3: Rotations from the sailcraft frame, C, to the body-fixed frame, S.

tion in Eq. (B.7). The angular velocity vector, ${}^{I}\boldsymbol{\omega}^{S}$, is constructed from the angular velocities of each rotation, that is,

$$ {}^{I}\boldsymbol{\omega}^{S} = {}^{I}\boldsymbol{\omega}^{F} + {}^{F}\boldsymbol{\omega}^{D} + {}^{D}\boldsymbol{\omega}^{C} + {}^{C}\boldsymbol{\omega}^{S} \tag{B.8} $$

$$ = (1-\Omega)\hat{\mathbf{z}}' - \dot{\delta}\hat{\mathbf{x}}'' - \dot{\alpha}\hat{\mathbf{y}}''' + \dot{\psi}\hat{\mathbf{x}}^{iv} \tag{B.9} $$

When expressed in body-fixed coordinates,

$$ {}^{I}\boldsymbol{\omega}^{S} = \left\{ \begin{array}{r} \dot{\psi} - \dot{\delta}\cos\alpha + (1-\Omega)\sin\alpha\cos\delta \\ -\dot{\alpha}\cos\psi + \dot{\delta}\sin\psi\sin\alpha + (1-\Omega)(\sin\psi\cos\alpha\cos\delta - \cos\psi\sin\delta) \\ \dot{\alpha}\sin\psi + \dot{\delta}\cos\psi\sin\alpha + (1-\Omega)(\cos\psi\cos\alpha\cos\delta + \sin\psi\sin\delta) \end{array} \right\} \tag{B.10} $$

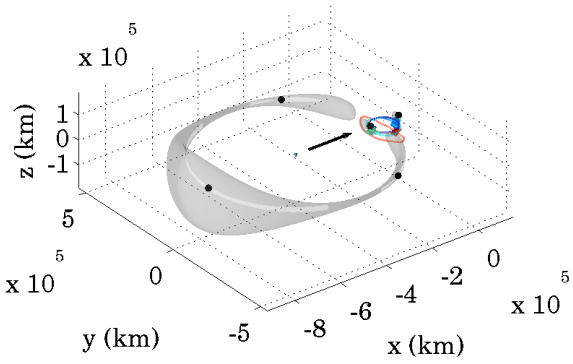
If the spacecraft possesses a fixed spin rate, whether it is three-axis stabilized ($\omega_{x0} = 0$) or spinning ($\omega_{x0} \neq 0$), the relative rotation angle, ψ, is integrated from

$$ \dot{\psi} = \omega_{x0} + \dot{\delta}\cos\alpha + (1-\Omega)\sin\alpha\cos\delta \tag{B.11} $$

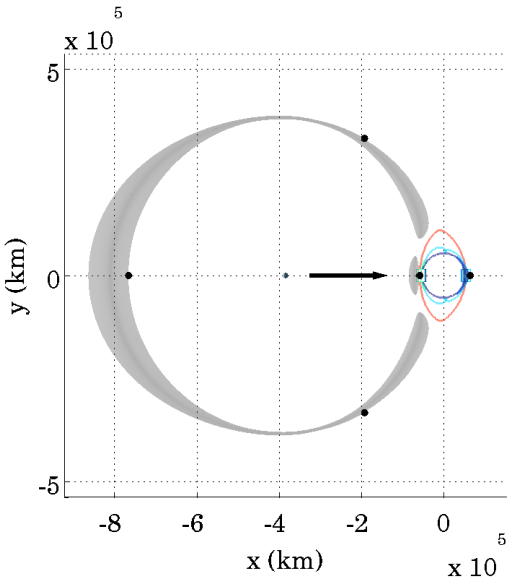
Finally, because the angular velocity vector is expressed in terms of the body-fixed frame, a derivative of Eq. (B.10) is required to determine the angular acceleration vector, ${}^{I}\dot{\boldsymbol{\omega}}^{S}$; a central-difference approximation of ${}^{I}\boldsymbol{\omega}^{S}$ is sufficient. Both the angular velocity and the angular acceleration are required to calculate the specific transverse torque, $\mathscr{M}_t$, that is required to physically reorient the sailcraft.

C. Instantaneous Equilibrium Surfaces Associated with $a_c = 1.70$ mm/s^2 in the Earth–Moon System

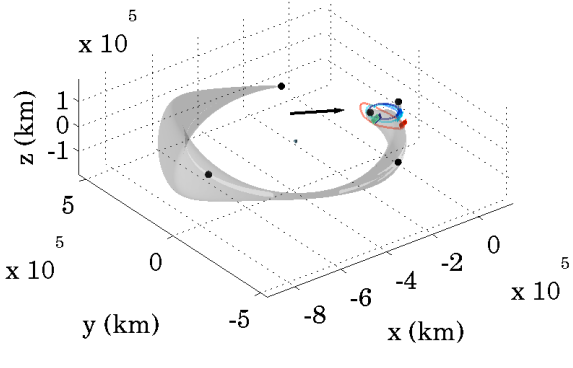
This Appendix includes additional views of the instantaneous equilibrium surfaces from Section 5.3. As the Sun moves around the Earth–Moon system, the torus changes shape, separates, and rejoins. The surfaces associated with a sail possessing a characteristic acceleration of $a_c = 1.70$ mm/s^2 appear in Figs. C.1 through C.6. The location and orientation of a sailcraft along each of the five sample trajectories from Section 5.2 appears as a colored plane along each trajectory at the time associated with the specified direction of the Sun-line. In the figures, the sunlight direction, or Sun-line, is represented by the large black arrow, the Moon is centered at the origin of the coordinate system, and the location and orientation of the sailcraft along each trajectory at the time associated with the indicated direction of the Sun-line appears as a colored plane. The Earth and the Moon, as well as the five Lagrange points in the Earth–Moon CR3B system, are included for scale.



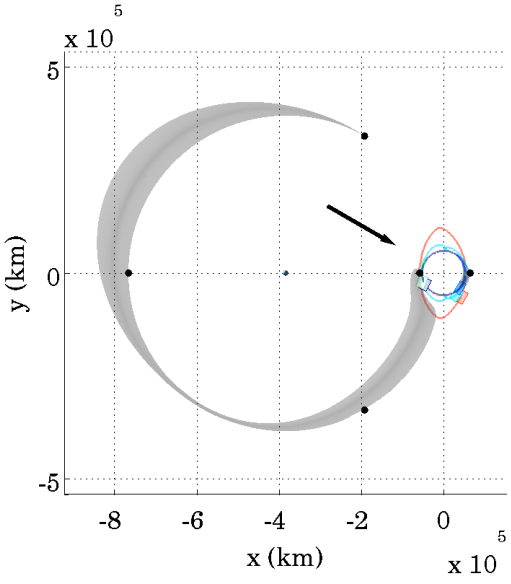
(a) $\Omega t = 0°$, 3-D



(b) $\Omega t = 0°$, xy view

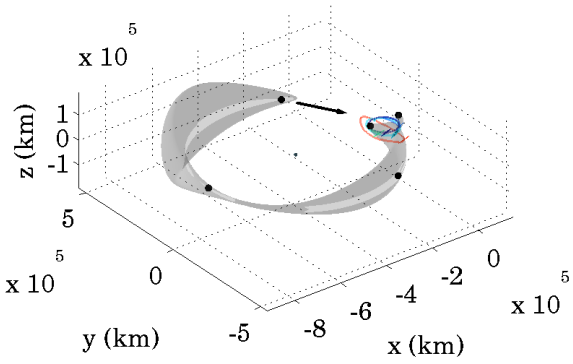


(c) $\Omega t = 30°$, 3-D

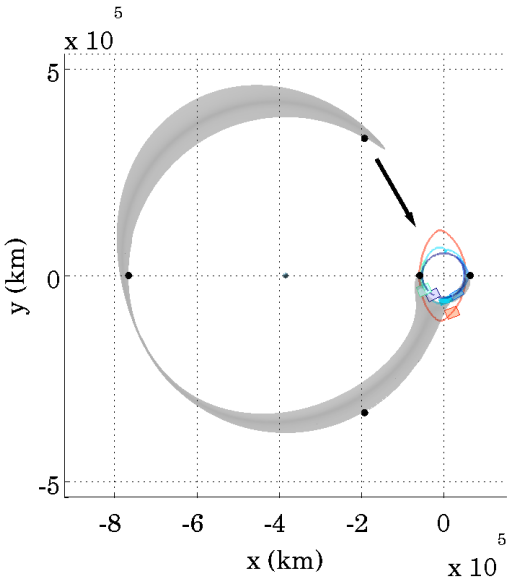


(d) $\Omega t = 30°$, xy view

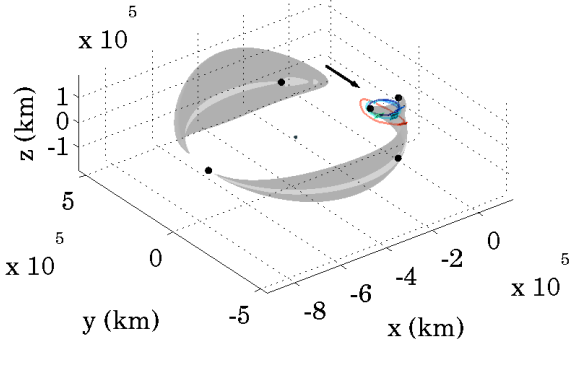
Fig. C.1: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 0° and 30°.



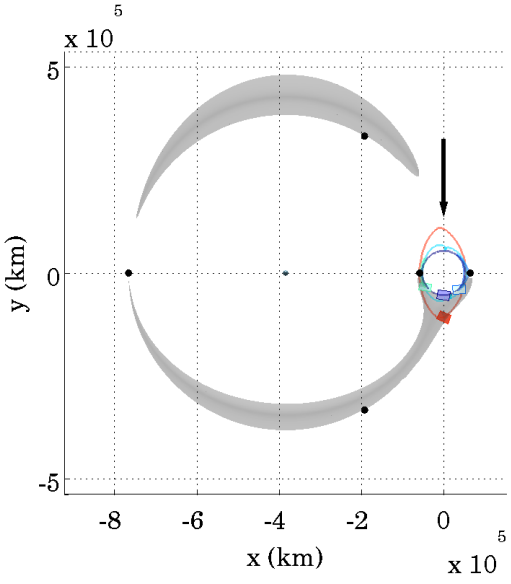
(a) $\Omega t = 60°$, 3-D



(b) $\Omega t = 60°$, xy view

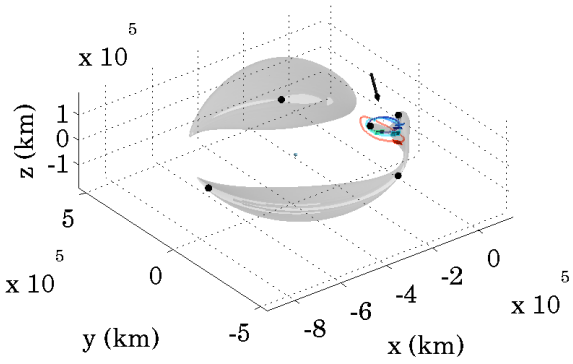


(c) $\Omega t = 90°$, 3-D

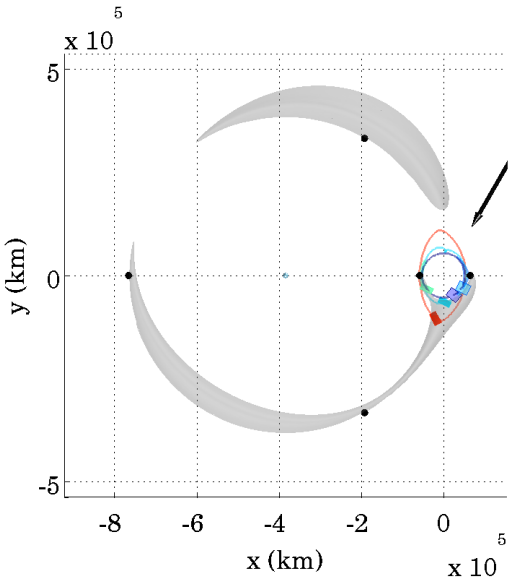


(d) $\Omega t = 90°$, xy view

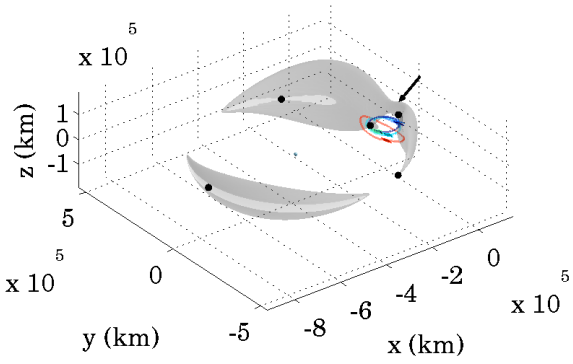
Fig. C.2: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 60° and 90°.



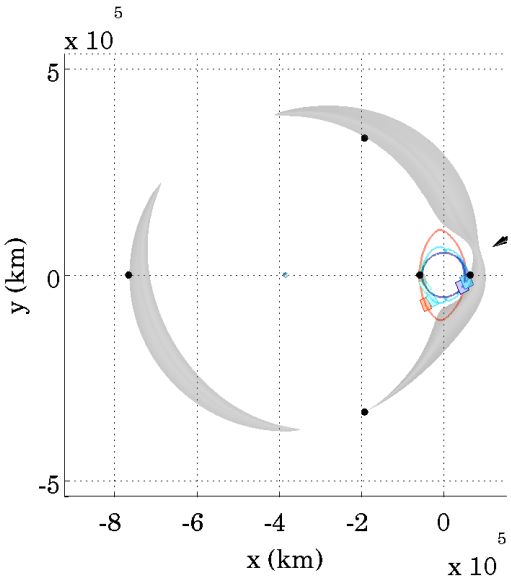
(a) $\Omega t = 120°$, 3-D



(b) $\Omega t = 120°$, xy view

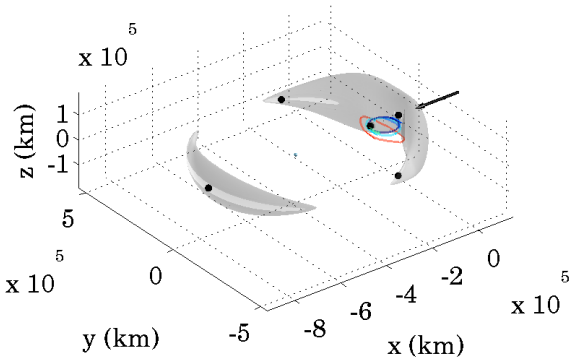


(c) $\Omega t = 150°$, 3-D

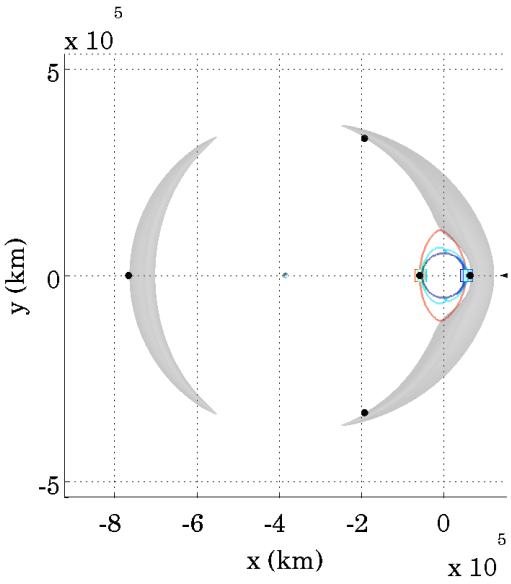


(d) $\Omega t = 150°$, xy view

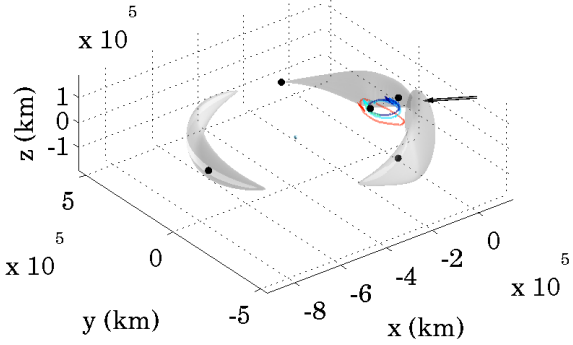
Fig. C.3: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 120° and 150°.



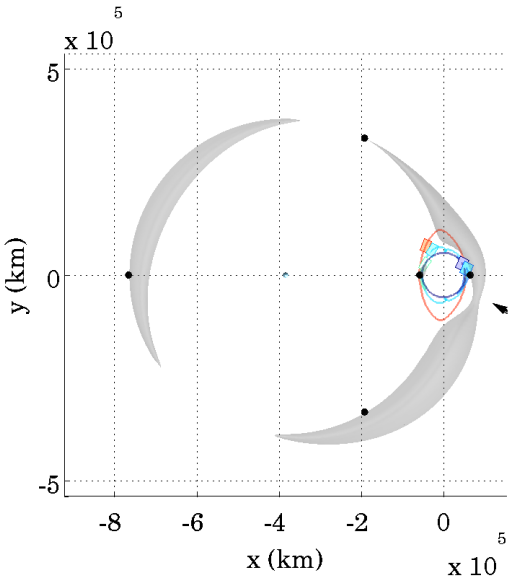
(a) $\Omega t = 180°$, 3-D



(b) $\Omega t = 180°$, xy view

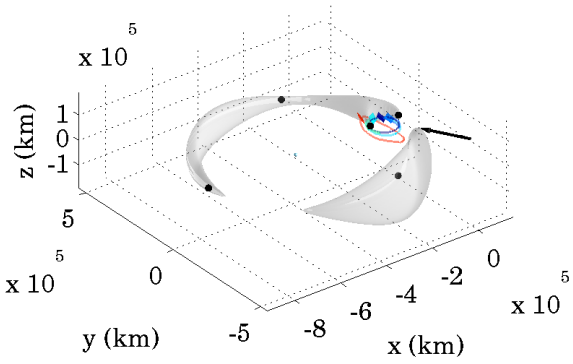


(c) $\Omega t = 210°$, 3-D

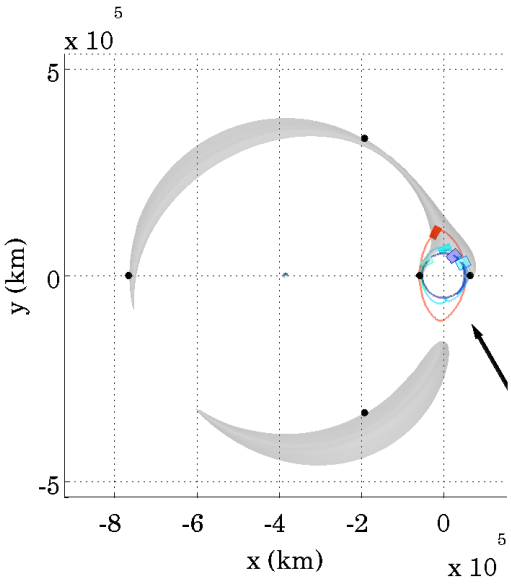


(d) $\Omega t = 210°$, xy view

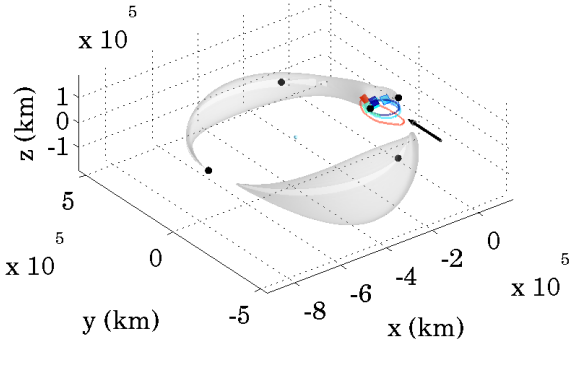
Fig. C.4: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 180° and 210°.



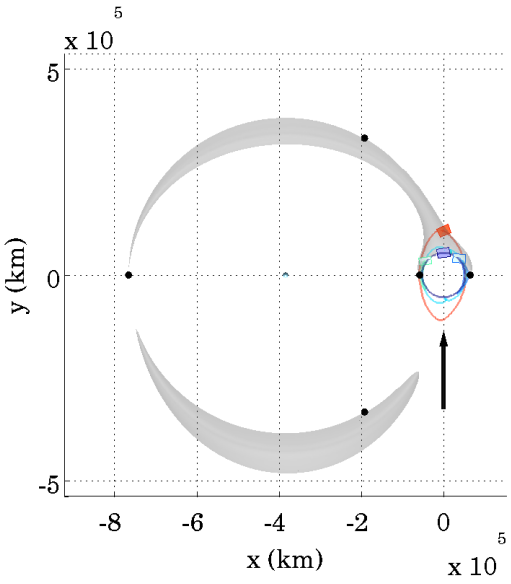
(a) $\Omega t = 240°$, 3-D



(b) $\Omega t = 240°$, xy view

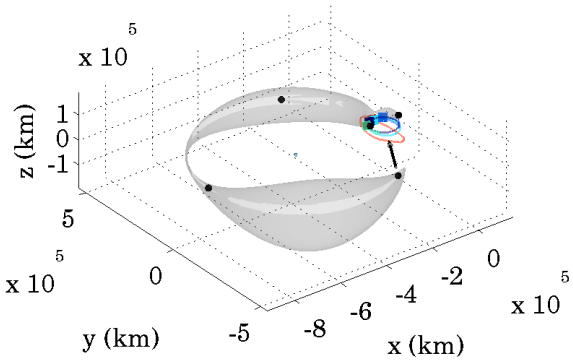


(c) $\Omega t = 270°$, 3-D

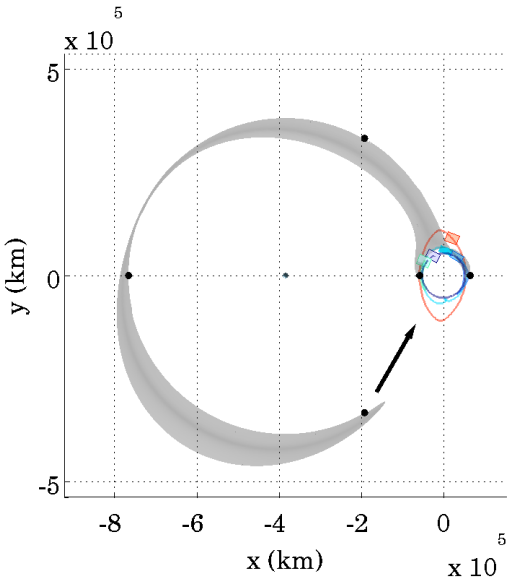


(d) $\Omega t = 270°$, xy view

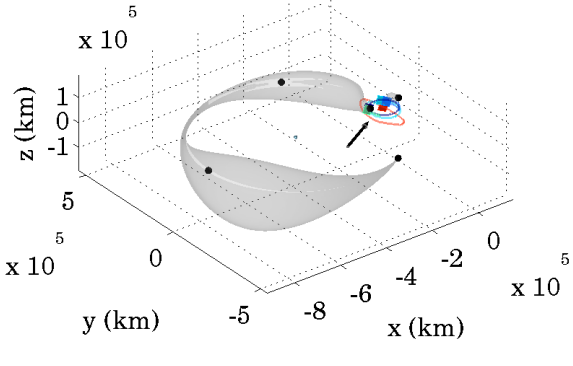
Fig. C.5: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 240° and 270°.



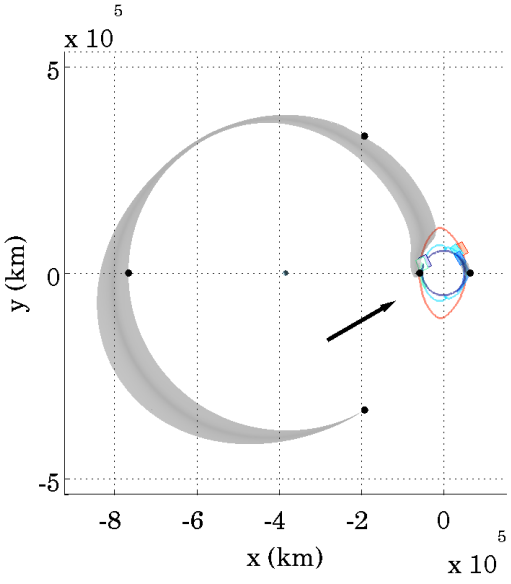
(a) $\Omega t = 300°$, 3-D



(b) $\Omega t = 300°$, xy view



(c) $\Omega t = 330°$, 3-D



(d) $\Omega t = 330°$, xy view

Fig. C.6: Instantaneous equilibrium surfaces in the Earth–Moon system for a sail with $a_c = 1.70$ mm/s^2 corresponding to solar angles of 300° and 330°.

D. Supplemental results for trajectory control strategies

This Appendix includes additional results from the control strategies that are initially presented in Section 7.5. The figures and the tables representing the results from simulations involving the control schemes appear throughout this document. For convenience, the locations of those figures and tables are summarized in Table D.1. The table includes results from simulations based on applications of the variations of the multiple-shooting flight-path control scheme. The full-arc strategy fails to track any reference orbits for more than a few revolutions in the presence of errors; thus, the full-arc results are limited.

Table D.1: Locations of results for various control schemes

	Error scheme and flight-simulation turn modeling		
	No errors, Instantaneous	10 km, 10 cm/s, 1° (3σ), 10°/hr slew rate	10 km, 10 cm/s, 3° (3σ), 10°/hr slew rate
Full-arc	Section 7.5.1	Not reported	Not reported
Look-ahead	Appendix D.1	Section 7.5.2	Appendix D.2
Adaptive	Appendix D.3	Section 7.5.3	Appendix D.4

D.1 Look-ahead control strategy with no errors

For comparison to other error configurations, the look-ahead scheme is examined in the presence of perfect knowledge of the state and no errors or uncertainty in the orientation as well as an assumption of instantaneous turns. Results for this configuration appear in Figs. D.1 through D.5. Sample results are highlighted in

Table D.2 (note that the ordinate refers to the number of days between turns and abscissa refers to the number of turns used to design an attitude profile).

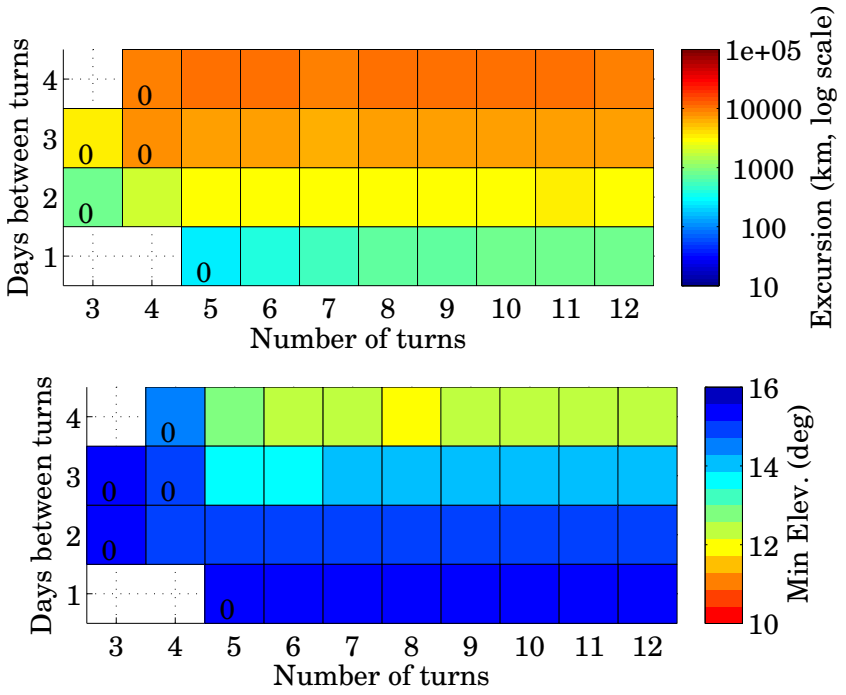


Fig. D.1: Maximum excursions and minimum elevation angles for the aqua reference trajectory based on an error-free configuration.

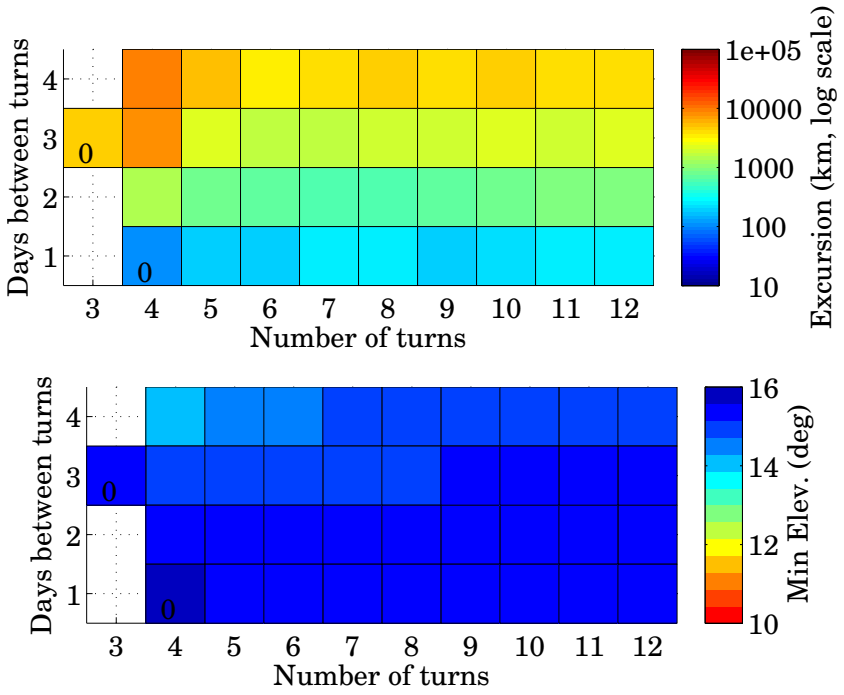


Fig. D.2: Maximum excursions and minimum elevation angles for the royal-blue reference trajectory based on an error-free configuration.

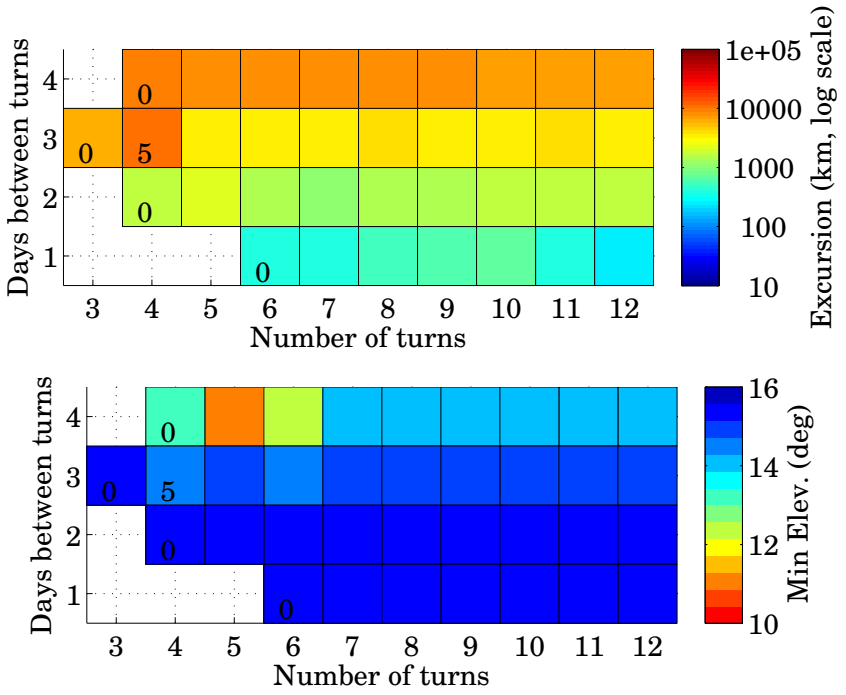


Fig. D.3: Maximum excursions and minimum elevation angles for the cyan reference trajectory based on an error-free configuration.

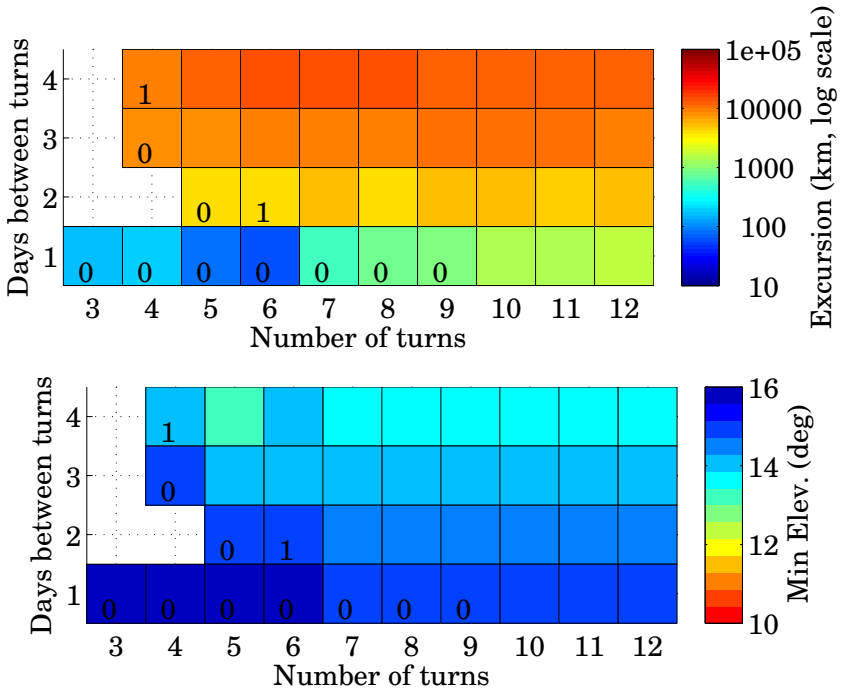


Fig. D.4: Maximum excursions and minimum elevation angles for the red-orange reference trajectory based on an error-free configuration.

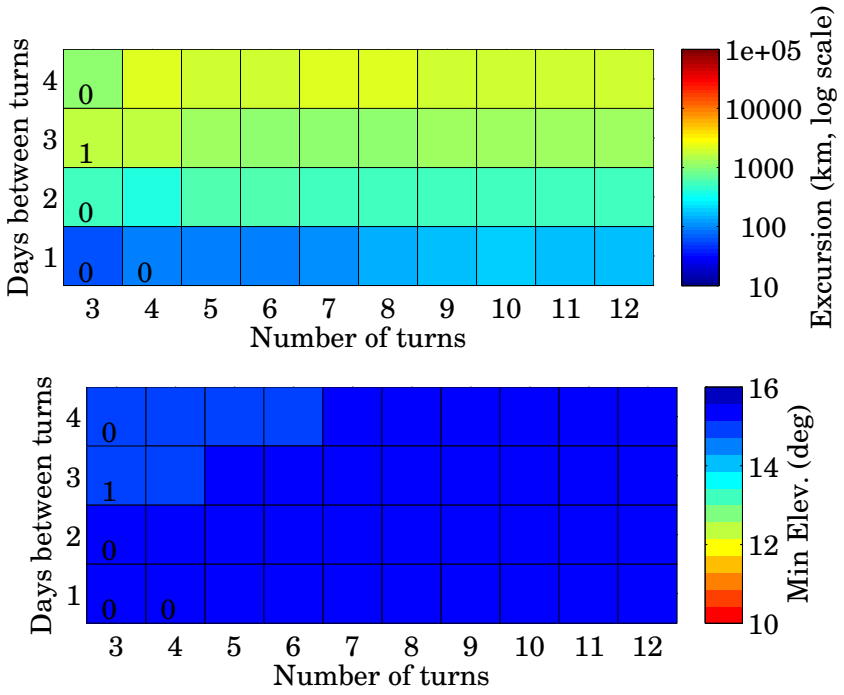


Fig. D.5: Maximum excursions and minimum elevation angles for the dark-blue reference trajectory based on an error-free configuration.

Table D.2: Look-ahead targeting: Excursions from the reference paths and elevation angles along controlled paths, based on an error-free configuration

Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
Extreme values					
Max Excur. (km)	11051	9700	10134	14953	2360
g-by-h	4-by-6	4-by-4	3-by-4	4-by-6	4-by-4
No. of revs.	36	36	5	36	36
Min Elev. Angle	12.1°	14.1°	10.9°	13.3°	14.9°
g-by-h	4-by-8	4-by-4	4-by-5	4-by-5	3-by-3
No. of revs.	36	36	36	36	1
2 days between turns, 7 turns per arc in designed sequence					
Max Excur. (km)	2781	617	1117	5405	555
Min Elev. Angle	14.9°	15.3°	15.2°	14.6°	15.3°

D.2 Look-ahead control strategy with large random attitude errors

To illustrate the sensitivity of the solar sail trajectories to large random attitude errors, attitude control errors of 3° (3σ) are incorporated into the look-ahead scheme. Results for the look-ahead scheme with random errors in the state knowledge (10 km and 10 cm/sec, 3σ) and attitude control (3°, 3σ), as well as 10°/hr turns incorporated into the flight simulation, appear in Figs. D.6 through D.10. Turns are assumed to be instantaneous when developing a turn profile. Sample results are highlighted in Table D.3 (note that the ordinate refers to the number of days between turns and the abscissa refers to the number of turns used to design an attitude profile). The criteria for selecting the "best options" in Table D.3 the greatest number of days between turns, followed by trajectories that possess minimum elevation angles greater than 13°, and the smallest excursion from the reference path. Most combinations of

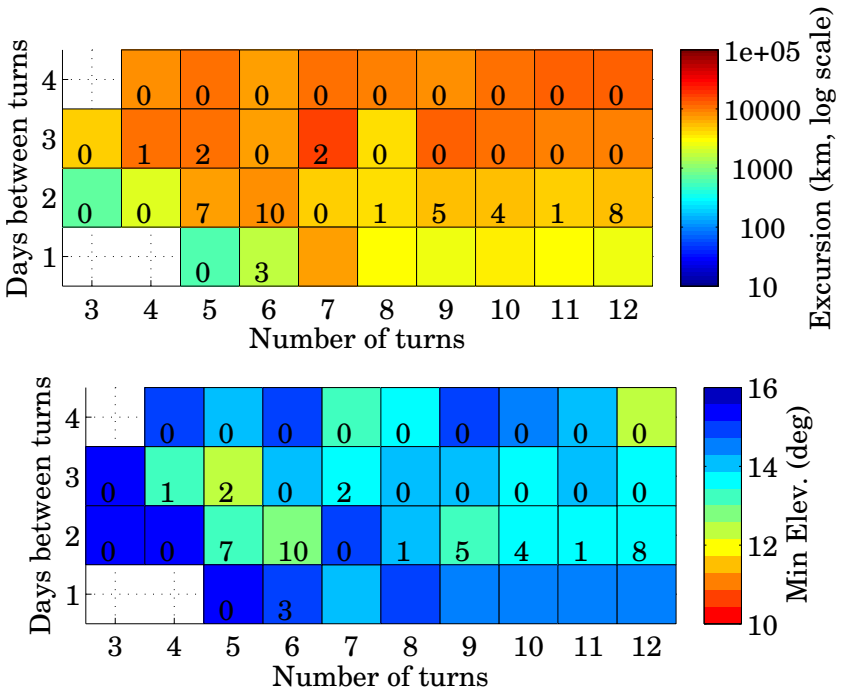


Fig. D.6: Maximum excursions and minimum elevation angles for the aqua reference trajectory based on a large attitude-error configuration.

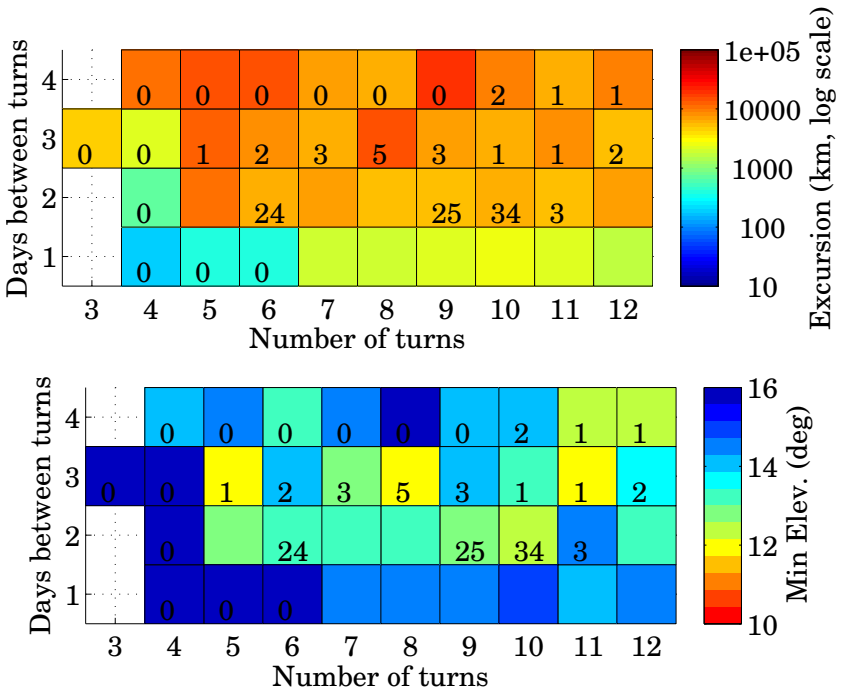


Fig. D.7: Maximum excursions and minimum elevation angles for the royal-blue reference trajectory based on a large attitude-error configuration.

days between turns (i.e., subarc length) and number of turns employed to design the reorientation sequence that enable the look-ahead algorithm to track the reference path result in controlled paths that appear similar to the respective reference path. The 4-day, 8-turn combination employed to track the red-orange reference path results in a controlled path that possesses the greatest excursions from the red-orange path (Fig. D.9). Nevertheless, the algorithm converges on a solution for the full 36-month simulation. The resulting path appears in Fig. D.11.

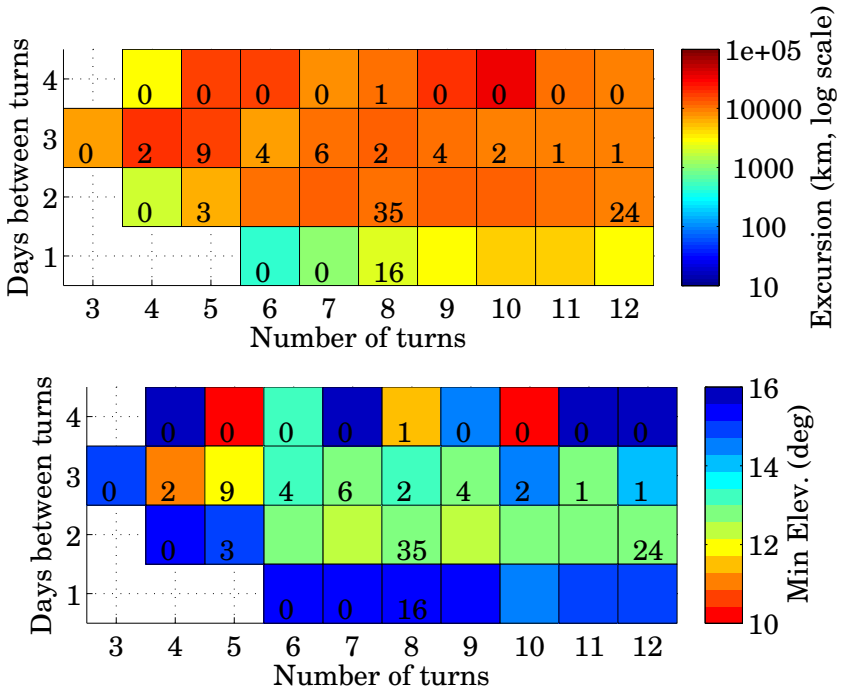


Fig. D.8: Maximum excursions and minimum elevation angles for the cyan reference trajectory based on a large attitude-error configuration.

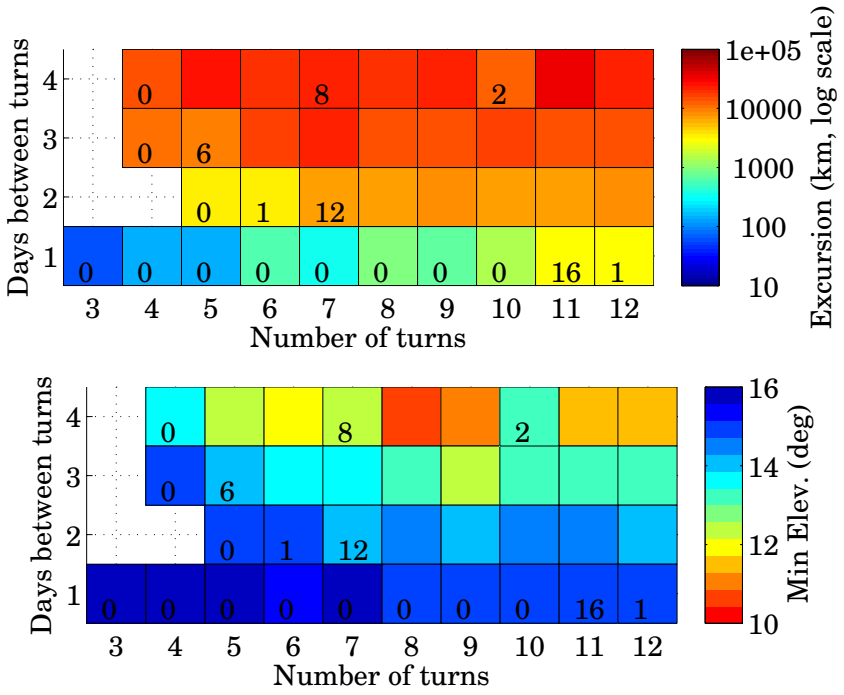


Fig. D.9: Maximum excursions and minimum elevation angles for the red-orange reference trajectory based on a large attitude-error configuration.

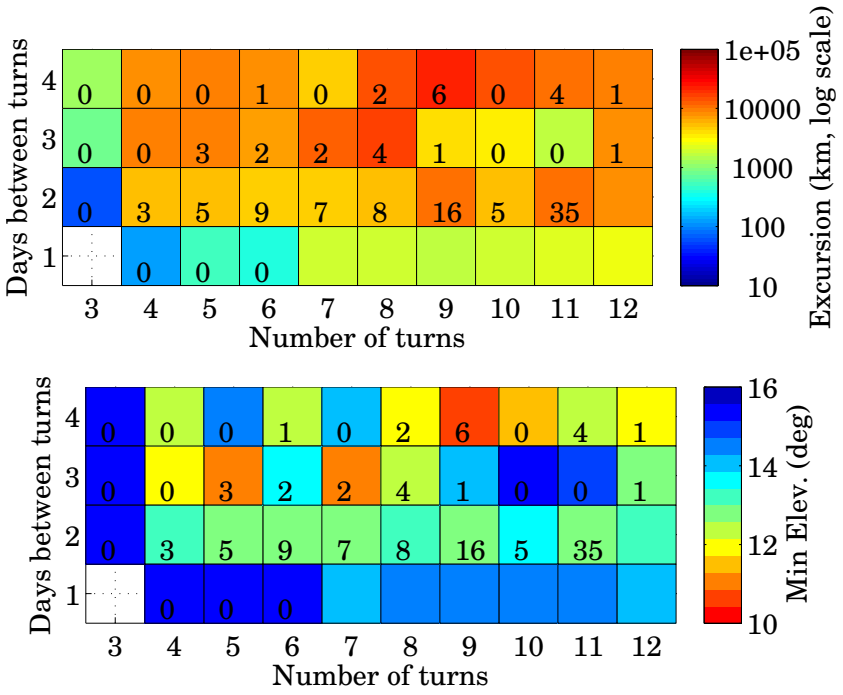
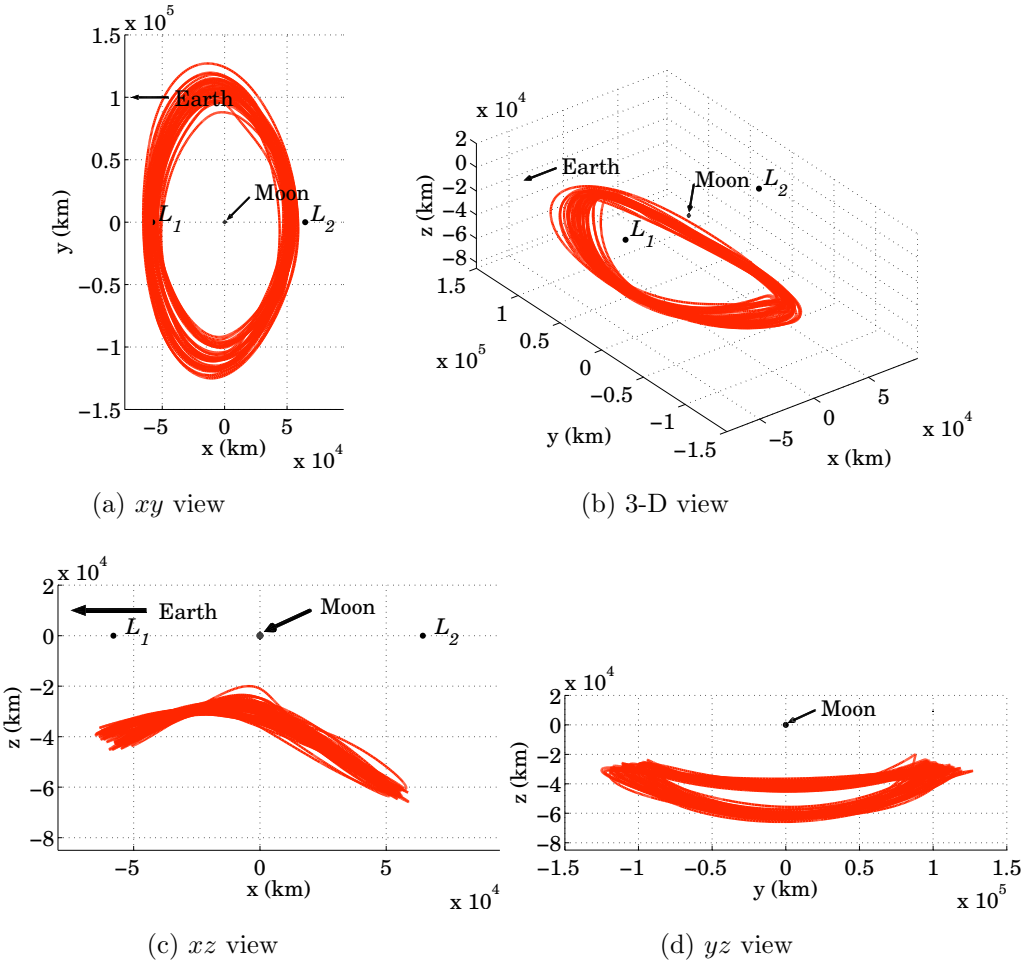


Fig. D.10: Maximum excursions and minimum elevation angles for the dark-blue reference trajectory based on a large attitude-error configuration.

Table D.3: Look-ahead targeting: Excursions from the reference paths and elevation angles along controlled paths, based on a configuration that incorporates 3° errors in the control

Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
Best options					
Max Excur. (km)	6759	5040	10298	16099	7776
Min Elev. Angle	14.1°	13°	13°	13.5°	13.1°
g-by-h	1-by-7	2-by-8	2-by-6	3-by-6	3-by-12
No. of revs.	36	36	36	36	36
Extreme values					
Max Excur. (km)	16383	18383	34477	32322	21501
g-by-h	3-by-7	4-by-9	4-by-10	4-by-11	4-by-9
No. of revs.	2	0	0	36	6
Min Elev. Angle	12.2°	11.9°	8.95°	10.8°	10.6°
g-by-h	3-by-5	3-by-5	4-by-10	4-by-8	4-by-9
No. of revs.	2	1	0	36	6



(a) xy view

(b) 3-D view

(c) xz view

(d) yz view

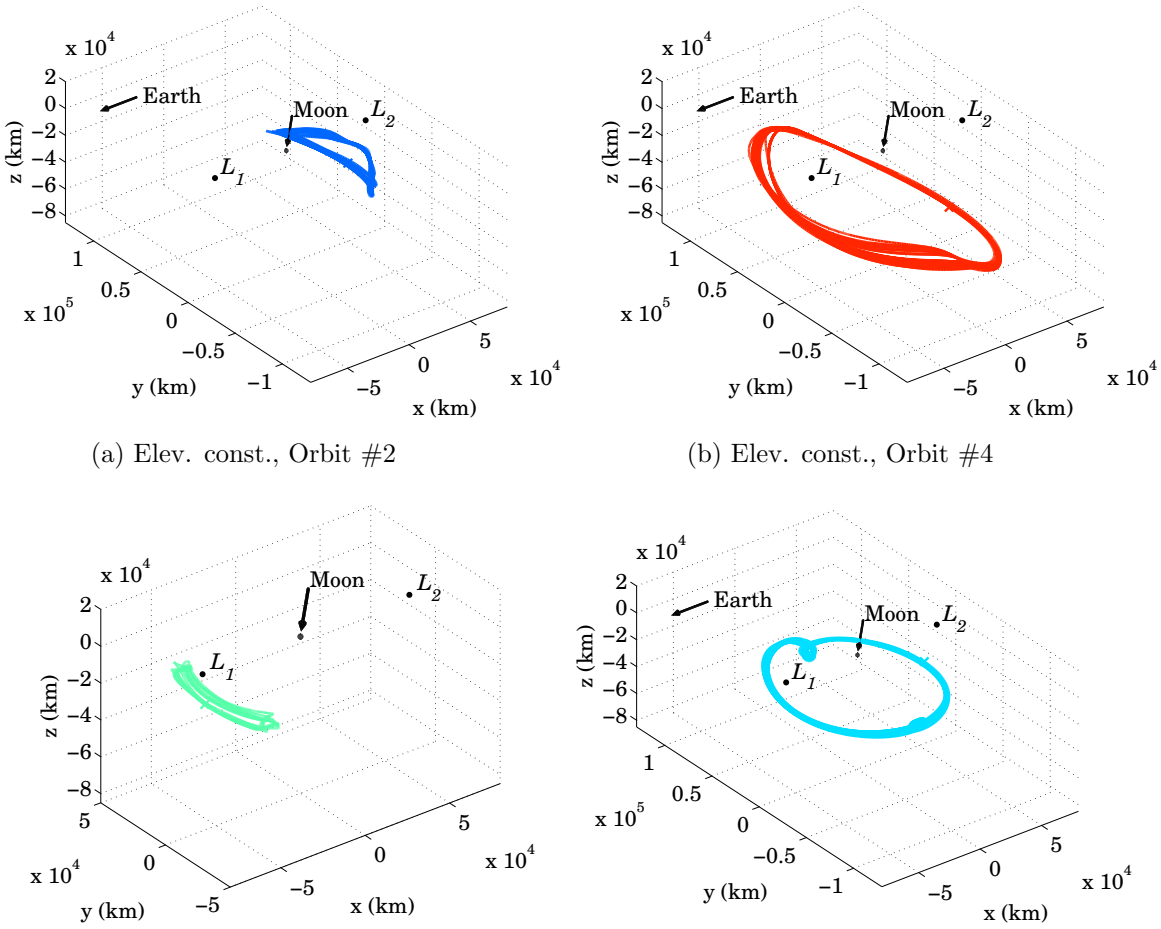
Fig. D.11: The resulting 36-month trajectory from the 4-day, 8-turn combination of the large-error look-ahead scheme for the red-orange reference path.

D.3 Adaptive controller with no errors

The adaptive control scheme is applied to each of the sample reference paths assuming zero error with either a 13° elevation-angle constraint or a 5000 km excursion constraint. Essentially, this configuration evaluates a series of fixed turns again the continuous solution. Initially using, 4 days between turns, the adaptive scheme attempts longer arc lengths (i.e., more turns) until reaching the limiting number of turns, h, (e.g., 12 turns). Then, a reduction of the subarc length, g, to 3 days between turns is attempted. With this configuration, the adaptive scheme does not successfully track all of the reference paths, as is clear in Table D.4. Sample trajectories from Orbits #2 and #4 subject to the elevation-angle constraint, as well as Orbits #1 and #3 subject to the excursion constraint, appear in Fig. D.12.

Table D.4: Adaptive targeting: 4-day window, no knowledge, control, or turn modeling errors

Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
13° elevation-angle constraint					
Max Excur. (km)	19532	15597	35816	22462	2522
Min Elev. Angle	12.6°	12.5°	12.9°	13.3°	14.7°
No. of revs.	36	36	5	36	36
Ave. g (days)	4	4	4	4	4
Ave. h (turns)	3.54	3.25	3.54	3.54	3.07
5000 km excursion constraint					
Max Excur. (km)	5374	6126	5913	7018	2736
Min Elev. Angle	12.3°	13.3°	13.9°	13.8°	14.7°
No. of revs.	8	36	36	19	36
Ave. g (days)	3.13	4	4	3.71	4
Ave. h (turns)	3.53	3.56	4.17	4.15	3.08



(a) Elev. const., Orbit #2

(b) Elev. const., Orbit #4

(c) Excur. const., Orbit #1

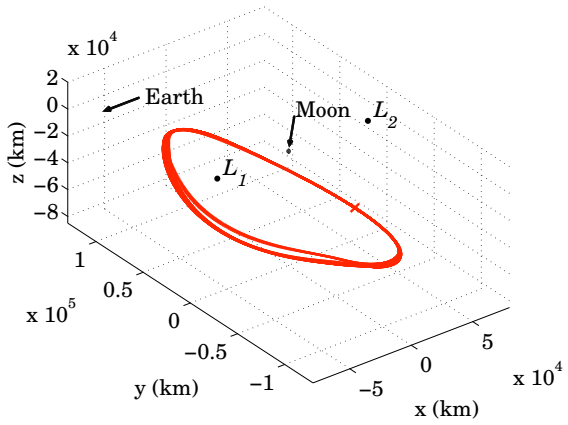
(d) Excur. const., Orbit #3

Fig. D.12: Four successful trajectories from the adaptive scheme, initiated with 4 days between turns and not subject to errors.

The long length of time between turns (and the imposition of path constraints at the time of those turns) contributes to the divergence from the controlled path. Employing the same error configuration as before, the adaptive scheme is initialized with two days between turns. All reference paths are tracked for the full 36 months, as demonstrated in Table D.5. With the exception of Orbit #4 (subject to an elevation-angle constraint), the controlled paths are visually identical to their respective reference trajectories. This exception appears in Fig. D.13.

Table D.5: Adaptive targeting: 2-day window, no knowledge, control or turn modeling errors

Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
13° elevation-angle constraint					
Max Excur. (km)	3550	3561	2972	8830	748
Min Elev. Angle	15°	15.2°	15.4°	14.6°	15.2°
No. of revs.	36	36	36	36	36
Ave. g (days)	2	2	2	2	2
Ave. h (turns)	3.31	3.26	3.28	3.67	3.14
5000 km excursion constraint					
Max Excur. (km)	3552	3552	2972	4408	749
Min Elev. Angle	15°	15.2°	15.4°	14.6°	15.2°
No. of revs.	36	36	36	36	36
Ave. g (days)	2	2	2	2	2
Ave. h (turns)	3.31	3.26	3.28	3.68	3.14



(a) Elev. const., Orbit #4

Fig. D.13: Sample successful trajectory from the adaptive scheme, initiated with 2 days between turns and not subject to errors.

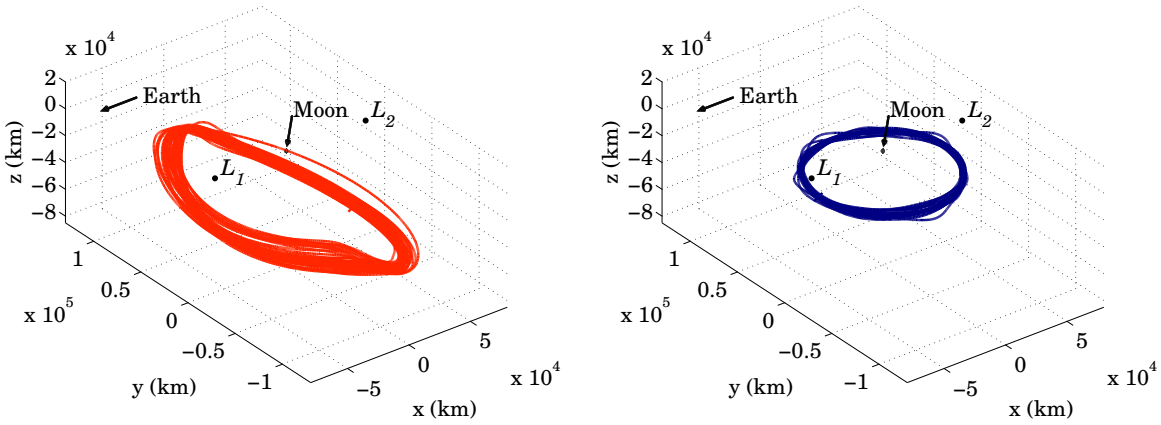
D.4 Adaptive controller with large control errors

Solar sail trajectories are sensitive to errors in attitude control. To test the effects of these errors against the adaptive tracking control scheme, random state-trajectory errors (10 km, 10 cm/sec, 3σ) and attitude errors ($3°$, 3σ), as well as $10°$/hr turn rates, are introduced in the flight simulation. In all but one case—the red-orange orbit—the adaptive controller fails to track the reference paths for the full 36 orbits. The results appear in Table D.6. Orbits #4 and #5, subject to elevation-angle constraints, appear in Fig. D.14.

Table D.6: Adaptive targeting: 4-day window, $3°$ control error (3σ)

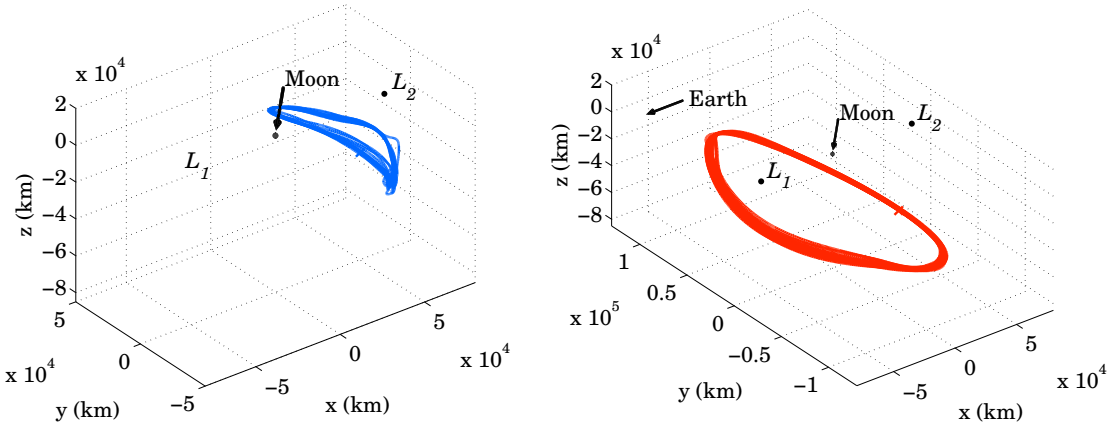
Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
$13°$ elevation-angle constraint					
Max Excur. (km)	8741	38662	2731	27708	13266
Min Elev. Angle	$14.7°$	$11°$	$19.3°$	$12.2°$	$11.2°$
No. of revs.	0	6	0	36	23
Ave. g (days)	4	3.96	4	3.99	3.96
Ave. h (turns)	4.5	3.91	4	3.67	3.37
5000 km excursion constraint					
Max Excur. (km)	5528	6036	2633	5297	8238
Min Elev. Angle	$14.7°$	$15.5°$	$19.3°$	$28°$	$12.8°$
No. of revs.	0	0	0	0	1
Ave. g (days)	3	4	4	4	4
Ave. h (turns)	3.33	3.67	4	4	3.27

When the initial time between turns is reduced to 2 days, the adaptive scheme tracks the reference paths better, but only two trajectories are able to be tracked for the full 36 months. The results appear in Table D.7 and Fig. D.15.



(a) Elev. const., Orbit #4 (b) Elev. const., Orbit #5

Fig. D.14: Two sample trajectories from the adaptive scheme, initiated with 2 days between turns and not subject to errors. Orbit #4 is tracked for the full 36 months, but Orbit #5 is only tracked for 23 months.



(a) Elev. const., Orbit #2 (b) Elev. const., Orbit #4

Fig. D.15: Two sample trajectories from the adaptive scheme, initiated with 2 days between turns and not subject to errors. Orbit #4 is tracked for the full 36 months, but Orbit #2 is only tracked for 13 months.

Table D.7: Adaptive targeting: 2-day window, 3° control error (3σ)

Orbit Number	1	2	3	4	5
Color	aqua	royal-blue	cyan	red-orange	dark-blue
13° elevation-angle constraint					
Max Excur. (km)	7758	11928	17778	9990	4915
Min Elev. Angle	12.9°	13.2°	13.5°	14.2°	12.5°
No. of revs.	14	13	19	36	36
Ave. g (days)	2	2	2	2	2
Ave. h (turns)	3.52	3.6	3.64	4.42	3.4
5000 km excursion constraint					
Max Excur. (km)	5677	3370	4946	4322	4561
Min Elev. Angle	13.3°	17.6°	14.2°	14.5°	12.7°
No. of revs.	18	0	2	4	9
Ave. g (days)	2	2	2	2	2
Ave. h (turns)	3.51	3.5	3.97	4.16	3.42

VITA

VITA

Geoffrey Wawrzyniak graduated from the University of Wisconsin–Madison with a Bachelors of Science in Mechanical Engineering in 1999 and his Master of Science in Engineering from the Aerospace Engineering and Engineering Mechanics Department at The University of Texas at Austin in 2001. Prior to attending Purdue, he was a navigator for various interplanetary missions at the Jet Propulsion Laboratory in Pasadena, California.