Politecnico di Torino
Master's Degree in Aerospace Engineering

Master's Thesis

Optimal low Earth orbit transfers with drag sails

Supervisor: Prof. Lorenzo Casalino
Co-Supervisor: Prof. Francesco Topputo

Author:
Andrea Forestieri

April 2021

A Pippo,
ovunque tu sia ora, spero che sia ruscito a trovare la serenità che tanto meriti,
amico mio.

Acknowledgments

Ringrazio di cuore il Prof. Casalino e il Prof. Topputo, per avermi dato l'opportunità di sviluppare un argomento di mio grande interesse.

Un sentito ringraziamento anche al Dott. Fasano, per i suoi preziosi consigli e per tutte le opportunità che mi ha dato.

Ai miei genitori: grazie per aver sempre avuto fiducia nelle mie capacità, per aver sempre investito su di me, per avermi lasciato libero di seguire le mie ambizioni, per aver sempre fatto di tutto affinché io non perdessi nessuna opportunità che mi si presentava davanti e, prima di tutto, per essere sempre stati dei modelli da seguire.

A mio fratello Giacomo: per te niente frasi strappalacrime. In fondo, io e te siam tipi coincisi. Sappi solo che, nonostante tutte le azzuffate, ti ringrazio per essere un fratello generoso ed è bello sapere di avere le proprie spalle coperte.

Desidero inoltre ringraziare Mattia e Vittoria, con cui ho condiviso tanto in questo percorso ed ancor prima di esso. Grazie per esserci sempre stati e per avermi aiutato quando ne avevo più bisogno. Avervi come amici è un privilegio.

Ringrazio inoltre tutta la compagnia de "gli Olbicelli". Un ringraziamento sentito a Fra, per essere stato un fratello maggiore e l'artefice di tutte le avventure più pazze che ho vissuto in questi anni. Lalla e Loris, da quando vi ho conosciuto ci avete messo poco a diventare delle persone importanti. Grazie per essermi stati accanto, soprattutto nell'ultimo difficile periodo. A Manu, per tutte le discussioni filosofiche davanti alle tante, forse troppe, birre del MacGilly's. Ad Andre e Vale, per la vostra gentilezza e per avere portato un po' di vivacità nel freddo di Torino.

Come non ringraziare anche tutti i miei colleghi di corso per tutte le giornate passate tra i corridoi del Poli, per i pranzi insieme, per le chiacchiere scambiate in fila alle macchinette del caffè, per le ore passate a ripetere e per quelle passate a far festa. Senza di voi non sarebbe stata la stessa esperienza.

E grazie a te, Chiara, che riesci sempre a tirare fuori il meglio di me. Senza di te, la redazione di questa tesi sarebbe stata immensamente più difficoltosa. Grazie per essere la mia voce quando non sono abbastanza forte per parlare e per rendermi ogni giorno un uomo migliore.

Grazie.

Abstract

Satellite operations in low Earth orbit will be extremely frequent in the near future, and consequently the optimization of trajectories between such orbits is becoming of extreme interest. Low-thrust electric propulsion will arguably be the preferred option for many missions, as it provides benefits in terms of propellant consumption. Low-thrust trajectories in low Earth orbit require several revolutions around the Earth and they are carried out in an environment where the two-body problem approximation is not suited to describe the motion of a satellite. The objective of this thesis is to illustrate a general methodology to take into account the effects of the oblateness of the Earth and aerodynamic drag, and obtain useful solutions to the minimum-time and minimum-propellant problems. An indirect optimization method based on Edelbaum's approximation is applied to transfers between almost circular low Earth orbits, considering the effects of drag and the asphericity of the Earth. The exploitation of drag sails is investigated in a case study, considering a small 15-kg spacecraft in an initial orbit similar to that of the International Space Station. Depending on the maximum frontal area of the sail, minor to significant improvements can be obtained in terms of minimum-time. The propellant consumption is improved significantly for a negative altitude change and in some cases nearly-zero-consumption maneuvers are achieved by deploying the sail at the right time.

Contents

List of Figures

List of Tables

Introduction

Active debris removal (ADR), on-orbit servicing (OOS) and small-sat deployment from the International Space Station (ISS) are just a few examples of missions the require a spacecraft to maneuver in low Earth orbit (LEO). Therefore, the optimization of trajectories between LEOs is becoming of extreme interest. Low-thrust electric propulsion will arguably be the preferred option for such missions, as it provides benefits in terms of propellant consumption. However, low-thrust trajectories require several revolutions around the Earth and they are carried out in an environment where the two-body problem approximation is not suited to describe the motion of satellites. As a result, the presence of perturbations must be accounted for in order to obtain significant solutions. The objective of this thesis is to illustrate a general method to take into account the effects of the oblateness of the Earth and aerodynamic drag for a quick and accurate estimation of minimum-time and minimum-propellant trajectories.

Given an orbital transfer, the optimization of a space trajectory consists in the determination of a control law (such as the time history of the thrust magnitude and direction) that either maximizes or minimizes a certain performance index. This work will be focused on the search of optimal control laws that either minimize the transfer time or maximize the final mass of a satellite or, equivalently, minimize propellant consumption. The optimization problem will be addressed with an optimal control theory (OCT) approach, that is a mathematical optimization method which aims to determine the control law that satisfies well determined physical constraints while maximizing (or minimizing) a performance index. OCT, explained in full detail in [3, 15, 22], applies the principles of the calculus of variations [16, 18] to obtain a set of optimality conditions, that result in boundary-value problem, the solution of which yields the optimal control law. This *indirect optimization* method has the advantage of having a high numerical precision, although the boundary value problem is somewhat difficult to solve [5].

LEOs have an altitude below 2000 km, so the eccentricity values are often less than 0.1. Therefore, the complexity of the optimization problem can be reduced by considering such orbits as circular. This approximation was used by Edelbaum for the most famous solution [8] to the minimum-time problem, published in 1961. His approach was based on an averaged dynamical model; he first determined the optimal controls for the one-revolution transfer and then used the obtained results to solve the multiple-revolutions transfer. Although the work of Edelbaum didn't include the precession of the line of nodes due to the oblatebess of the Earth, he laid the foundations for several modifications which added complexity to the original work. Other contributions to the analysis of optimal low-thrust orbit transfers between circular orbits appear in references [9, 11, 13, 27]. In [4] the Edelbaum's approach is enhanced by accounting for the variation of the mass of the S/C and in [7, 12] the precession of the line of nodes is accounted for by the Earth's gravitational harmonic J2 coefficient. This thesis further extends the method to deal with aerodynamic drag without being far more demanding in terms of computational effort.

The atmospheric density in LEO is quite difficult to be estimated. As a matter of fact, the upper atmosphere is subject to fluctuations that depend on a variaty of factors, such as altitude, local solar time and solar activity. One of the most precise models for the Earth's atmosphere is the Jacchia-Bowman 2008 model [2]. However, a less accurate exponential model based on the U.S. Standard Atmosphere 1976 [20] is used here for simplicity and much faster computation. Nonetheless, the proposed methodology is not dependent on the atmospheric model, so any model can be readily adopted instead. In addition, a constant value for the drag coefficient is adopted, although a more rigorous analysis could be carried out by considering its variation at different altitudes and at different solar times.

Since the orbits are almost-circular, changes of eccentricity and argument of periapsis are not considered. In addition, the rendezvouz problem is not addressed and the true anomaly time history is neglected as well. Therefore, the state of the system is described by semimajor axis, inclination and RAAN. The optimization problem is formulated by fixing such parameters at the initial time for a chaser spacecraft. The solution yields the control law to achieve rendezvous with a target spacecraft in either minimum time or with minimum propellant consumption. Both orbits are perturbed by J2, whereas aerodynamic drag doesn't affect the target orbit. Namely, it is assumed that the target spacecraft perfoms station-keeping maneuvers to maintain the same semimajor axis and inclination. In addition to thrust magnitude and direction, the frontal area of the chaser spacecraft is treated as a control variable. This assumption entails that the spacecraft is equipped with a drag sail and can perform aeroassisted maneuvers.

The thesis is structured as follows. Chapter 1 introduces the context of this work, Chapters 2 and 3 give a brief overview of astrodynamics and space propulsion. Chapter 4 describes the adopted mathematical model for the perturbations. Chapter 5 presents the basics of OCT and shows how to apply it to LEO transfers. A case study is then presented in Chapter 6 and the advantages brought by drag sails are evaluated.

Chapter 1

Research context

This work will apply an indirect optimization method to LEO trajectories. Thus, this chapter is meant to give the reader an insight into the reasons why the optimization of such trajectories is becoming of extreme interest among the scientific community.

1.1 Active debris removal

Sixty-four years ago Sputnik 1 was inserted into orbit. Since then, we've launched so many satellites for so many different applications that we can't even think of living without their endless benefits. Unfortunately, the bill always comes due. For all these years we've chosen to ignore the byproduct of space activities: orbital debris. The so called «space-junk» comes in different shapes and sizes. They can be as tiny as a fleck of paint or as big as a whole satellite. Collision-avoidance maneuvers are becoming routine to deal with the problem. However, this is not a permanent solution. The average impact speed of a piece of debris running into another object is roughly 10 km/s and even a tiny piece can cause a lot of damage. The most significant collision happened in 2009, when the dead Russian military satellite Kosmos-2251 accidentally collided with the active American commercial satellite Iridium 33. The collision had some serious consequences and scattered more than 1000 pieces of debris larger than 10 cm [26], let alone the smaller ones. The following year, the U.S. Space Surveillance Network (SSN) was able to catalog over 2000 debris fragments from the collision. The causes for this accident are to be found in the lack of precise and up-to-date information of current satellite positions and velocities. Although sometimes this information is available, it is often affected by errors. In this specific case, the two satellites were expected to miss by 584 meters. In addition, close approaches are becoming more frequent and planning an avoidance maneuver is a challenging task, also considering its effects on the satellite's normal functioning. Although the fallout of this event is under control, we can't ignore the fact it's a warning of the

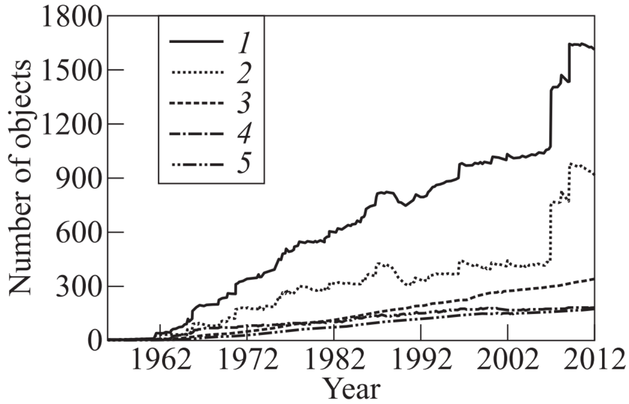


Figure 1.1: History of the increase of the debris population cataloged by the SSN: 1 - total objects; 2 - fragmentation debris; 3 - satellites; 4 - debris related to missions; and 5 - rocket bodies

potential collision cascade effect, known as the *Kessler Syndrome* [14]. This collision cascading was first predicted by Kessler and Cour-Palais in 1978. They formulated a theoretical scenario in which the density of orbital debris in LEO is so high that a satellite collision could cause a cascade in which each collision is the cause of further ones. By using a mathematical model they were able to predict the rate at which a belt of orbital debris might form. This event could have dangerous implications, as it could make the use of satellites in specific orbits difficult for many generations.

As shown in fig. 1.1, the historical increase of the debris population is mainly driven by fragmentation debris. The top curve is the increment of objects and further detail about the population breakdown is brought by the the four curves below. The well evident recent jumps coincide with the Chinese anti-satellite test of 2007 end the collision between Cosmos 2251 and Iridium 33. The majority of the 22000 objects the SSN was able to catalog in 2012 were larger than 10 cm. However, radar observations show that the number goes up to 500,000 if 1 cm level objects are counted. Moreover, at the 1 mm level the population is likely in the order of hundreds of millions. However, at the relative speed of impacts, even this tiny debris can pose serious concerns. As a matter of fact, an impact by a debris larger than 5 mm is likely to end the mission of a satellite. Fig. 1.2 shows the predicted growth of space debris, as calculated by the NASA Orbital Debris Program Office in a study that didn't assume any mitigation measure in the future. The geosynchronous (GEO) zone is located in a range of 400 km around the geosynchronous altitude, and the zone between LEO and GEO is the medium Earth orbit (MEO). As it can be readily seen, the increase in LEO is projected to be the most substantial. Indeed, the recent increase in LEO has lead to the introduction of the 25-year rule and the measure known as passivation. The first one states that satellites and orbital stages of rockets shall reenter the Earth's atmosphere within 25 years of mission completion if their

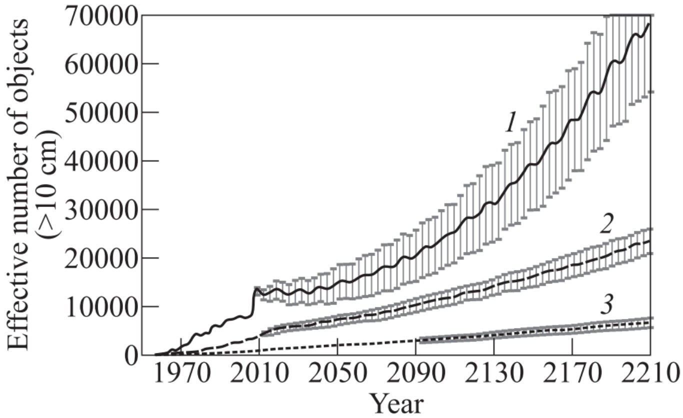


Figure 1.2: Growth projection of the debris population larger than 10 cm in LEO, MEO and GEO. The projection assumes no mitigation measures implemented in the future: 1 - LEO (between 200 km and 2000 km altitude); 2 - MEO (between 2000 km and 35586 km altitude); and GEO (between 35586 km and 35986 km altitude)

deployment orbit altitude is in the LEO region, whereas the latter consist in the depletion of all latent energy reservoirs of a satellite or orbital stage to prevent an accidental explosion after the mission.

Although the number of objects is mostly comprised of fragmentation debris, the total mass is mainly made up of rocket bodies and spacecraft (S/C). In fact, the latter account for more than 96% of the total mass of debris in orbit, with fragmentation debris representing just over 3%. This is a crucial point to keep into account; the mass of an object is also an important factor, as it can represent fuel for the cascade effect. As of 2012, the total mass of space debris was more than 6000 t, and almost half of it is in LEO (below 2000 km altitude).

Recent studies [19] have shown that the measures adopted so far won't be enough to safeguard the environment from possible accidents and LEO requires more drastic measures, such as active debris removal. As for MEO and GEO, the projection in the future isn't as harsh as the one in LEO and even without any mitigation measure, the situation will be very much under control. In addition, for GEO there's the option of maneuvering the S/C after the end of the mission to a graveyard orbit, a few hundred kilometers above. Nevertheless, even though there's no urgent need for ADR in MEO and GEO for the near future, we have to keep into account the the build-up of debris will continue in these zones. As an author's note, I'd like to point out that we, as a species, ought to start thinking about how to clean our mess up there as well, so that we won't face the same situation we have to deal with now.

Given the projection for the near future, it appears obvious that we should start cleaning the LEO environment. Many objectives will drive future missions and many different paths can be taken. What is for sure is that all ADR mission concepts will have at least one thing in common: maneuvering a S/C between LEOs. Hence, this explains the first reason why the optimization of LEO trajectories is of primary importance.

1.2 On-orbit servicing

The execution of operations in orbit such as refueling, maintenance, repair and assembly is known as on-orbit servicing. From the 1970s up to the 1990s, OOS has been executed many times, a lot of which concerned the Space Shuttle program. The main reason why OOS was born was the huge interest in large space structures. As a matter of fact, the first OOS experiences were carried out for the American Skylab space station. These kind of activities required humans to perform extra-vehicular activities (EVA) in order to carry out the tasks. At first, concepts of robotic OOS were deemed unsuited, as the technology was still immature. In fact, it presented many difficulties; the speed of telecommunication was not sufficient to enable teleoperations and the subsystems didn't go anywhere near the capabilities of a human being. In addition, the economic side of the challenge was highly unfavorable. However, thanks to the turn of the new millennium, which saw the birth of the International Space Station, the interest in large structure came back strongly and brought by a renewed interest in robotic OOS. This resurgence was also aided by new economic analysis [17] that showed a market exists for the technology and also by the fact that the technology readiness level (TRL) is greatly improved [23]. These two factors culminated in a great interest for such missions at present.

Exactly as for ADR, all robotic OOS missions will require transfers between LEOs. Therefore, the results of this thesis will be of extreme interest for this application.

1.3 Small-sat deployment from platforms

The miniaturization of electronics and the increased availability of commercial-of-the-shelf (COTS) components, resulted in a growth of small satellite missions over the past decade. In addition, these spacecraft are often launched as secondary payload, that is they share a ride on a launch vehicle that is mostly paid for by the organization that commissions the launch of a larger satellite. Consequently, small satellite operators usually obtain reduced prices for orbit insertion, but they have no control over the launch date or the orbital trajectory of the launcher. Moreover, secondary payload opportunities entail requirements due to deployment mechanisms, as satellites must be compatible with such platforms. This access-to-orbit problem can be addressed with an alternative strategy. Namely, these small-sats can be transported to an orbiting platform (e.g. the ISS) and then deployed into orbit.

Figure 1.3: Deployment from the International Space Station

Up to this date, over 50 satellites have been placed in cargo transportation bags with other equipment and supplies bound for the ISS. Once there, they have been integrated into suitable deployers and placed into orbit (figure 1.3). This deployment strategy brings some major advantages. In the first place, since the satellites are placed into pressurized capsules, they are not exposed to the same kind of shock, vibration and depressurization loads as a rideshare mission. As a result, this orbit insertion strategy lowers risk and decreases costs for ground testing. In addition, resupply missions to the ISS are quite frequent and highly reliable. This means that the launch schedule is steady and predictable and small satellites operators can choose between 4-5 launch windows per year.

Since the operative orbit may be quite different from that of the ISS (or other deployment platforms), an orbit transfer may be needed. For this reason, satellite operators may be interested in either minimum-time or minimum-propellant trajectories, depending on the mission requirements.

Chapter 2

Orbital mechanics

This thesis applies the principles of optimal control theory (OCT) to the optimization of space trajectories, and thus its theoretical content is deeply rooted in the calculus of variations. However, without a basic understanding of astrodynamics, its substance and results would be meaningless. Therefore, this chapter is meant to give the reader all the fundamental prerequisites needed to grasp the contents of this work. For more details, the reader may refer to [1, 25].

2.1 Two-body orbital mechanics

The Kepler's laws, published by Johann Kepler between 1609 and 1619, marked a fundamental step towards unraveling the mysteries of planetary motion. They are:

Kepler's first Law The orbit of each planet is an ellipse, with the sun at a focus.

Kepler's second Law The line joining the planet to the sun sweeps out equal areas in equal times.

Kepler's third Law The square of the period of a planet is proportional to the cube of its mean distance form the sun.

Although these laws were a result of observations, and thus represented only a description of planetary motion, they also laid the foundations for Isaac Newton, who 50 years later figured out the reason behind the laws. In arguably one of the greatest work ever conceived by a human mind, *Philosophiae Naturalis Principia Mathematica*, Newton introduced his three laws of motion:

Newton's first Law Every body continues in its state of rest or of uniform motion in a straight line unless it is compelled to change that state by forces impressed upon it.

Newton's second Law The rate of change of momentum is proportional to the force impressed and is in the same direction of that force.

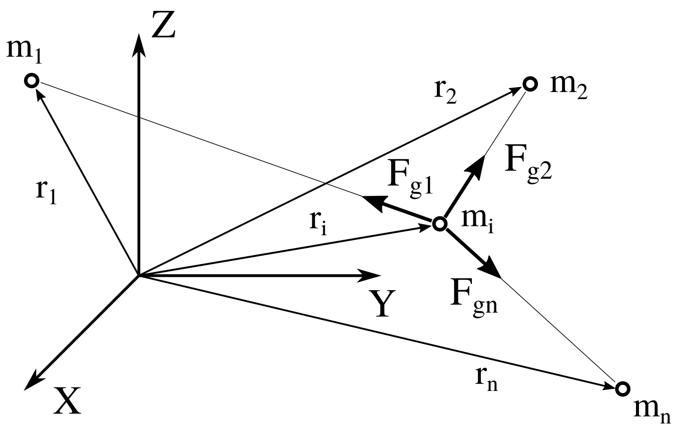


Figure 2.1: The n-body problem

Newton's third Law To every action there is always opposed an equal reaction.

For a fixed mass system, the second law can be expressed as

$$\sum \mathbf{F} = m\ddot{\mathbf{r}} \tag{2.1}$$

where $\sum \mathbf{F}$ is the vector sum of the forces acting on the fixed mass m and $\ddot{\mathbf{r}}$ is the vector acceleration of the mass relative to an inertial reference frame. In the same work, Newton introduced his law of universal gravitation:

$$\mathbf{F}_g = -G\frac{Mm}{r^2}\frac{\mathbf{r}}{r} \tag{2.2}$$

where m and M are two masses,$\mathbf{F}_g$ is the force on mass m and $\mathbf{r}$ is the vector from M to m. The universal gravitational constant G has the value $6,67 \times 10^{-11}\,\mathrm{Nm}^2/\mathrm{kg}^2$.

2.1.1 The n-body problem

At any given time, the motion of a body (which could be an artificial satellite, a planet or an interplanetary probe) is being determined by several gravitational masses and other forces, such as drag, thrust and solar radiation pressure. The objective of the problem is finding the law that describes how the position vectors of n-bodies evolve through time. Let's study the motion of the body m_i from figure 2.1,assuming spherical distribution of mass and spherical geometry of the body. These assumptions are necessary to apply Newton's law of gravitation and from Gauss's theorem we know it's equivalent to considering punctiform masses. In reality, planets and moons are not perfectly spherical, and the gravitational effects due to the shape of the bodies is responsible for many effects not described by Kepler's and Newton's laws. These effects will be discussed in Section 2.4.2. Additionally, let's assume only gravitational forces are acting upon the bodies and all the masses are constant through time. These two assumption are false, for example, when the body is moving through an atmosphere where drag effects are present, when it's expelling mass (propellant) to produce thrust or when solar radiation pressure is present. With respect to an inertial coordinate system (X,Y,Z) the position vectors of the n bodies are $\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_n$. Applying Newton's law of gravitation, the force $\mathbf{F}_{\text{gj}}$ exerted on m_i by a generic mass m_j is

$$\mathbf{F}_{\text{gj}} = -G\frac{m_i m_j}{\|\mathbf{r}_i - \mathbf{r}_j\|^2}\frac{\mathbf{r}_i - \mathbf{r}_j}{\|\mathbf{r}_i - \mathbf{r}_j\|} \tag{2.3}$$

The vector sum of all the forces acting on m_i is

$$\mathbf{F}_{\text{g}} = -Gm_i\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{m_j}{\|\mathbf{r}_i - \mathbf{r}_j\|^2}\frac{\mathbf{r}_i - \mathbf{r}_j}{\|\mathbf{r}_i - \mathbf{r}_j\|} \tag{2.4}$$

or

$$\mathbf{F}_{\text{g}} = -Gm_i\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{m_j}{\|\mathbf{r}_{ji}\|^2}\frac{\mathbf{r}_{ji}}{\|\mathbf{r}_{ji}\|} \tag{2.5}$$

where $\mathbf{r}_{ji} = \mathbf{r}_i - \mathbf{r}_j$ and $j \neq i$ is present because m_i does not exert a force on itself. Applying Newton's second law,

$$\ddot{\mathbf{r}}_i = -G\sum_{\substack{j=1\\ j\neq i}}^{n}\frac{m_j}{\|\mathbf{r}_{ji}\|^2}\frac{\mathbf{r}_{ji}}{\|\mathbf{r}_{ji}\|} \tag{2.6}$$

Extending to all bodies,

$$\begin{cases} \ddot{\mathbf{r}}_1 = -G\sum_{j=2}^{n}\frac{m_j}{\|\mathbf{r}_{j1}\|^2}\frac{\mathbf{r}_{j1}}{\|\mathbf{r}_{j1}\|} \\ \ddot{\mathbf{r}}_2 = -G\sum_{\substack{j=1\\ j\neq 2}}^{n}\frac{m_j}{\|\mathbf{r}_{j2}\|^2}\frac{\mathbf{r}_{j2}}{\|\mathbf{r}_{j2}\|} \\ \vdots \\ \ddot{\mathbf{r}}_n = -G\sum_{j=1}^{n-1}\frac{m_j}{\|\mathbf{r}_{jn}\|^2}\frac{\mathbf{r}_{jn}}{\|\mathbf{r}_{jn}\|} \end{cases} \tag{2.7}$$

gives n coupled vector differential equations and the problem has no analytical solution.

2.1.2 The two-body problem

Let us assume we want to study the motion of a body relative to another one, keeping the simplifying assumption made so far. By using the results from the

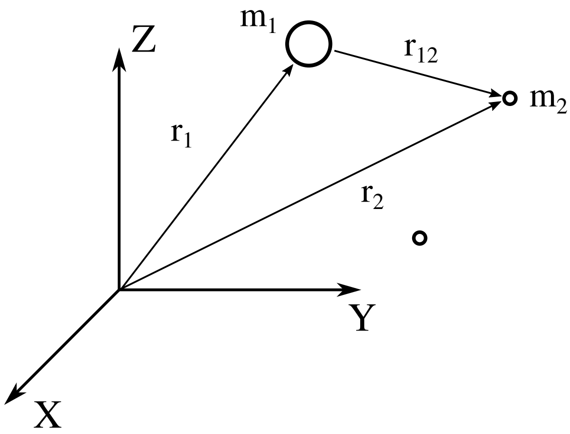


Figure 2.2: The two-body problem

previous section, let us study the motion of m_2, which for example could be an Earth satellite, relative to m_1, which could be the Earth (figure 2.2). The acceleration of m_1 relative to m_2 is

$$\ddot{\mathbf{r}}_{12} = \ddot{\mathbf{r}}_2 - \ddot{\mathbf{r}}_1 \tag{2.8}$$

Substituting the first 2 equations from (2.7) into equation (2.8) gives

$$\ddot{\mathbf{r}}_{12} = -G\sum_{\substack{j=1\\ j\neq 2}}^{n} \frac{m_j}{\|\mathbf{r}_{j2}\|^2}\frac{\mathbf{r}_{j2}}{\|\mathbf{r}_{j2}\|} - G\sum_{j=2}^{n} \frac{m_j}{\|\mathbf{r}_{j1}\|^2}\frac{\mathbf{r}_{j1}}{\|\mathbf{r}_{j1}\|} \tag{2.9}$$

or, expanding,

$$\begin{aligned}\ddot{\mathbf{r}}_{12} = &- G\frac{m_1}{\|\mathbf{r}_{12}\|^2}\frac{\mathbf{r}_{12}}{\|\mathbf{r}_{12}\|} - G\sum_{j=3}^{n} \frac{m_j}{\|\mathbf{r}_{j2}\|^2}\frac{\mathbf{r}_{j2}}{\|\mathbf{r}_{j2}\|}\\ &- G\frac{m_2}{\|\mathbf{r}_{21}\|^2}\frac{\mathbf{r}_{21}}{\|\mathbf{r}_{21}\|} - G\sum_{j=3}^{n} \frac{m_j}{\|\mathbf{r}_{j1}\|^2}\frac{\mathbf{r}_{j1}}{\|\mathbf{r}_{j1}\|}\end{aligned} \tag{2.10}$$

Since $\mathbf{r}_{12} = -\mathbf{r}_{21}$,

$$\ddot{\mathbf{r}}_{12} = -G\frac{m_1+m_2}{\|\mathbf{r}_{12}\|^2}\frac{\mathbf{r}_{12}}{\|\mathbf{r}_{12}\|} - \sum_{j=3}^{n} Gm_j\left(\frac{\mathbf{r}_{j2}}{\|\mathbf{r}_{j2}\|^3} - \frac{\mathbf{r}_{j1}}{\|\mathbf{r}_{j1}\|^3}\right) \tag{2.11}$$

Therefore, the motion of m_2 relative to m_1 depends on the gravitational effects between the two masses (first term of equation (2.11)) and on all of the other gravitational forces acting between the two masses and all of the other bodies, which could be the sun, the moon and the other planets. Since the gravitational forces between Earth and an artificial satellite are much bigger than those between them and the other bodies, the last term of equation (2.11) represents the perturbing effects. Replacing $\mathbf{r}_{12}$ with $\mathbf{r}$, that is the vector that goes from the primary body to the secondary one, m_1 with M and m_2 with m, and neglecting the perturbing effects, gives

$$\ddot{\mathbf{r}} = -G\frac{M+m}{r^3}\mathbf{r} \tag{2.12}$$

If the mass of the primary body is much bigger than that of the secondary one, we can make another simplifying assumption by saying that $G\,(M+m) \approx GM$. Let us also introduce a convenient parameter, μ, called the gravitational parameter:

$$\mu \equiv GM \tag{2.13}$$

For any given body, its gravitational parameter is the product of the gravitational constant and its mass. Table 2.1 lists the gravitational parameters of the sun and the planets of our solar system. We can now write equation (2.12) as

$$\ddot{\mathbf{r}} + \frac{\mu}{r^3}\mathbf{r} = \mathbf{0} \tag{2.14}$$

Equation (2.14) is the restricted two-body equation of motion, where the term «restricted» indicates that the gravitational pull of m on M is neglected.

Table 2.1: Gravitational parameter of the Sun and planets

Planet	**Mass/Mass Earth**	**Gravitational parameter** [km^3/s^2]
Sun	333432	1.327×10^{11}
Mercury	0.056	2.232×10^4
Venus	0.817	3.257×10^5
Earth	1	3.986×10^5
Mars	0.108	4.305×10^4
Jupiter	318.0	1.268×10^8
Saturn	95.2	3.795×10^7
Uranus	14.6	5.820×10^6
Neptune	17.3	6.896×10^6

2.1.3 Potential energy

The mechanical work per unit mass done by the gravitational force of the primary body to move the secondary one from a certain point 1 to point 2 is

$$\mathcal{L} = \int_1^2 \mathbf{F}_g \cdot d\mathbf{s} = \int_1^2 -\frac{\mu}{r^3}\mathbf{r} \cdot \mathbf{ds} = \int_1^2 -\frac{\mu}{r^3} dr = \frac{\mu}{r_2} - \frac{\mu}{r_1} \tag{2.15}$$

and does not depend on the trajectory of the body. Therefore, the gravitational field is conservative and an object moving through only a gravitational field does not lose or gain any mechanical energy, but only exchanges its kinetic and potential energy. As a matter of fact, the work done by the gravitational force can be expressed as the opposite of the change of potential energy of the body. Per unit mass one has

$$\mathcal{L} = -\Delta\mathcal{E}_g = -\left(\mathcal{E}_{g2} - \mathcal{E}_{g1}\right) = \mathcal{E}_{g1} - \mathcal{E}_{g2} \tag{2.16}$$

When the work is positive and the secondary body moves towards the primary one, the potential energy decreases, whereas when the secondary body moves away the potential energy increases. Therefore, from equations (2.15) and (2.16) one has

$$\mathcal{E}_{g1} - \mathcal{E}_{g2} = -\frac{\mu}{r_1} + \frac{\mu}{r_2} \tag{2.17}$$

with the specific potential energy defined as

$$\mathcal{E}_g = -\frac{\mu}{r} + c \tag{2.18}$$

where «specific» indicates it is per unit mass and c is an arbitrary constant. Its values depends on which reference point is used as zero of potential energy. This choice is completely arbitrary, but in astrodynamics c is conventionally set to zero. This makes the zero reference of potential energy at infinity from the primary body and also makes the potential energy always negative. Therefore, one has

$$\mathcal{E}_g = -\frac{\mu}{r} \tag{2.19}$$

2.1.4 Constants of the motion

Of course, the objective of the problem is to solve the equation of motion in order to describe the trajectory of the satellite. Before doing that, let us collect some useful insights into the physics of a body moving in a gravitational field.

Let us take the scalar product of the equation of motion with the first derivative of the position vector, that is the velocity vector:

$$\ddot{\mathbf{r}} \cdot \dot{\mathbf{r}} + \frac{\mu}{r^3}\mathbf{r} \cdot \dot{\mathbf{r}} = 0 \tag{2.20}$$

Since $\mathbf{v} = \dot{\mathbf{r}}$ and, given any vector $\mathbf{a}$, one has $\mathbf{a} \cdot \dot{\mathbf{a}} = a\dot{a}$,

$$\dot{v}v + \frac{\mu}{r^3}r\dot{r} = 0 \tag{2.21}$$

that is,

$$\frac{d}{dt}\left(\frac{v^2}{2} - \frac{\mu}{r}\right) = 0 \tag{2.22}$$

The term within the brackets is the specific mechanical energy of the body

$$\mathcal{E} = \frac{v^2}{2} - \frac{\mu}{r} \tag{2.23}$$

that is the sum of the kinetic and potential energy. As already anticipated, during the motion it is constant and an object moving under the influence of gravity alone exchanges one form of energy with the other.

Let us now take the vector product of $\mathbf{r}$ with the equation of motion,

$$\mathbf{r} \times \ddot{\mathbf{r}} + \mathbf{r} \times \frac{\mu}{r^3}\mathbf{r} = \mathbf{0} \tag{2.24}$$

which gives

$$\mathbf{r} \times \ddot{\mathbf{r}} = \mathbf{0} \tag{2.25}$$

that is,

$$\frac{d}{dt}\left(\mathbf{r} \times \mathbf{v}\right) = \mathbf{0} \tag{2.26}$$

Therefore, the specific angular momentum

$$\mathbf{h} = \mathbf{r} \times \mathbf{v} \tag{2.27}$$

is a constant vector. Since $\mathbf{h}$ is constant and must be always perpendicular to both $\mathbf{r}$ and $\mathbf{v}$, the motion of the secondary body must occur in the same plane, that is $\mathbf{r}$ and $\mathbf{v}$ must remain in the same plane. This plane is known as *orbital plane*. This result is common to all central forces, since there's no torque acting on the body. The magnitude of the specific angular momentum vector can be expressed with the *flight path angle*. As shown in figure 2.3, it's always possible to define a vertical direction and an horizontal direction, no matter where the secondary body is located in space. In the orbital plane, the local vertical at the location of the S/C is the same as the direction of $\mathbf{r}$ and the local horizontal is perpendicular to it. Thus, one may always say «up» by meaning away from the center of the primary body and «down» by meaning towards its center. Therefore, the direction of the velocity vector can be specified by the angle φ it makes with the local horizontal, known as flight path angle. From the definition of cross product the magnitude of $\mathbf{h}$ is

$$h = rv\sin\left(\pi - \varphi\right) = rv\cos\varphi \tag{2.28}$$

The angle $\pi - \varphi$ is also referred to as γ, the zenith angle. However, it's more convenient to express the specific angular momentum in terms of φ.

The cross product of the equation of motion with $\mathbf{h}$ gives

$$\ddot{\mathbf{r}} \times \mathbf{h} + \frac{\mu}{r^3}\mathbf{r} \times \mathbf{h} = \mathbf{0} \tag{2.29}$$

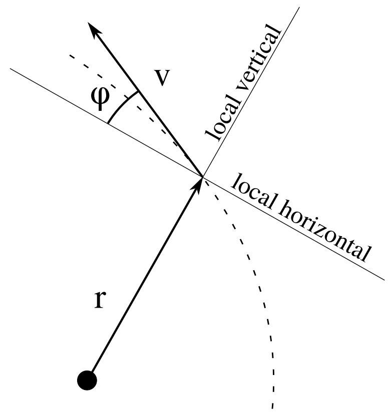


Figure 2.3: Flight path angle

which can be rewritten as

$$\frac{d}{dt}\left(\dot{\mathbf{r}} \times \mathbf{h}\right) + \frac{\mu}{r^3}\mathbf{r} \times \left(\mathbf{r} \times \dot{\mathbf{r}}\right) = \mathbf{0} \tag{2.30}$$

We can now use equations (A.6) and (A.16) from Appendix A to rewrite equation (2.30) as

$$\begin{aligned} \frac{d}{dt}\left(\dot{\mathbf{r}} \times \mathbf{h}\right) + \frac{\mu}{r^3}\mathbf{r} \times \left(\mathbf{r} \times \dot{\mathbf{r}}\right) &= \\ \frac{d}{dt}\left(\dot{\mathbf{r}} \times \mathbf{h}\right) + \frac{\mu}{r^3}\left[\left(\mathbf{r} \cdot \dot{\mathbf{r}}\right)\mathbf{r} - \left(\mathbf{r} \cdot \mathbf{r}\right)\dot{\mathbf{r}}\right] &= \\ \frac{d}{dt}\left(\dot{\mathbf{r}} \times \mathbf{h}\right) + \mu\left(\frac{\dot{r}}{r^2}\mathbf{r} - \frac{\dot{\mathbf{r}}}{r}\right) &= \\ \frac{d}{dt}\left(\dot{\mathbf{r}} \times \mathbf{h}\right) - \mu\frac{d}{dt}\left(\frac{\mathbf{r}}{r}\right) &= \mathbf{0} \end{aligned} \tag{2.31}$$

By integrating, one finds that the vector

$$\mathbf{B} = \mathbf{v} \times \mathbf{h} - \mu\frac{\mathbf{r}}{r} \tag{2.32}$$

is another constant of the motion in the two-body problem.

2.1.5 Trajectory equation

Let us now take the scalar product of $\mathbf{r}$ with equation (2.32),

$$\mathbf{r} \cdot \mathbf{B} = \mathbf{r} \cdot \mathbf{v} \times \mathbf{h} - \mathbf{r} \cdot \mu\frac{\mathbf{r}}{r} \tag{2.33}$$

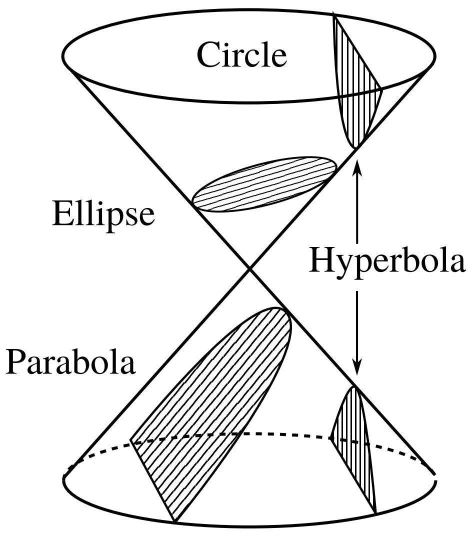


Figure 2.4: Conic sections

By using equation (A.18) one has

$$rB\cos\nu = h^2 - \mu r \tag{2.34}$$

where ν is the angle between $\mathbf{B}$ and $\mathbf{r}$. Let us finally solve for r and obtain

$$r = \frac{h^2/\mu}{1 + B/\mu\cos\nu} \tag{2.35}$$

Equation (2.35) is the trajectory equation in polar coordinates and it describes how the secondary body moves with respect to the primary body. Before making any observation with regard to this equation, let us revise some basic concepts of conic sections.

2.1.6 Conic sections

A conic section can be defined as a curve obtained from the intersection of a plane and a right circular cone (figure 2.4). If the plane crosses one half of the cone the section is called *ellipse*. If the plane is also parallel to the base of the cone, the section is a *circle*. On the other hand, if the plane is parallel to a line in the surface of the cone, the section is a *parabola*. Finally, if the plane cuts across both halves of the cone, the section is called *hyperbola* and it has two branches. There also can be degenerate conics, like one or two straight lines and a single point, when the plane cuts across the apex of the cone. The mathematical translation of this geometric definition is the following. A conic section is the locus of points such that the ratio of the distance r from a given point, called *focus,* to its distance d from a given line,

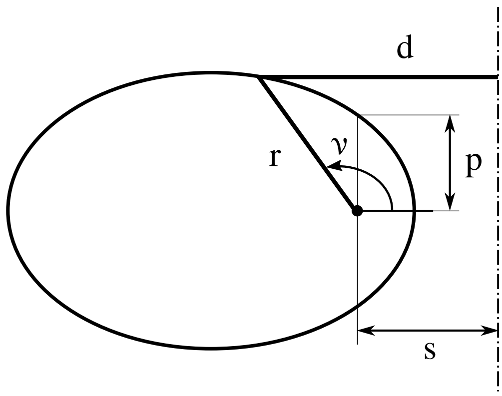


Figure 2.5: Conic sections parameters

called *directrix,* is a positive constant e, called *eccentricity*:

$$e = \frac{r}{d} \tag{2.36}$$

Letting s be the distance between the focus and the directrix, as shown in figure 2.5, one has

$$d = s - r\cos\nu \tag{2.37}$$

which can be rewritten as

$$r = \frac{p}{1 + e\cos\nu} \tag{2.38}$$

where $p = es$ is a geometrical constant of the conic section known as *parameter* or *semilatus rectum*. Equation (2.38) is the general equation of a conic section in polar coordinates with the origin located at a focus and the polar angle ν defined as the angle between the position vector and the point of the conic section nearest to the focus. Equation (2.38) is formally identical to the trajectory equation. This verifies Kepler's first law and extends it to include orbital motion along any conic section, not just ellipses. The semilatus rectum of the conic, that is the distance between the focus and the point in the trajectory where $\nu = \pi/2$, is related to the angular momentum of the S/C:

$$p = \frac{h^2}{\mu} \tag{2.39}$$

In addition, the eccentricity of the conic section is the magnitude of $\mathbf{B}/\mu$ and

$$\mathbf{e} = \frac{\mathbf{B}}{\mu} \tag{2.40}$$

is called eccentricity vector.

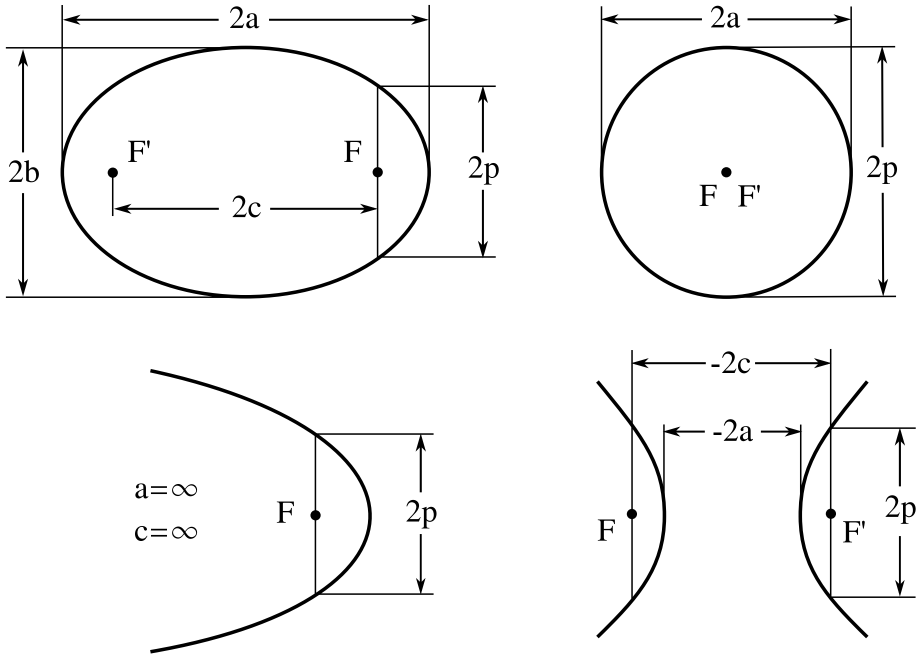


Figure 2.6: Geometry of conic sections

Among the conic parameters, for orbital mechanics the directrix has no physical significance. On the other hand, the focus, the eccentricity and the semilatus rectum are important parameters. Figure 2.6 shows the geometrical parameters of conic sections. Physically, the prime focus F represents the location of the primary body and the second focus F' has no physical significance. The parabola constitutes the limit between closed and open orbits and its second focus lies at an infinite distance from the prime. The length of the chord denoted as $2p$ is the *latus rectum* and the length of the chord between the line of the foci and its intersection with the conic section, denoted as $2a$, is the *major axis*. Following from this definition, the dimension a is known as *semimajor axis*. The semimajor axis of the circle is its radius, for the parabola it is infinite and for the hyperbola is taken as negative. The width of an ellipse at the center, denoted as $2b$, is called *minor axis* and the dimension b is known as *semiminor axis*. The distance between the two foci is denoted as $2c$. For the circle it is zero, as the two foci coincide, for the parabola it is infinite and for the hyperbola it is taken as negative. Following from the definition of a conic section one can easily prove that

$$e = \frac{c}{a} \tag{2.41}$$

and

$$p = a\left(1 - e^2\right) \tag{2.42}$$

always hold true except for the parabola. The extreme points of the major axis are called *apses*. The point nearest the prime focus is called *periapsis* and the point farthest from the prime focus is called *apoapsis*. This nomenclature may also vary depending on the primary body. As a matter of fact, if the primary body is the Earth we may also specify it by saying *perigee* and *apogee*, or, in the case of the Sun, *perihelion* and *apohelion*. For the circle this nomenclature is obviously not definable and the *apoapsis* has no meaning for open trajectories.

From equation (2.35) we see that the radius reaches its minimum value for $\nu = 0$. Therefore, the eccentricity vector is directed towards the periapsis. The polar angle ν, between the position vector and the vector the points towards the periapsis, is known as *true anomaly*. The distance from the primary body to either periapsis or apoapsis can be obtained by inserting 0 or π as true anomaly in equation (2.35). Therefore, for any conic section one has

$$r_P = \frac{p}{1+e} \tag{2.43}$$

$$r_A = \frac{p}{1-e} \tag{2.44}$$

where the subscripts P and A indicate the periapsis and the apoapsis, respectively. Combining equations (2.43) and (2.44) with equation (2.42) gives

$$r_P = a\,(1-e) \tag{2.45}$$

$$r_A = a\,(1+e) \tag{2.46}$$

2.1.7 Relating energy to the geometry of an orbit

The total specific energy can be evaluated at any point in the trajectory. In particular, at the periapsis one has

$$\mathcal{E} = \frac{v_P^2}{2} - \frac{\mu}{r_P} \tag{2.47}$$

The velocity at the periapsis can be expressed in terms of specific angular momentum. From equation (2.28) one has

$$v_P = \frac{h}{r_P \cos\varphi} \tag{2.48}$$

From the definition of flight path angle, one has

$$\tan\varphi = \frac{v_r}{v_t} \tag{2.49}$$

where

$$v_r = v\sin\varphi \tag{2.50}$$

is the radial component of the velocity and

$$v_t = v \cos\varphi \tag{2.51}$$

its tangential component. Substituting equation (2.40) in (2.35) we can write

$$r\left(1 + e\cos\nu\right) = \frac{h^2}{\mu} \tag{2.52}$$

Differentiating both side,

$$\dot{r}\left(1 + e\cos\nu\right) - r\dot{\nu}e\sin\nu = 0 \tag{2.53}$$

Since $v_r = \dot{r}$ and $v_t = r\dot{\nu}$, one has

$$\frac{v_r}{v_t} = \frac{e\sin\nu}{1 + e\cos\nu} \tag{2.54}$$

Substituting (2.54) into (2.49) gives

$$\tan\varphi = \frac{e\sin\nu}{1 + e\cos\nu} \tag{2.55}$$

At the periapsis one has

$$\varphi_P = \frac{e\sin 0}{1 + e\cos 0} = 0 \tag{2.56}$$

We can now use this result and rewrite (2.48) as

$$v_P = \frac{h}{r_P} \tag{2.57}$$

Let us know substitute (2.57) into (2.47) and obtain

$$\mathcal{E} = \frac{h^2}{2r_p^2} - \frac{\mu}{r_P} \tag{2.58}$$

Combining equations (2.39) and (2.42) gives

$$h^2 = \mu a\left(1 - e^2\right) \tag{2.59}$$

Let us now substitute equations (2.59) and (2.45) into (2.58) to finally obtain

$$\mathcal{E} = -\frac{\mu}{2a} \tag{2.60}$$

Equation (2.60) is valid for all the types of orbit. For closed orbits (circle and ellipse) a is positive and the total mechanical energy of the S/C is negative, while for the parabola a is infinite and the energy is zero and for the hyperbola a is negative and the energy is positive. Combining equations (2.59) and (2.60) gives

$$e = \sqrt{1 + \frac{2\mathcal{E}h^2}{\mu^2}} \tag{2.61}$$

For any conic orbit, the energy and the angular momentum determine the eccentricity of the orbit, which specifies the shape of the orbit. When the orbit is closed $\mathcal{E}$ is negative and $e < 1$; if $\mathcal{E}$ is positive the orbit is an hyperbola $e > 1$; if $\mathcal{E}$ is zero the orbit is a parabola and $e = 1$. However, when h is zero e is also equal to 1; as a matter of fact, in this case the orbit is a degenerate conic (straight downfall to the primary body). Therefore, all parabolas have $e = 1$ but an orbit whose eccentricity is 1 may not be a parabola.

2.1.8 Closed orbits

The orbits of all Earth satellites are ellipses. Furthermore, the orbits object of this thesis are almost circular and their eccentricity is close to zero. For this reason, this section will go over some basic concepts of closed orbits, whereas parabolas and hyperbolas won't be treated in this work.

2.1.8.1 Period of an elliptical orbit

Since an ellipse is a closed curve, a S/C on an elliptical orbit travels the same path at each revolution. The time it takes the S/C to complete one revolution is called the *period*. From equations (2.28) and (2.51) we know that

$$h = rv_t \tag{2.62}$$

Using $v_t = r\dot{\nu}$ we can write

$$h = r^2 \frac{d\nu}{dt} \tag{2.63}$$

The differential element of area dA swept out by $\mathbf{r}$ as it moves through an infinitesimal angle $d\nu$ is given by

$$dA = \frac{r^2}{2} d\nu \tag{2.64}$$

We can now use this expression to rewrite 2.64 as

$$\frac{dA}{dt} = \frac{h}{2} \tag{2.65}$$

which proves Kepler's second law that the radius vector sweeps equal areas in equal time, as h is a constant. Rewriting equation (2.65) as

$$dt = \frac{2}{h} dA \tag{2.66}$$

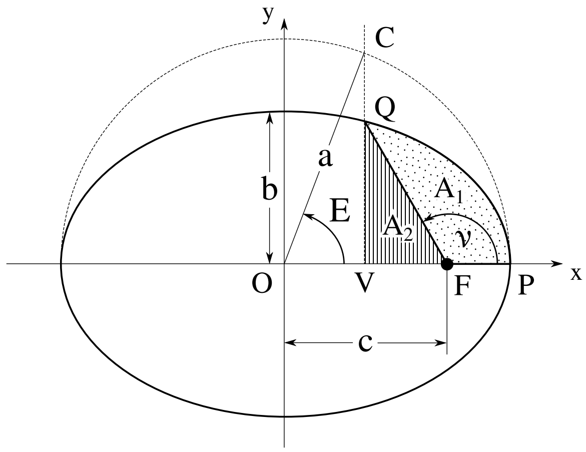


Figure 2.7: Eccentric anomaly, E

and integrating for one period gives

$$T_E = \frac{2}{h}\int\limits_{A_E} dA = \frac{2\pi ab}{h} \tag{2.67}$$

where $A_E = \pi ab$ is the area of an ellipse and T_E is the period of the orbit. From simple geometry, one can easily prove that

$$a^2 = b^2 + c^2 \tag{2.68}$$

so we can now write, by also using equations (2.41), (2.42),

$$T_E = \frac{2\pi a\sqrt{a^2 - c^2}}{h} = \frac{2\pi a\sqrt{a^2\left(1-e^2\right)}}{h} = \frac{2\pi a\sqrt{ap}}{h} \tag{2.69}$$

and, from equation (2.39), we can finally write

$$T_E = 2\pi\sqrt{\frac{a^3}{\mu}} \tag{2.70}$$

The period of an elliptical orbit depends only on the size of the semimajor axis. Furthermore, equation (2.70) proves Kepler's third law, since the semimajor axis is the «mean distance» of a S/C from the focus.

2.1.8.2 Time of flight on an elliptical orbit

The time of flight t_Q from the periapsis P to any point Q of the elliptical orbit can be evaluated with the area A_1 of figure 2.7, swept out by $\mathbf{r}$. As a matter if fact, from Kepler's second law one has

$$t_Q = T_E \frac{A_1}{A_E} = 2\sqrt{\frac{a^3}{\mu}} \frac{A_1}{ab} \tag{2.71}$$

In order to evaluate A_1, let us introduce a circle of radius a, centered in the center of the ellipse O. The point C on the circle with the same abscissa as the point Q and the angle E is called *eccentric anomaly*. Let us write the Cartesian equations of a circle and an ellipse centered in O:

$$\begin{cases} \frac{x_Q}{a^2} + \frac{y_Q}{b^2} = 1 \\ \frac{x_C}{a^2} + \frac{y_C}{a^2} = 1 \end{cases} \tag{2.72}$$

Since $x_Q \equiv x_C$ one realizes that

$$\frac{y_Q}{y_C} = \frac{b}{a} \tag{2.73}$$

From figure 2.7 one has that the area A_1 is given by area QVP minus area A_2:

$$A_1 = A_{QVP} - A_2 \tag{2.74}$$

Since A_2 is a right triangle whose base is $c - a\cos E$ and whose altitude is $(a \sin E)\, b/a$, we can write

$$\begin{aligned} A_2 &= \frac{b}{2a} a \sin E \,(c - a\cos E) = \\ &= \frac{b}{2} \sin E \,(ea - a\cos E) = \\ &= \frac{ab}{2} (e \sin E - \cos E \sin E) \end{aligned} \tag{2.75}$$

As for A_{QVP}, one has

$$A_{QVP} = A_{CVP} \frac{b}{a} \tag{2.76}$$

where A_{CVP} is the area of the sector COP minus the triangle whose base is $a\cos E$ and whose altitude is $a \sin E$. Thus,

$$A_{QVP} = \left(\frac{1}{2} a^2 E - \frac{a^2}{2} \cos E \sin E \right) \frac{b}{a} = \frac{ab}{2} (E - \cos E \sin E) \tag{2.77}$$

Substituting equations (2.77) and (2.75) into (2.74) gives

$$A_1 = \frac{ab}{2} (E - e \sin E) \tag{2.78}$$

Finally, substituting into equation (2.71) yields

$$t_Q = \sqrt{\frac{a^3}{\mu}}\,(E - e\sin E) = \frac{M}{n} \tag{2.79}$$

where, according to Kepler, the *mean motion* n is

$$n = \sqrt{\frac{\mu}{a^3}} \tag{2.80}$$

and the *mean anomaly* M is

$$M = E - e\sin E \tag{2.81}$$

In order to use equation (2.79) to determine the time of flight, we have to relate the eccentric anomaly to the true anomaly. From figure 2.7,

$$\cos E = \frac{c + r\cos\nu}{a} = \frac{ea + r\cos\nu}{a} \tag{2.82}$$

Combining equations (2.38) and (2.42) gives

$$r = \frac{a\left(1 - e^2\right)}{1 + e\cos\nu} \tag{2.83}$$

Therefore, we can rewrite equation (2.82) as

$$\cos E = \frac{e + \cos\nu}{1 + e\cos\nu} \tag{2.84}$$

and the correct quadrant for E is obtained by considering that ν and E are always in the same half-plane.

2.1.8.3 Circular orbit

The circular orbit is a special case of an elliptical orbit and the semimajor axis of a circle is just its radius, so equation (2.70) becomes

$$T_c = 2\pi\sqrt{\frac{r^3}{\mu}} \tag{2.85}$$

The speed of a S/C in a circular orbit v_c is called circular speed, which can be easily derived from equations (2.23) and 20:

$$\mathcal{E} = \frac{v_c^2}{2} - \frac{\mu}{r} = -\frac{\mu}{2a} \tag{2.86}$$

Since $a = r$ we have

$$v_c = \sqrt{\frac{\mu}{r}} \tag{2.87}$$

The greater the radius the less is the speed of a S/C in a circular orbit. For a S/C in LEO the circular speed may vary from 7.7 km/s at 200 km altitude to 6.9 km/s at 2000 km.

The angular velocity of a S/C in a circular orbit is

$$n_c = \frac{v_c}{r} = \sqrt{\frac{\mu}{r^3}} \tag{2.88}$$

The true anomaly swept out by the radius vector in t_Q is

$$\nu = n_c t_Q \tag{2.89}$$

Equation (2.89) unveils the meaning of the mean anomaly M. As a matter of fact, from equations (2.80) and (2.88), the mean motion n is, for a given orbit, the angular velocity of a circular orbit with radius equal to the semimajor axis. Therefore, rewriting equation (2.79),

$$M = n t_Q \tag{2.90}$$

and comparing it to equation (2.89) one realizes that the mean anomaly of a given orbit is, after t_Q time is passed from the passage at the periapsis, the same as the true anomaly swept out by a S/C in a circular orbit with the same semimajor axis during t_Q.

2.2 Coordinate systems

Depending on the type of space mission, it is necessary to provide a suitable coordinate system and a set of coordinates. The choice of a «suitable» coordinate system is far from being a simple task, as the spatial description of an orbit ideally requires an inertial reference frame. However, any coordinate system one can define has a certain degree of uncertainty to its inertial qualities. In practice, a coordinate system centered in the primary body and with axes pointing a fixed direction with respect to the fixed stars does the job of being almost inertial. Rectangular coordinate systems are the most used in astrodynamics, although sometimes spherical polar coordinates are more practical. In order to describe a rectangular coordinate system, one has to give the position of the origin, the orientation of the fundamental plane, on which the X and the Y axes lie, the positive direction of the Z axis and the principal X direction (assuming X, Y and Z form a right handed set of coordinate axes).

2.2.1 The heliocentric-ecliptic coordinate system

The origin of the heliocentric-ecliptic coordinate system is the center of the Sun. The fundamental plane is the *ecliptic*, that is the plane the Earth orbits on. As

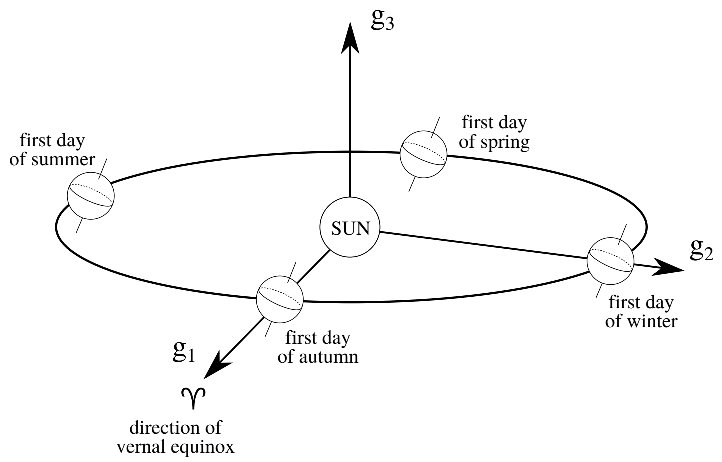


Figure 2.8: Heliocentric-ecliptic coordinate system (seasons are for Northern Hemisphere)

shown in figure 2.8, the direction of the the X axis, labeled as $\mathbf{g}_1$, is given by the line of intersection between the ecliptic plane and the Earth's equatorial plane. The positive direction of $\mathbf{g}_1$, called the vernal equinox direction (symbol Υ because it used to point in the direction of the constellation of Aries), is given by a line that, on the first day of spring points, from the center of the Earth points towards the center of the Sun. As the Earth's axis precesses over the centuries, the line of intersection between the ecliptic and Earth's equatorial plane slowly drifts clockwise in what is known as precession of the equinoxes. Therefore, this coordinate system is not really inertial and, where precision is required, the set of coordinates of an object are usually specified by saying they were based on the vernal equinox direction of a given year, or *epoch*.

2.2.2 The geocentric-equatorial coordinate system

The origin of the geocentric-equatorial coordinate system (figure 2.9) is the center of the Earth. The fundamental plane is the equator and the positive X axis points in the vernal equinox direction. The Z axis points in the direction of the north pole. Therefore, this coordinate system is not turning with the Earth, but rather it is fixed with respect to the stars (neglecting the precession of the equinoxes). This reference system, defined by unit vectors $\mathbf{I}$, $\mathbf{J}$ and $\mathbf{K}$, is obviously extremely useful for describing the motion of Earth's satellites.

2.2.3 The right ascension-declination coordinate system

The right ascension-declination coordinate system (figure 2.10) is strictly related to the geocentric-equatorial system. The fundamental plane is the celestial equator,

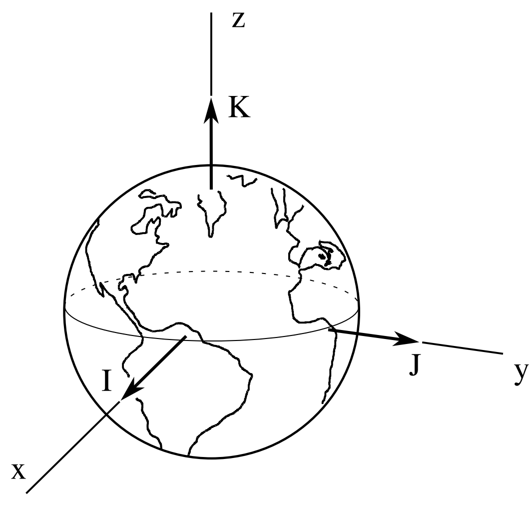


Figure 2.9: Geocentric-equatorial coordinate system

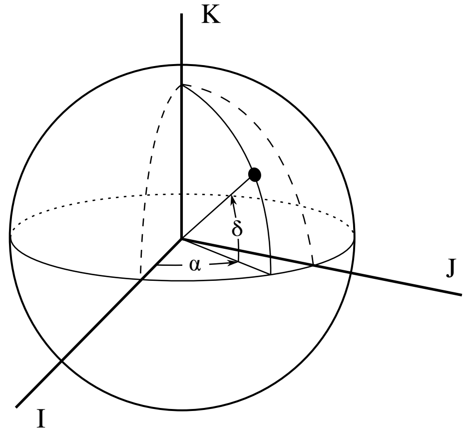


Figure 2.10: Right ascension-declination coordinate system

i.e. the extension of the equatorial plane to a fictitious sphere of infinite radius known as *celestial sphere*. The projection of an object on the celestial sphere is described by two angles, called *right ascension* and *declination*. The right ascension α is measured eastward in the plane of celestial equator from the vernal equinox direction. The declination δ is measured northward from the celestial equator to the line from the origin of the system to the projection of the object on the celestial sphere. Since the celestial sphere is infinite, its center may be any point. Therefore, we may choose the center of the Earth as origin of the system as well as a point on its surface.

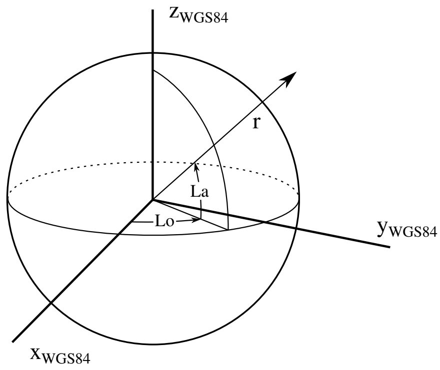


Figure 2.11: WGS 84

2.2.4 The world geodetic system 84

The origin of World Geodetic System 84 (WGS 84) is the center of mass of the Earth. The fundamental plane is the equatorial plane and the positive X axis points towards the intersection between the equatorial plane and the Greenwich meridian. Therefore, this coordinate system rotates about its Z axis with the same angular velocity as the Earth. Any point can be identified with a radius $\mathbf{r}$ and two angles, Lo (Longitude) and La (Latitude), as shown in figure 2.11.

2.2.5 The perifocal coordinate system

The origin of the perifocal coordinate system (figure 2.12) is the center of the primary body. The fundamental plane is the plane of the orbit of the secondary body. The X axis (unit vector $\mathbf{p}$) points towards the periapsis; the Y axis (unit vector $\mathbf{q}$) lies in the orbital plane and its rotated in the direction of the orbital motion. The positive direction of the Z axis (unit vector $\mathbf{w}$) is the positive direction of the angular momentum vector. The vector $\mathbf{r}$ is given by

$$\mathbf{r} = r\cos\nu\mathbf{p} + r\sin\nu\mathbf{q} + 0\mathbf{w} \tag{2.91}$$

Taking the time derivative, one finds that $\mathbf{v}$ is given by

$$\mathbf{v} = (\dot{r}\cos\nu - r\dot{\nu}\sin\nu)\,\mathbf{p} + (\dot{r}\sin\nu + r\dot{\nu}\cos\nu)\,\mathbf{q} + 0\mathbf{w} \tag{2.92}$$

Since $v_r = \dot{r}$ and $v_t = r\dot{\nu}$ one has

$$\mathbf{v} = (v_r\cos\nu - v_t\sin\nu)\,\mathbf{p} + (v_r\sin\nu + v_t\cos\nu)\,\mathbf{q} + 0\mathbf{w} \tag{2.93}$$

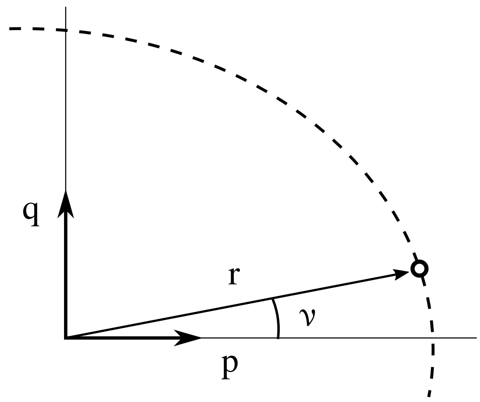


Figure 2.12: Perifocal coordinate system

Combining equations (2.28), (2.35) and (2.51)gives

$$v_t = \frac{h}{r} = \frac{\mu}{h}\left(1 + e\cos\nu\right) \tag{2.94}$$

From equation (2.54) one has

$$v_r = v_t \frac{e\sin\nu}{1 + e\cos\nu} \tag{2.95}$$

Combining (2.94) and (2.95) yields

$$v_r = \frac{\mu}{h} e\sin\nu \tag{2.96}$$

Since $\varphi = \arctan\left(v_r/v_t\right)$, from equation 2.96 we deduce that φ is negative when $\sin\nu$ is negative and vice versa:

$$\begin{cases} \varphi > 0 & \text{if } 0 < \nu < \pi \\ \varphi < 0 & \text{if } \pi < \nu < 2\pi \end{cases} \tag{2.97}$$

We can finally insert (2.94) and (2.96) into equation (2.93) to obtain

$$\mathbf{v} = -\frac{\mu}{h}\sin\nu\mathbf{p} + \frac{\mu}{h}\left(e + \cos\nu\right)\mathbf{q} + 0\mathbf{w} \tag{2.98}$$

If we know calculate

$$\mathbf{r}\cdot\mathbf{v} = \frac{\mu}{h} re\sin\nu \tag{2.99}$$

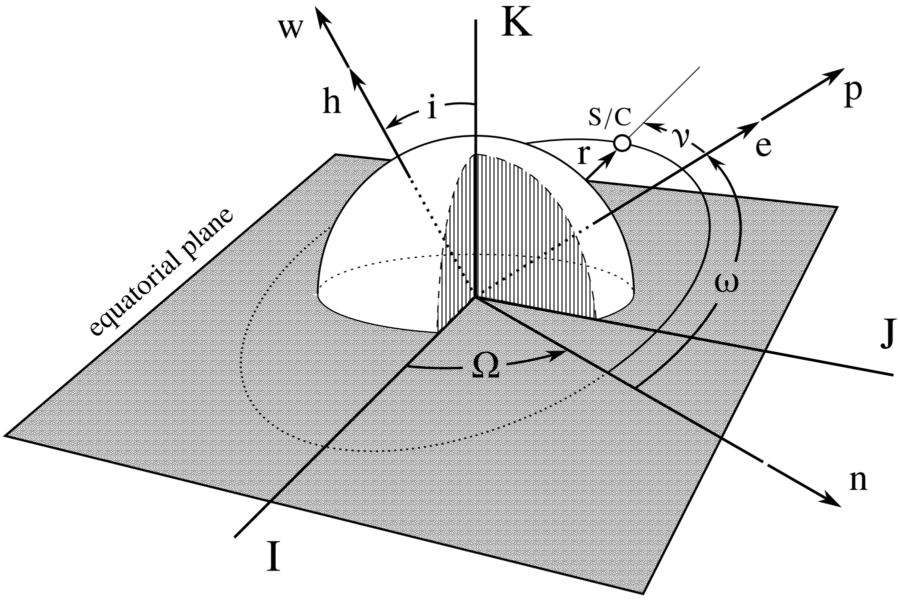


Figure 2.13: Classical orbital elements

we realize that, from equation (2.97),

$$\begin{cases} \varphi > 0 & \text{if } \mathbf{r} \cdot \mathbf{v} > 0 \\ \varphi < 0 & \text{if } \mathbf{r} \cdot \mathbf{v} < 0 \end{cases} \tag{2.100}$$

2.3 Classical orbital elements

The shape and the orientation of an orbit, as well as the position of a S/C on that orbit, are completely described by the position vector and the velocity vector. Indeed, any consistent set of six parameters can describe a two-body problem orbit. Five independent quantities unambiguosly define an orbit's shape and orientation, and a sixth quantity is needed to pinpoint the position of a S/C along that orbit. Using the position vector and the velocity vector (together they give six independent quantities) is extremely unpractical and other six parameters are preferred: the classical orbital elements (figure 2.13). These parameters are obviously dependent on $\mathbf{r}$ and $\mathbf{v}$ and can be determined directly from them. The classical orbital elements are:

1. a, semimajor axis. It defines the dimension of the orbit.

2. e, eccentricity. It defines the shape of the orbit.

3. $\imath$, inclination. It is defined by the angle between the $\mathbf{K}$ unit vector and the and

the angular momentum vector:

$$\imath = \arccos\left(\mathbf{K} \cdot \frac{\mathbf{h}}{h}\right) \tag{2.101}$$

The inclination is always greater than or equal to zero radians and less than or equal to π radians: $0 \leq \imath \leq \pi$

4. Ω, right ascension of the ascending node (RAAN). It is defined as the angle in the fundamental plane between the $\mathbf{I}$ unit vector and the point where the S/C crosses through the fundamental plane in a northerly direction, hence «ascending», measured counterclockwise when viewed from the north side of the fundamental plane. Mathematically, we can define the line of nodes with unit vector
$$\mathbf{n} = \frac{\mathbf{K} \times \mathbf{h}}{\|\mathbf{K} \times \mathbf{h}\|} \tag{2.102}$$
pointing towards the ascending node. Following this definition, we can write
$$\begin{cases} \Omega = \arccos\left(\mathbf{I} \cdot \mathbf{n}\right) & \text{if } \mathbf{n} \cdot \mathbf{J} \geq 0 \\ \Omega = 2\pi - \arccos\left(\mathbf{I} \cdot \mathbf{n}\right) & \text{if } \mathbf{n} \cdot \mathbf{J} < 0 \end{cases} \tag{2.103}$$
since Ω can assume any value from zero radians to 2π radians.

5. ω, argument of periapsis. It is defined as the angle in the plane of the orbit of the S/C between the ascending node and the periapsis, measured in the direction of the motion of the S/C. Mathematically,
$$\begin{cases} \omega = \arccos\left(\mathbf{n} \cdot \mathbf{p}\right) & \text{if } \mathbf{p} \cdot \mathbf{K} \geq 0 \\ \omega = 2\pi - \arccos\left(\mathbf{n} \cdot \mathbf{p}\right) & \text{if } \mathbf{p} \cdot \mathbf{K} < 0 \end{cases} \tag{2.104}$$

6. ν, true anomaly. Given the other orbital elements, it pinpoints the position of the S/C along the orbit. The mathematical definition is, taking into account equation (2.100),
$$\begin{cases} \nu = \arccos\left(\frac{\mathbf{r}}{r} \cdot \mathbf{p}\right) & \text{if } \mathbf{r} \cdot \mathbf{v} \geq 0 \\ \nu = 2\pi - \arccos\left(\frac{\mathbf{r}}{r} \cdot \mathbf{p}\right) & \text{if } \mathbf{r} \cdot \mathbf{v} < 0 \end{cases} \tag{2.105}$$

The orbital elements a and e define the geometry of the orbit, that is the dimension and the shape. The orbital elements i, Ω and ω define the orientation of the orbit and, finally, ν reveals where the S/C is along the orbit.

Sometimes, other angle are also used. The angle

$$\Pi = \Omega + \omega \tag{2.106}$$

is known as *longitude of periapsis*. It is the angle from $\mathbf{I}$ to periapsis, measured in the equatorial plane and then in the orbital plane. If the orbit is equatorial, that is the inclination is zero, the line of nodes is not definable and it is convenient to use this angle.

If the orbit is circular, the argument of periapsis ω is not definable and it is convenient to use the angle

$$\vartheta = \omega + \nu \tag{2.107}$$

known as *argument of latitude at epoch*, that is the angle in the plane of the orbit between the ascending node and the radius vector at a particular time (epoch).

In the case of circular orbit Π is not definable, whereas for equatorial orbit μ is not definable. Therefore, in the case of circular equatorial orbit, it is convenient to use the angle

$$l = \Omega + \omega + \nu \tag{2.108}$$

also known as *true longitude at epoch*.

2.4 Perturbations

So far we have discussed trajectories that describe the relative motion of two spherically symmetric bodies under the action of gravity alone. However, real S/C trajectories are not described accurately by the two-body restricted problem. As a matter of fact, a S/C is subject to several perturbations, such as the presence of other attractive bodies, atmospheric drag and lift, the asphericity of the attractive bodies, solar radiation effects, magnetic effects and thrust. These perturbations, defined as deviation from the expected motion, may have different consequences on the trajectory of a S/C depending on the space mission. However, in the short run the orbits of planets and satellites are well approximated by the two-body problem, which makes it very convenient to describe the motion of celestial bodies. Nevertheless, in the long run, the perturbations will make the trajectory of the S/C very different from the one predicted by the two-body problem. The orbital elements slowly change and the S/C appears to continuously pass from a Keplerian orbit to another. Given a specific time, the Keplerian orbit that has the same orbital elements of the real perturbed trajectory is called *osculating orbit*. Following this definition, the orbital elements of the S/C at a given time are the osculating orbital elements. Generally speaking, the variation of an orbital element can be secular, i.e. it follows a linear trend with time, or it can be subject to long period oscillations as a result of the interaction with the variation of other orbital elements, or also subject to short-period oscillations, usually caused by phenomena that periodically occur during each revolution. In all cases, the restricted two-body equation of motion can be formally written as

$$\ddot{\mathbf{r}} = -\frac{\mu}{r^3}\mathbf{r} + \mathbf{f}_P \tag{2.109}$$

where $\mathbf{f}_P$ is the perturbing acceleration. In general, the equation doesn't have an analytical solution and different techniques can be adopted to integrate it. These techniques are referred to as *special perturbations* and *general perturbations*. Special perturbations methods consist in the numerical integration of the equation of motion and yield the trajectory of a S/C given a set of initial condition at a particular time, therefore obtaining a specific solution for a well-defined case. On the other side, general perturbation methods provide an approximated analytical solution by expanding perturbing accelerations into series, which are suitably truncated and integrated.

2.4.1 Variation of parameters

As already discussed, a two-body problem orbit can be described by any suitable set of six parameters. The variation of parameters method describes how any of these sets of parameters vary with time as a result of perturbations. The method consists in the analytical description of the rate of change of the parameters due to the perturbations. Since in this work the method will be applied with the classical orbital elements, it can also be called *variation of elements*. In addition, this method can be considered either as a special perturbation method or a general perturbation method, depending on how the expression that the method yields are then integrated. This methodology is associated with two formulations; the first one leads to the *Gauss planetary equations*, which relate the components of the perturbing acceleration to the rate of change of the orbital elements; the second one, not of interest for this work, leads to the *Lagrange planetary equations*.

2.4.1.1 The Gauss planetary equations

As already stated, the Gauss planetary equations relate the components of a perturbing acceleration to the rate of change of the orbital elements. Depending on convenience, the perturbing acceleration can be projected in two different ways. One coordinate system has its principle axis R with unit vector $\mathbf{R}$ along the instantaneous radius vector $\mathbf{r}$. The axis T with unit vector $\mathbf{T}$ is rotated 90° in the orbital plane in the direction of motion. The third axis W with unit vector $\mathbf{W}$ is perpendicular to the other two. The other coordinate system has its principle axis V with unit vector $\mathbf{V}$ along the instantaneous velocity vector $\mathbf{v}$. The axis N, with unit vector $\mathbf{N}$, is rotated through a 90° angle in the orbital plane about $\mathbf{W}$. The third axis $\mathbf{W}$ is the same as for the other coordinate system. Therefore, following this definitions we have $\mathbf{f}_P = f_R\mathbf{R} + f_T\mathbf{T} + f_W\mathbf{W} = f_V\mathbf{V} + f_N\mathbf{N} + f_w\mathbf{W}$. The relations between the two types of projections are

$$f_R = \frac{h}{pV}\left[(e\sin\nu)\, f_V - (1+e\cos\nu)\, f_N\right] \tag{2.110}$$

$$f_T = \frac{h}{pV}\left[(1+e\cos\nu)\, f_V + (e\sin\nu)\, f_N\right] \tag{2.111}$$

The derivation of the rate of change of the orbital elements, shown in depth in [25], leads to the following results:

$$\begin{aligned}
\dot{a} &= \frac{2a^2}{\sqrt{\mu p}}\left[(e\sin\nu)\, f_R - \left(\frac{p}{r}\right) f_T\right] = \left(\frac{2a^2V}{\mu}\right) f_V \\
\dot{e} &= \sqrt{\frac{p}{\mu}}\left[(\sin\nu)\, f_R + \left(\frac{r}{p}\right)\left(e\cos^2\nu + 2\cos\nu + e\right) f_T\right] = \\
&= \frac{1}{V}\left[2\,(e+\cos\nu)\, f_V - \left(\frac{a\sin\nu}{r}\right) f_N\right] \\
\dot{\imath} &= \frac{r}{\sqrt{\mu p}}\cos(\omega+\nu)\, f_W \\
\dot{\Omega} &= \frac{r}{\sqrt{\mu p}}\frac{\sin(\omega+\nu)}{\sin(\imath)} f_W \\
\dot{\omega} &= -\dot{\Omega}\cos\imath - \frac{\sqrt{\frac{p}{\mu}}}{e}\left[(\cos\nu)\, f_R - \sin\nu\left(1+\frac{r}{p}\right) f_T\right] = \\
&= -\dot{\Omega}\cos\imath + \frac{1}{eV}\left[(2\sin\nu)\, f_V + \frac{r}{p}\left(2e + \cos\nu + e^2\cos\nu\right) f_N\right] \\
\dot{M} &= n - \frac{2r}{\sqrt{ua}} f_R - \sqrt{1-e^2}\left(\dot{\omega} + \dot{\Omega}\cos\imath\right) = \\
&= n - \frac{2\sqrt{1-e^2}}{V}\left[\left(\frac{er\sin\nu}{p}\right) f_V - f_N\right] - \sqrt{1-e^2}\left(\dot{\omega} + \dot{\Omega}\cos\imath\right)
\end{aligned} \tag{2.112}$$

Where the mean anomaly is used instead of the true anomaly (they are related by equations (2.81) and (2.84)).

By looking at the equation for the rate of change of the semimajor axis, one notes that the only component of the perturbing acceleration that has a role in its variation is the one along the velocity vector. The eccentricity can be varied by either a V component or a N component. However, they have different effects depending on where the S/C is along the orbit. One interesting observation is that the geometry of the orbit, i.e. its dimension and its shape, is only changed by perturbing accelerations in the orbital plane. The equations for the rate of the inclination and RAAN show that the orbital plane is changed by a perturbing acceleration along W; depending on where the S/C is along the orbit, different effects on the orbital plane change can be obtained. At the ascending node, that is when $\nu = -\omega$, a perturbing acceleration along W varies only the inclination. At the antinodes, namely at $\Delta\nu = \pi/2$ from the nodes, a perturbing acceleration along W varies only the RAAN. Given a certain f_W, the cost of changing Ω is lower if the inclination of the orbit is lower. The argument of periapsis is varied by a change of RAAN and also by in-plane perturbing accelerations. The rate of change of the mean anomaly depends on in-plane perturbing accelerations and on ω and Ω variations.

2.4.2 Perturbations in LEO

In LEO the most important perturbing forces acting on a S/C are due to aerodynamic effects, asphericity of the Earth, solar radiation and electromagnetic effects. In this thesis, only the first two will be accounted for, as the latter are usually quite small and negligible. Indeed, electromagnetic effects have a far bigger impact on the attitude dynamics of a S/C than they do on its trajectory.

2.4.2.1 Atmospheric drag

Since in most cases aerodynamic lift is negligible for satellites, only the effects of drag will be analyzed in this work. The analytic formulation of the perturbing acceleration due to atmospheric drag is not a simple task, as many parameters that define it are afflicted by uncertainties. As a matter of fact, the aerodynamic forces acting on a S/C are dependent on atmospheric fluctuation, the frontal area of the S/C and the drag coefficient. The residual atmosphere in LEO is characterized by a density so low that conventional fluid mechanics is not applicable, as the interaction between the S/C and the atmosphere has to be evaluated at the molecular level. Without losing generality, the perturbing acceleration due to drag is given by

$$\dot{\mathbf{v}} = -\frac{1}{2}\rho\frac{SC_D}{m}\left|\mathbf{v}-\mathbf{v}_{\text{atm}}\right|^2\frac{\mathbf{v}-\mathbf{v}_{\text{atm}}}{\left|\mathbf{v}-\mathbf{v}_{\text{atm}}\right|} \tag{2.113}$$

where C_D is the drag coefficient associated with A, m is the mass of the S/C, $\mathbf{v}_{\text{atm}}$ is the absolute velocity of the atmosphere and ρ is the atmospheric density at the altitude of the S/C. The drag coefficient depends on the type of reflection of the particles that impact the surface of the S/C and typically it's close to the value 2.2. The atmospheric density depends on a vast variety of factors, such as altitude, season, local solar time, solar activity. It's evident that its exact value is never accurately predicted and, as a consequence, the effect of drag can solely be roughly estimated. S is the projection of the area of the satellite perpendicular to the vector $\mathbf{v}-\mathbf{v}_{\text{atm}}$, that is the relative velocity of the S/C with respect to the atmosphere. The frontal area S depends on the attitude of the satellite. In this work, it will also depend on the deployment of a sail. By introducing the vector $\mathbf{v}_{\text{rel}} = \mathbf{v}-\mathbf{v}_{\text{atm}}$ we can rewrite equation (2.113) as

$$\dot{\mathbf{v}} = -\frac{1}{2}\rho\frac{SC_D}{m}\left|\mathbf{v}_{\text{rel}}\right|\mathbf{v}_{\text{rel}} \tag{2.114}$$

2.4.2.2 Asphericity of the Earth

The Earth doesn't have neither geometrical symmetry nor mass symmetry. As a matter of fact, it is an oblate body, flattened at the poles and bulged at the equator. In addition, it is characterized by asymmetrical density variations. Therefore, the expression for the gravitational potential is not quite the one introduced in the two-body problem. The equatorial bulge produces a torque on the satellite and the orbit plane precesses. In addition, the line of the nodes moves westward for direct orbits, that is whose inclination is less than 90°, and westward for retrograde orbits, that is whose inclination is more than 90°. The gravitational potential of the Earth in the WGS84 coordinate system can be expressed by

$$\begin{aligned}\mathcal{E}_g =& \frac{\mu}{r} - \frac{\mu}{r}\sum_{n=2}^{\infty} J_n \left(\frac{r_\oplus}{r}\right)^n \mathcal{P}_n \sin \mathrm{La} \\ & - \frac{\mu}{r}\sum_{n=2}^{\infty}\sum_{m=2}^{\infty} J_{n,m} \left(\frac{r_\oplus}{r}\right)^m P_n^m \sin \mathrm{La} \cos\left[m\left(\mathrm{Lo} - \mathrm{Lo}_{m,n}\right)\right]\end{aligned} \tag{2.115}$$

The function $\mathcal{P}_n(x)$ is the Legendre's polynomial of degree n:

$$\mathcal{P}_n(x) = \frac{1}{2^n n!}\frac{d^n}{dx^n}\left[\left(x^2-1\right)\right] \tag{2.116}$$

whereas $P_n^m(x)$ is the associated Legendre's polynomial of degree n and order m:

$$\mathcal{P}_n(x) = (-1)^m \left(1-x^2\right)^{\frac{m}{2}} \frac{d^m}{dx^m}\left[\mathcal{P}_n(x)\right] \tag{2.117}$$

The terms J_n are called zonal harmonics, whereas $J_{n,m}$ are called tesseral and sectorial harmonics for $n \neq m$ and $n = m$, respectively. The zonal harmonics describe the asymmetry with respect to the equator and depend on latitude. The tesseral and sectorial harmonics describe the asymmetry of the Earth about its axis of rotation. The tesseral harmonics depend on latitude and longitude, while the sectorial harmonics depend on longitude only.

The zonal harmonic J_2, which describes Earth's oblateness, is the most important term and its value ($J_2 = 1.082629 \cdot 10^{-3}$) is three orders of magnitude greater then any other. In addition, LEO orbits aren't affected by tesseral and sectorial harmonics, as their effect is averaged by the rotation of the Earth inside the orbit of the S/C (this isn't the case for satellites in geostationary orbit, which maintain a fixed position relative to the Earth). In practice, for LEO orbit it is sufficient to consider the Earth as axisymmetric and neglect both tesseral and sectorial harmonics. In addition, if only J_2 is considered, the method of general perturbation yields the following orbital elements variations after one revolution:

$$\Delta\Omega = -3\pi J_2 \left(\frac{r_\oplus}{p}\right)^2 \cos\imath \tag{2.118}$$

$$\Delta\omega = \frac{3}{2}\pi J_2 \left(\frac{r_\oplus}{p}\right)^2 \left(5\cos^2\imath - 1\right) \tag{2.119}$$

The regression of the line of the nodes is an important effect on LEO orbits, and its entity increases as the orbit is less inclined. For direct orbits the line of the nodes regresses, whereas for retrograde orbits it precesses. On the other hand, the line of apses precesses for $\imath < 63.4°$ or $\imath > 116.6°$ and regresses for $63.4° < \imath < 116.6°$. For either $\imath = 63.4°$ or $\imath = 116.6°$, that is when $5\cos^2\imath - 1 = 0$, J_2 has no effect on the line of apses.

Chapter 3

Propulsion

Since ADR and OOS missions will require the S/C to perform many LEO transfers, electric propulsion (EP) will be arguably the preferred option. As a matter of fact, although the acceleration levels EP can achieve are far lower than those of chemical propulsion, the propellant consumption of electric thrusters is very limited. Although transfer times will be generally longer than the ones achievable with chemical propulsion, the low propellant consumption will allow the S/C to complete more transfers. This chapter will give the reader an insight into the generalities of space propulsion and will highlight the differences between chemical and electric propulsion.

3.1 Generalities of space propulsion

In space there is almost nothing to exchange momentum with. Therefore, the S/C is an isolated system and the only way to generate thrust consists in having on board something to exchange momentum with, namely the propellant. The source of the energy used to accelerate the propellant defines the main two fields of space propulsion: chemical propulsion and electric propulsion. The energy used to accelerate the propellant comes from a chemical reaction for chemical propulsion, whereas for electric propulsion it comes from a electric power generator.

Let us consider an isolated S/C, far away from any other mass that could generate a gravitational pull. The S/C has a certain velocity v at the time t. Let us study a uni-dimensional case. After an infinitesimal time dt the S/C has ejected a mass of propellant dm_p and changes its velocity to $v + dv$. The mass dm_p has a velocity c with respect to the S/C, thus having a $v - c$ absolute velocity. Let us apply the law of conservation of momentum:

$$mv = (m - dm_p)(v + dv) - dm_p(c - v) \tag{3.1}$$

which rewrites as

$$m dv = dm_p c \tag{3.2}$$

Let us introduce the propellant mass flow rate:

$$\dot{m}_p = \frac{dm_p}{dt} \tag{3.3}$$

The differential dm_p can be rewritten as the product of the mass flow rate and dt, giving

$$m\frac{dv}{dt} = \dot{m}_p c \tag{3.4}$$

This equation gives an expression for the product of the mass of the S/C and its acceleration, that is thrust:

$$T = \dot{m}_p c \tag{3.5}$$

In order to accelerate the propellant from zero velocity to c with respect to the S/C, a certain amount of energy has to be spent. Let us introduce the thrust power:

$$P_T = \frac{1}{2}\dot{m}_p c^2 = \frac{1}{2}Tc \tag{3.6}$$

which is the power needed to continuously accelerate the propellant from 0 to c.

In reality, the propellant that is being ejected interacts with the propellant inside the /C. If the exit velocity of the propellant is u_e, taking into account the momentum exchange with the ejected propellant, the thrust formula rewrites as

$$T = \dot{m}_p u_e + A_e\left(p_e - p_0\right) \tag{3.7}$$

where A_e is the exit area and p_0 is the atmospheric (null if the S/C is in a void). In any case, it is far more convenient to use equation (3.5), compressing in c the effects of accelerating the propellant and the pressure of the ejected propellant. By using (3.5), the detail on the value of the exit velocity and the value of the exit pressure is lost. However, this is not a problem as the detail on how thrust is generated is not of particular interest. As a matter of fact, two thrusters can exert the same thrust on a S/C with different u_e and p_e, but the effects on the S/C are the same. For this reason, it is far more useful to define the ratio between thrust and mass flow rate:

$$c = \frac{T}{\dot{m}_p} \tag{3.8}$$

where c is called *effective exhaust velocity*. If $p_e = p_0$, c is exactly equal to the exhaust velocity. In space propulsion, this is almost always the case (approximately).

3.2 Performance of the propulsion system

The aim of the propulsion system is to change the velocity of the S/C. Thus, the thrust is not itself as much of interest as its effects on the S/C. The integral of thrust over the complete duration of a mission,

$$I_t = \int_{t_0}^{t_f} T dt \tag{3.9}$$

is called *total impulse*. The total mass of propellant used is

$$m_p = \int_{t_0}^{t_f} \dot{m}_p dt \tag{3.10}$$

The ratio between the total impulse and the weight of the mass of the propellant used on the surface of the Earth,

$$I_{sp} = \frac{I_t}{m_p g_0} \tag{3.11}$$

is called *specific impulse*. Historically, the specific impulse is defined with the weight instead of the mass of the propellant. If both thrust and mass flow rate are constant throughout the mission, one has

$$I_{sp} = \frac{I_t}{m_p g_0} = \frac{T\Delta t}{\dot{m}_p \Delta t g_0} = \frac{T}{\dot{m}_p g_0} = \frac{c}{g_0} \tag{3.12}$$

The meanings of I_{sp} and c are fundamentally the same. The specific impulse is a measure of how effectively the propellant mass is ejected. Let us consider given mass of propellant m_p and let's imagine we want to generate thrust with a certain effective exhaust velocity c. Let us also imagine that the generated thrust is equal to the weight of the propellant: $T = m_p g_0$. Let us find how much time the thruster can give us this level of thrust. From the definition of total impulse, we have $\Delta t = I_t/T$. This expression rewrites, using the definition of specific impulse, as $\Delta t = I_{sp} m_p g_0/T$. Since $T = m_p g_0$, we finally have $\Delta t = I_{sp}$. Given m_p, if c increases the same thrust can be exerted for more time. The specific impulse can be considered as the time a certain quantity of propellant can give a thrust equal to its weight on Earth. In practice, if I_{sp} increase, the same thrust can be exerted for more time or more thrust can be exerted for a given time.

Even though the total impulse $I_t = m_p c$ is a measure of the performance of the propulsion system, as it shows how it affects the S/C based on how much propellant is on board and on the value of effective exhaust velocity the system is capable to achieve, it is not enough to have a complete picture of the performance. As a matter of fact, the velocity change of the S/C is the primary aim of propulsion. The same total impulse has different effects on different S/Cs, depending on the mass. The integral of the acceleration due to thrust over the course of a mission,

$$\Delta V = \int_{t_0}^{t_f} \frac{T}{m} dt \tag{3.13}$$

is called *delta-V*. It represents the change in velocity that a S/C would be subjected to if the only force acting on it was thrust parallel to the velocity. The actual change in velocity depends on all of the other forces that interact with the S/C. However, the meaning on ΔV is measuring of the effects of thrust. For example, a body in a gravitational field will have $\Delta V = 0$ in the absence of thrust, even when accelerating or decelerating. The delta-V can be rewritten by using equation (3.8):

$$\Delta V = \int_{t_0}^{t_f} \frac{c\dot{m}_p}{m} dt \tag{3.14}$$

The propellant mass flow rate is equal and opposite to the time derivative of the mass of the S/C:

$$\dot{m}_p = -\frac{dm}{dt} \tag{3.15}$$

Substituting equation (3.15) into (3.14) yields

$$\Delta V = \int_{m_0}^{m_f} -c\frac{dm}{m} \tag{3.16}$$

This expression can be simplified by assuming c is a constant. The resulting equation is the *Tsiolkovsky rocket equation*:

$$\Delta V = c \ln\left(\frac{m_0}{m_f}\right) \tag{3.17}$$

which can be written as

$$m_f = m_0 e^{-\frac{\Delta V}{c}} \tag{3.18}$$

The rocket equation relates the initial and the final mass of a S/C depending on the delta-V and c. Thus, it gives an expression for the propellant needed for a maneuver:

$$m_p = m_0 - m_f = m_0 \left(1 - e^{-\frac{\Delta V}{c}}\right) \tag{3.19}$$

The following table shows the ration between initial and final mass as a function of the ratio between ΔV and c:

$\Delta V/c$	m_f/m_0	m_0/m_f
5	0.0067	148
2	0.135	7.39
1	0.368	2.72
0.5	0.606	1.65
0.2	0.819	1.22
0.1	0.905	1.11

Table 3.1: Results of the rocket equation

Since m_0/m_f is a measure of the cost of the mission (the more it increases the more propellant is needed), given a certain ΔV, c shall be comparable. The bigger the effective exhaust velocity, the lesser propellant is needed for a given mission/maneuver. The typical delta-V for some missions are shown on the following table:

Mission	ΔV [km/s]
LEO insertion	10
Station keeping (1 year)	0.5
LEO-GEO (impulsive)	3.5
LEO-GEO (low-thrust)	6
Earth escape (impulsive)	3.2
Earth escape (low-thrust)	8
Earth-Mars (impulsive)	5.5
Earth-Mars (low-thrust)	6
Earth-Jupiter (low-thrust)	16.7
Earth-Alpha Centauri	30000

Table 3.2: Typical delta-Vs

As a first approximation, given the mass of a S/C, the total impulse solely depends on the delta-V:

$$I_t \approx m_{\text{avg}} \Delta V \tag{3.20}$$

From the definition of I_{sp},

$$m_p = \frac{I_t}{g_0 I_{sp}} \tag{3.21}$$

one notes that, given the mission, the mass of propellant is inversely proportional to the specific impulse of the thruster. Given a certain mission, doubling I_{sp} means completing the mission with half of the propellant. For chemical propulsion, I_{sp} is limited by the maximum energy that the chemical reaction can generate ($I_{sp} = 450\,\text{s}$ with liquid oxygen LOX and liquid hydrogen LH2). This fact has repercussions on the ratio m_0/m_f. For example, insertion in LEO requires $\Delta V = 10\,\text{km/s}$. By using LOX/LH2, which by the way is the best combination in terms of I_{sp}, the ratio $\Delta V/c$

is roughly equal to 2. From table (3.2) we see that the final mass is limited to a 10% of the initial mass. For bigger delta-Vs chemical propulsion is almost always not practical. On the other hand, electric propulsion is not characterized by a direct dependence between the propellant and the energy available. Therefore, bigger c can be easily obtained.

3.3 Chemical propulsion and electric propulsion

In chemical propulsion a propellant mass flow rate $\dot{m}_p$ reacts and releases chemical energy, which is converted to kinetic energy by letting the gas expand and accelerate. The energy balance can be written as

$$\dot{m}_p \left(h_e + \frac{u_e^2}{2} - h_0 \right) = \eta \dot{m}_p E_{ch} \tag{3.22}$$

Where E_{ch} is the energy per unit mass released by the reaction, h_e and h_0 are the specific enthalpy of the propellant when it enters and when it exits the thrust chamber. The efficiency η takes into account how much chemical energy is actually transformed into enthalpy. By neglecting the enthalpy of the propellant when it enters the thrust chamber, assuming maximum efficiency and substituting $c \approx u_e$, we can write

$$c_{\text{chemical}} \approx \sqrt{2E_{ch}} \tag{3.23}$$

Equation (3.23) shows that c is limited by the energy of the reaction and gives an upper limit for c. This is also intuitive, as the energy used to accelerate the propellant is contained in the propellant itself. Therefore, the extent to which it can accelerate depends on the energy the propellant itself is capable of releasing. On the other hand, this fact is what makes chemical rockets extremely scalable, as it's possible to either increase or decrease the propellant mass flow rate for a given value of c. In theory, the thrust achievable can be arbitrarily big, with the drawback of having a large propellant consumption.

As for electric propulsion, the propellant mass flow rate $\dot{m}_p$ receives a certain amount of electric power P_E. The energy balance is:

$$\dot{m}_p \frac{c^2}{2} \approx \eta P_E \tag{3.24}$$

While chemical rockets can achieve high values for η, this is not the case for electric thrusters. As a matter of fact, the transformation from electrical to kinetic energy is characterized by high losses. From equation (3.24) we have

$$c_{\text{electric}} \approx \sqrt{\frac{2\eta P_E}{\dot{m}_p}} \tag{3.25}$$

Theoretically, arbitrary high values of c can be obtained by either increasing P_E or decreasing $\dot{m}_p$.

In a chemical rocket, the energy available $\dot{m}_p E_{ch}$ and requested $\dot{m}_p c^2/2$ are bound by the mass of propellant. If one increases, the other one increases as well. On the other side, an electric thruster doesn't have this limitation, as the available power P_E is completely unrelated to the mass of propellant. Therefore, given a certain power, an electric thruster can accelerate more a smaller mass of propellant, and vice versa. Substituting $T = \dot{m}_p c$ into equation (3.24) yields

$$c \approx \frac{2\eta P_E}{T} \tag{3.26}$$

which shows that c can be increased by either increasing P_E or decreasing T. This is the fundamental difference between chemical and electric propulsion. Chemical rockets can generate arbitrarily high values of thrust (they span from mN to MN range) with a limited c, whereas electric thruster can achieve high values of c with the drawbacks of having on board a power source (solar arrays, typically) and limited thrust (order of N maximum). The available thrust is bounded by technological limits, as the increase of P_E is accompanied by an increase of the mass of the power source. The choice between chemical or electric depends on the mission. Certainly, orbit insertion requires high levels of thrust and chemical is the only option. On the other hand, an interplanetary transfer can be achieved with a lower propellant consumption by choosing electric propulsion, taking into account that the mission will require more time.

3.4 Limits of electric propulsion

Electric propulsion requires a power generator. The total mass m of a S/C can be broken down in various terms:

$$m = m_u + m_p + m_s \tag{3.27}$$

where m_u is the mass of the payload, m_p is the mass of propellant and m_s is the mass of the power generator. The acceleration due to thrust is

$$a = \frac{T}{m} \tag{3.28}$$

which is less than the acceleration the S/C would experience if m_s was the only mass

$$a < \frac{T}{m_s} \tag{3.29}$$

With good approximation, the mass of the power generator can be assumed as proportional to the electric power the generator produces:

$$m_s = \alpha P_E \tag{3.30}$$

Lower value of α correspond to a better technology. Therefore, we can write:

$$a < \frac{T}{\alpha P_E} \tag{3.31}$$

Substituting (3.26) into (3.31) yields

$$a < \frac{T}{\alpha \frac{Tc}{2\eta}} \tag{3.32}$$

which rewrites as

$$a < \frac{\eta}{\alpha} \frac{2}{c} \tag{3.33}$$

The acceleration is limited by the right hand side of the inequality. In order grasp the orders of magnitude one can expect, let us assume very optimistic values for η and α. If $\eta = 0.5$, $\alpha = 1\,\mathrm{kg/kW}$ and $c = 10\,\mathrm{km/s}$ one has

$$a < 0.1\,\frac{\mathrm{m}}{\mathrm{s}^2} \approx \frac{g_0}{100} \tag{3.34}$$

where g_0 is the gravitational acceleration on the surface of the Earth. This thesis will deal with orbital maneuvers in LEO. This means that the acceleration due to thrust will be orders of magnitude less than the one due to gravity. Therefore, we can expect very gradual changes in the trajectory of the S/C.

Chapter 4

Mathematical model

4.1 Atmospheric model

As stated in Section 2.4.2, the atmospheric fluctuations that characterize LEOs are dependent on a vast variety of factors, such as altitude, season, local solar time and solar activity. Since the evaluation of drag depends on the knowledge of the atmospheric density, it is necessary to have an atmospheric model that accurately predicts the density along the orbit of the S/C. Since this thesis is meant to give a methodology in order to take into account the effects of drag on the optimization problem, its aim is not to give results for a specific case study. For this reason, the U.S. Standard Atmosphere 1976 was chosen a suitable atmospheric model. The methodology is not dependent on the atmospheric model, so any model can be readily adopted instead of the one chosen here.

The U.S. Standard Atmosphere is «an idealized, steady-state representation of the earth's atmosphere from the surface to 1000 km, as it is assumed to exist in a period of moderate solar activity». This atmospheric model was adopted by the United States Committee on Extension to the Standard Atmosphere (COESA), a group of organizations which aimed to provide the missile industry a realistic atmospheric model beyond the altitudes of aircraft operations.

The full detail on how the model is built is in [20]. Using the tabulated data, the density profile $\rho(z)$ between altitudes from $z = 86\,\mathrm{km}$ and $z = 1000\,\mathrm{km}$ can be calculated with the basic equation form

$$\rho(z) = \mathrm{e}^{Az^4+Bz^3+Cz^2+Dz+E} \tag{4.1}$$

where the value of the A through E coefficients are listed in table 4.1.

Table 4.1: Density formula coefficients

	Density [kg/m^3]				
Altitude [km]	**A**	**B**	**C**	**D**	**E**
86-91	0.000000	-3.322622×10^{-6}	9.111460×10^{-4}	-0.2609971	5.944694
91-100	0.000000	2.873405×10^{-5}	-0.008492037	0.6541179	-23.62010
100-110	-1.240774×10^{-5}	0.005162063	-0.8048342	55.55996	-1443.338
110-120	0.000000	-8.854164×10^{-5}	0.03373254	-4.390837	176.5294
120-150	3.661771×10^{-7}	-2.154344×10^{-4}	0.04809214	-4.884744	172.3597
150-200	1.906032×10^{-8}	-1.527799×10^{-5}	0.004724294	-0.6992340	20.50921
200-300	1.199282×10^{-9}	-1.451051×10^{-6}	6.910474×10^{-4}	-0.1736220	-5.321644
500-750	8.105631×10^{-12}	-2.358417×10^{-9}	-2.635110×10^{-6}	-0.01562608	-20.02246
750-1000	$-3.701195 \times 10^{-12}$	-8.608611×10^{-9}	5.118829×10^{-5}	-0.06600998	-6.137674

4.2 Analytical description of the perturbations

As already stated, this work takes into account the effects of drag and J2. The trajectories of interest are characterized by long transfer times between almost circular LEOs, so the equations for eccentricity, argument of periapsis and mean anomaly can be neglected. In addition, the eccentricity can be considered always null. Therefore, the radius, the semimajor axis and the semilatus rectum can be considered to be equal and the velocity of the S/C is always equal to the circular velocity. The differential equations that describe the rate of change of semimajor axis, inclination and RAAN are obtained by substituting $e = 0$, $r = a = p$ and $V = \sqrt{\mu/r}$ in equations (2.112):

$$
\begin{aligned}
\dot{a} &= 2\sqrt{\frac{a^3}{\mu}} f_V \\
\dot{\imath} &= \sqrt{\frac{a}{\mu}} \cos\vartheta f_W \\
\dot{\Omega} &= \sqrt{\frac{a}{\mu}} \frac{\sin\vartheta}{\sin\imath} f_W
\end{aligned} \tag{4.2}
$$

In order to apply the variation of elements method and to describe analytically the effects of the perturbations on the orbital elements, let us formulate the expression of the projections of the perturbing accelerations in the coordinate system VNW introduced in Section 2.4.1.1.

First, let us derive the projection of the acceleration due to drag, given by equation (2.114):

$$\dot{\mathbf{v}} = -\frac{1}{2}\rho\frac{SC_D}{m}\left|\mathbf{v}_{\text{rel}}\right|\mathbf{v}_{\text{rel}} \tag{4.3}$$

where $\mathbf{v}_{\text{rel}} = \mathbf{v} - \mathbf{v}_{\text{atm}}$. The velocity of the atmosphere at a given point at a distance r from the center of the Earth can be approximated by

$$\mathbf{v}_{\text{atm}} = \boldsymbol{\omega}_{\oplus} \times \mathbf{r} \tag{4.4}$$

where $\boldsymbol{\omega}_\oplus$ is the angular velocity of the Earth. As a matter of fact, equation (4.4) means that the atmosphere is turning with the Earth with the same angular velocity. Therefore, if the equation holds true, the velocity of the atmosphere is always contained in the local horizontal plane, as it must be perpendicular to $\mathbf{r}$. Since almost circular orbits are considered, the flight path angle of the S/C is always zero and the velocity is also contained in the local horizontal plane as well. As a consequence, at any given point along the orbit the velocity of the atmosphere has a component along V and a component along W, but never has a N component. In addition, since the velocity is perpendicular to the radius vector, the unit vector $\mathbf{V}$ from the VNW system is always coincident to the unit vector $\mathbf{T}$ from the RTW system, as can be easily proved by substituting $e = 0$ in equation (2.111). Therefore, the components of the acceleration due to drag in VNW are

$$(f_V)_{\text{drag}} = (f_T)_{\text{drag}} = \dot{\mathbf{v}} \cdot \mathbf{T} \tag{4.5}$$

$$(f_W)_{\text{drag}} = \dot{\mathbf{v}} \cdot \mathbf{W} \tag{4.6}$$

Substituting equation (2.114) yields

$$(f_V)_{\text{drag}} = -\frac{1}{2}\rho\frac{SC_D}{m}\left|\mathbf{v}_{\text{rel}}\right|(v_{\text{rel}})_T \tag{4.7}$$

$$(f_W)_{\text{drag}} = -\frac{1}{2}\rho\frac{SC_D}{m}\left|\mathbf{v}_{\text{rel}}\right|(v_{\text{rel}})_W \tag{4.8}$$

where $(v_{\text{rel}})_T$ and $(v_{\text{rel}})_W$ are the T and W components of $\mathbf{v}_{\text{rel}}$. In order to write their mathematical expressions, we have to write the components of $\mathbf{v}_{\text{rel}}$ in the perifocal frame and then transform the coordinates to RTW.

The angular velocity of the Earth only has a K component in the geocentric-equatorial coordinate system:

$$\boldsymbol{\omega}_\oplus = 0\mathbf{I} + 0\mathbf{J} + \omega_\oplus\mathbf{K} \tag{4.9}$$

Let us transform the coordinates to the perifocal frame. Since the orbit is circular, let us use the arguments of latitude from equation (2.107) instead of the true anomaly. This way, we can arbitrarily set ω to zero in the transformation matrix (B.5). Transforming the coordinates according to equation (B.6) yields

$$\boldsymbol{\omega}_\oplus = 0\mathbf{p} + \omega_\oplus \sin\imath\mathbf{q} + \omega_\oplus\cos\imath\mathbf{w} \tag{4.10}$$

Let us now calculate the cross product

$$(\mathbf{v}_{\text{atm}})_{\text{pqw}} = (0\mathbf{p} + \omega_\oplus \sin\imath\mathbf{q} + \omega_\oplus\cos\imath\mathbf{w}) \times (r_p\mathbf{p} + r_q\mathbf{q} + 0\mathbf{w}) \tag{4.11}$$

which gives

$$(\mathbf{v}_{\text{atm}})_{\text{pqw}} = -r_q\omega_\oplus \cos\imath\mathbf{p} + r_p\omega_\oplus \cos\imath\mathbf{q} - r_p\omega_\oplus \sin\imath\mathbf{w} \tag{4.12}$$

The p and q components of the radius are given by

$$r_p = r\cos\vartheta \tag{4.13}$$

$$r_p = r\sin\vartheta \tag{4.14}$$

where ϑ is the argument of latitude. We can now use the transformation matrix from equation (B.8) to finally obtain

$$(\mathbf{v}_{\text{atm}})_{\text{RTW}} = 0\mathbf{R} + r\omega_\oplus \cos\imath\mathbf{T} - r\omega_\oplus \sin\imath\cos\vartheta\mathbf{W} \tag{4.15}$$

The velocity of the S/C in the RTW frame is

$$(\mathbf{v})_{\text{RTW}} = 0\mathbf{R} + \sqrt{\frac{\mu}{r}}\mathbf{T} - 0\mathbf{W} \tag{4.16}$$

We can now write, substituting $r = a$, the expressions for $(v_{\text{rel}})_T$, $(v_{\text{rel}})_W$ and $|\mathbf{v}_{\text{rel}}|$:

$$(v_{\text{rel}})_T = \sqrt{\frac{\mu}{a}} - \omega_\oplus a\cos\imath \tag{4.17}$$

$$(v_{\text{rel}})_W = \omega_\oplus a\sin\imath\cos\vartheta \tag{4.18}$$

$$|\mathbf{v}_{\text{rel}}| = \sqrt{\left(\sqrt{\frac{\mu}{a}} - \omega_\oplus a\cos\imath\right)^2 + (\omega_\oplus a\sin\imath\cos\vartheta)^2} \tag{4.19}$$

Finally, substituting (4.17), (4.18) and (4.19) in equations (4.7) and (4.8), we have the expressions for the accelerations components due to drag:

$$(f_V)_{\text{drag}} = -\frac{1}{2}\rho\frac{SC_D}{m}\left(\sqrt{\frac{\mu}{a}} - \omega_\oplus a\cos\imath\right)\sqrt{\left(\sqrt{\frac{\mu}{a}} - \omega_\oplus a\cos\imath\right)^2 + (\omega_\oplus a\sin\imath\cos\vartheta)^2} \tag{4.20}$$

$$(f_W)_{\text{drag}} = -\frac{1}{2}\rho\frac{SC_D}{m}\left(\omega_\oplus a\sin\imath\cos\vartheta\right)\sqrt{\left(\sqrt{\frac{\mu}{a}} - \omega_\oplus a\cos\imath\right)^2 + (\omega_\oplus a\sin\imath\cos\vartheta)^2} \tag{4.21}$$

The perturbing drag acceleration is responsible for time variations of a, $\imath$ and Ω. In order to study the extent of these variations, let us analyze

$$\frac{\dot{a}_{\text{drag}}}{\dot{\imath}_{\text{drag}}} = 2\frac{a}{\cos\vartheta}\frac{(f_V)_{\text{drag}}}{(f_W)_{\text{drag}}} \tag{4.22}$$

$$\frac{\dot{a}_{\text{drag}}}{\dot{\Omega}_{\text{drag}}} = 2a\frac{\sin\imath}{\sin\vartheta}\frac{(f_V)_{\text{drag}}}{(f_W)_{\text{drag}}} \tag{4.23}$$

Substituting equations (4.20) and (4.21) in (4.22) and (4.23) yields

$$\frac{\dot{a}_{\text{drag}}}{\dot{\imath}_{\text{drag}}} = \frac{2}{\cos\vartheta}\frac{\sqrt{\frac{\mu}{a}} - \omega_{\oplus} a\cos\imath}{\omega_{\oplus}\sin\imath\cos\vartheta} \tag{4.24}$$

$$\frac{\dot{a}_{\text{drag}}}{\dot{\Omega}_{\text{drag}}} = \frac{2}{\sin\vartheta}\frac{\sqrt{\frac{\mu}{a}} - \omega_{\oplus} a\cos\imath}{\omega_{\oplus}\cos\vartheta} \tag{4.25}$$

We see that the the numerator of these equations is always greater than $2\left(\sqrt{\mu/a} - \omega_{\oplus} a\right)$, while the denominator is always less than $\omega_{\oplus}$. Therefore, we can write the following inequalities:

$$\frac{\dot{a}_{\text{drag}}}{\dot{\imath}_{\text{drag}}} \geq \frac{2\left(\sqrt{\frac{\mu}{a}} - \omega_{\oplus} a\right)}{\omega_{\oplus}} \tag{4.26}$$

$$\frac{\dot{a}_{\text{drag}}}{\dot{\Omega}_{\text{drag}}} \geq \frac{2\left(\sqrt{\frac{\mu}{a}} - \omega_{\oplus} a\right)}{\omega_{\oplus}} \tag{4.27}$$

It is easy to demonstrate that up until 1000 km altitude, above which the effects of drag are negligible, the numerator is a monotonically decreasing function of the altitude. Substituting $\omega_{\oplus} = 7.2921 \cdot 10^{-5}$ rad/s, for a 1000 km altitude we can write

$$\frac{\dot{a}_{\text{drag}}}{\dot{\imath}_{\text{drag}}} > 10^5\,\frac{\text{km/s}}{\text{rad/s}} \tag{4.28}$$

$$\frac{\dot{a}_{\text{drag}}}{\dot{\Omega}_{\text{drag}}} > 10^5\,\frac{\text{km/s}}{\text{rad/s}} \tag{4.29}$$

The effects of drag always have a far greater impact on the rate of change of the semimajor axis than they do on the rate of change of the inclination and RAAN.

Among the effects caused by the zonal harmonic J2, only the secular variation of the RAAN is accounted for in this work, as the short period variations are neglected. The time derivative of Ω is found by dividing the RAAN change over one revolution (equation (2.118)) by the orbital period (equation 2.85):

$$\dot{\Omega}_{J2} = -\frac{3}{2}J_2\left(\frac{r_E}{a}\right)^2\sqrt{\frac{\mu}{a^3}}\cos\imath \tag{4.30}$$

where r_E is the radius of the Earth.

By introducing the angle β between the orbital plane and the thrust vector, the acceleration components due to thrust are

$$(f_V)_{\text{thrust}} = \frac{T}{m}\cos\beta \tag{4.31}$$

$$(f_W)_{\text{thrust}} = \frac{T}{m}\sin\beta \tag{4.32}$$

We can finally write the time differential equations that mathematically describe the rate of change of the orbital elements:

$$\begin{aligned}
\frac{da}{dt} &= 2\sqrt{\frac{a^3}{\mu}}\left[\frac{T}{m}\cos\beta - (f_V)_{\text{drag}}\right] \\
\frac{d\imath}{dt} &= \sqrt{\frac{a}{\mu}}\cos\vartheta\left[\frac{T}{m}\sin\beta - (f_W)_{\text{drag}}\right] \\
\frac{d\Omega}{dt} &= \sqrt{\frac{a}{\mu}}\frac{\sin\vartheta}{\sin\imath}\left[\frac{T}{m}\sin\beta - (f_W)_{\text{drag}}\right] - \frac{3}{2}J_2\left(\frac{r_E}{a}\right)^2\sqrt{\frac{\mu}{a^3}}\cos\imath
\end{aligned} \tag{4.33}$$

Chapter 5

Indirect optimization of space trajectories

The aim of optimal control theory (OCT) is to determine a control law that will cause process to satisfy physical constraints while maximizing (or minimizing) a performance index. The formulation of an optimal control problem requires a mathematical model of the process, a mathematical description of the physical constraints and a definition of the performance index. The mathematical model of the process, already derived in Section 4.2, describes the response of the physical system to inputs. The physical system is associated with a state vector $\mathbf{x}$ and a set of differential equations, functions of $\mathbf{x}$, the control vector $\mathbf{u}$ and time. In our case the state vector is given by the semimajor axis, the inclination of the orbit and the RAAN, as they describe the physical state of the system. On the other hand, the controls are frontal area of the S/C and thrust magnitude and direction. This chapter shows how to apply OCT to the problem of optimal spacecraft trajectories, and how to improve numerical accuracy as described in [6].

5.1 Optimal control theory

Let us use the notation of Appendix A. The physical system is then described by a set of state differential equation

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}\left(\mathbf{x}, \mathbf{u}, t\right) \tag{5.1}$$

which rule the evolution from the initial to the final state. If variables exhibit discontinuity, it is convenient to split the trajectory between the initial and final point (external boundaries) into n arcs at the points where the state or control variables are discontinuous (internal boundaries). Arc j starts at time $t_{(j-1)_+}$ and end at time t_{j_-}. At those instants the state vector is $\mathbf{x}_{(j-1)_+}$ and $\mathbf{x}_{j_-}$, respectively. The signs -

and + denote the values of the variables immediately before or after a certain time instant. For example, possible discontinuities of variables in space trajectories may be modeled as instantaneous changes in velocity in the case of impulsive maneuvers. Moreover, time discontinuity may also arise in the case of planetary flybys, if the time inside the sphere of influence of a planet is not neglected. Without losing generality, the boundary conditions can be grouped into a vector $\boldsymbol{\chi}$ and written as

$$\boldsymbol{\chi}\left(\mathbf{x}_{(j-1)_+}, \mathbf{x}_{j_-}, t_{(j-1)_+}, t_{j_-}\right) = \mathbf{0} \tag{5.2}$$

for $j = 1, \ldots, n$. The optimal control problem is formulated as the minimization or maximization of a functional of generic form

$$J = \varphi\left(\mathbf{x}_{(j-1)_+}, \mathbf{x}_{j_-}, t_{(j-1)_+}, t_{j_-}\right) + \sum_j \int_{t_{(j-1)_+}}^{t_{j_-}} \Phi\left(\mathbf{x}\left(t\right), \mathbf{u}\left(t\right), t\right) dt \tag{5.3}$$

for $j = 1, \ldots, n$. The term φ depends on the values that the variables and time assume at the internal boundaries. On the other hand, the integral of Φ depends on the values that the variables and time assume at each point during the trajectory. The function φ is known as the Mayer term and identifies the cost related to the state of the system at the boundaries, whereas the function Φ is known as the Lagrange term and keeps track if the state and control costs during the entire time history. The functional J may be defined with just the Mayer term, just the Lagrange term or both. If $\Phi = 0$ one has the Mayer formulation, whereas if $\varphi = 0$ one has the Lagrange formulation. The Mayer formulation is here preferred. However, the two formulations are equivalent, and one can be obtained from the other. A necessary condition for optimality requires that the first variation of J is null for any variation $\delta\mathbf{x}$ and $\delta\mathbf{u}$ along the trajectory, and for any variation $\delta\mathbf{x}_{(j-1)_+}$, $\delta\mathbf{x}_{j_-}$, $\delta t_{(j-1)_+}$ and δt_{j_-} at the boundaries.

The functional J is rewritten in the Mayer formulation by introducing the Lagrange multipliers $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$:

$$J^* = \varphi + \boldsymbol{\mu}^T\boldsymbol{\chi} + \sum_j \int_{t_{(j-1)_+}}^{t_{j_-}} \left(\boldsymbol{\lambda}^T\left(\mathbf{f} - \frac{d\mathbf{x}}{dt}\right)\right) dt \qquad j = 1, \ldots, n \tag{5.4}$$

where the constants $\boldsymbol{\mu}$ are associated with the boundary conditions and the adjoint variables $\boldsymbol{\lambda}$ are associated with the state equations. The functionals J and J^* depend on the state variables $\mathbf{x}$, their time derivatives, the controls $\mathbf{u}$ and the values that the variables and time assume at the boundaries. From equations (5.3) and (5.4) it can be easily seen that J and J^* coincide if both the boundary conditions(5.2) and the state equations (5.1) are satisfied. Integrating by parts the expression for J^* yields

$$
\begin{aligned}
J^* = \varphi + \boldsymbol{\mu}^T \boldsymbol{\chi} + \sum_j \left(\boldsymbol{\lambda}^T_{(j-1)_+} \mathbf{x}_{(j-1)_+} - \boldsymbol{\lambda}^T_{j_-} \mathbf{x}_{j_-} \right) \\
+ \sum_j \int_{t_{(j-1)_+}}^{t_{j_-}} \left(\boldsymbol{\lambda}^T \mathbf{f} - \frac{d\boldsymbol{\lambda}^T}{dt} \mathbf{x} \right) dt \qquad j = 1, \dots, n
\end{aligned}
\tag{5.5}
$$

The first variation δJ^* can be obtained by differentiating equation (5.5):

$$
\begin{aligned}
\delta J^* \quad = \quad & \left(-H_{(j-1)_+} + \frac{\partial \varphi}{\partial t_{(j-1)_+}} + \boldsymbol{\mu}^T \frac{d\boldsymbol{\chi}}{dt_{(j-1)_+}} \right) \delta t_{(j-1)_+} + \\
& + \left(H_{j_-} + \frac{\partial \varphi}{\partial t_{j_-}} + \boldsymbol{\mu}^T \frac{d\boldsymbol{\chi}}{dt_{j_-}} \right) \delta t_{j_-} + \\
& + \left(\boldsymbol{\lambda}^T_{(j-1)_+} + \frac{\partial \varphi}{\partial \mathbf{x}_{(j-1)_+}} + \boldsymbol{\mu}^T \left[\frac{\partial \boldsymbol{\chi}}{\partial \mathbf{x}_{(j-1)_+}} \right] \right) \delta \mathbf{x}_{(j-1)_+} + \\
& + \left(\boldsymbol{\lambda}^T_{j_-} + \frac{\partial \varphi}{\partial \mathbf{x}_{j_-}} + \boldsymbol{\mu}^T \left[\frac{\partial \boldsymbol{\chi}}{\partial \mathbf{x}_{j_-}} \right] \right) \delta \mathbf{x}_{j_-} + \\
& + \sum_j \int_{t_{(j-1)_+}}^{t_{j_-}} \left(\left(\frac{dH}{d\mathbf{x}} + \frac{d\boldsymbol{\lambda}^T}{dt} \right) \delta \mathbf{x} + \frac{\partial H}{\partial \mathbf{u}} \delta \mathbf{u} \right) dt \qquad j = 1, \dots, n
\end{aligned}
\tag{5.6}
$$

where the Hamiltonian has been introduced:

$$
H = \boldsymbol{\lambda}^T \mathbf{f} \tag{5.7}
$$

The necessary condition for optimality requires $\delta J^* = 0$ for any admissible variation along the trajectory. By appropriately choosing the value of the adjoint constants and variables one can nullify the coefficients of any of the variations in equation (5.6). The Euler-Lagrange equations for the adjoint variables

$$
\frac{d\boldsymbol{\lambda}}{dt} = - \left(\frac{\partial H}{\partial \mathbf{x}} \right)^T \tag{5.8}
$$

and the algebraic equations for the control variables

$$
\left(\frac{\partial H}{\partial \mathbf{u}} \right)^T = 0 \tag{5.9}
$$

can be obtained by nullifying the coefficients of the variations $\delta\mathbf{x}$ and $\delta\mathbf{u}$. As one can observe from equation (5.9), the control laws are formally independent from the formulation of the problem, namely from whether the performance index has to be maximized or minimized. In addition, if a control variable is subject to constraints, the equation may not provide the optimal controls. This may be the case when the thrust magnitude can vary between a minimum or a maximum value. Therefore, it would make no sense to look for a solution that requires a thrust level outside the constraints. In our case, the frontal area is also bounded by such maximum and minimum value constraints. An admissible control is defined as a control that does not violate any constraint and the subset of admissible controls is called admissible region. The Pontryagin's maximum principle [3] states that the optimal control must maximize the Hamiltonian (if J^* is to be maximized). In practice two possibilities may occur:

- the optimal control is given by equation (5.9) if the control is in the admissible region;
- the optimal control is at the boundary of the admissible region, that is the control assumes its maximum or minimum value if equation (5.9) yields a control outside of the admissible region.

In particular, if the Hamiltonian is linear with respect to a control variable, two other possibilities mat occur:

- if the coefficient of the control in equation (5.7) is not null, H is maximized either for the maximum value of the control if the coefficient is positive or for the minimum value of the control if the coefficient is negative, in agreement with the Pontryagin's maximum principle;
- if the coefficient of the control in equation (5.7) is null for a finite interval of time, a *singular arc* arises and it is necessary to set all the successive time derivatives of the coefficient equal to zero, until one of the control appears specifically in one of them; the optimal control is determined by setting such time derivative equal to zero.

Finally, the boundary condition for optimality are determined by nullifying the coefficients of the other variations:

$$-\boldsymbol{\lambda}_{j_-}^T + \frac{\partial \varphi}{\partial \mathbf{x}_{j_-}} + \boldsymbol{\mu}^T \left[\frac{\partial \boldsymbol{\chi}}{\partial \mathbf{x}_{j_-}}\right] = 0 \qquad j = 1, \ldots, n \tag{5.10}$$

$$\boldsymbol{\lambda}_{j_+}^T + \frac{\partial \varphi}{\partial \mathbf{x}_{j_+}} + \boldsymbol{\mu}^T \left[\frac{\partial \boldsymbol{\chi}}{\partial \mathbf{x}_{j_+}}\right] = 0 \qquad j = 0, \ldots, n-1 \tag{5.11}$$

$$H_{j_-} + \frac{\partial \varphi}{\partial t_{j_-}} + \boldsymbol{\mu}^T \frac{d\boldsymbol{\chi}}{dt_{j_-}} = 0 \qquad j = 1, \ldots, n \tag{5.12}$$

$$-H_{j_-} + \frac{\partial \varphi}{\partial t_{j_-}} + \boldsymbol{\mu}^T \frac{d\boldsymbol{\chi}}{dt_{j_-}} = 0 \qquad j = 0, \ldots, n-1 \tag{5.13}$$

Equations (5.10) and (5.12) have no meaning at the starting time, while equations (5.11) and (5.13) have no meaning at the final time. If the generic state variable x is characterized by particular boundary condition, equations (5.10) and (5.11) yield particular boundary conditions for optimality for the corresponding adjoint variable λ_x:

- if the value of x is given at the starting time ($\boldsymbol{\chi} = \mathbf{0}$ contains the equation $x_0 - a = 0$ with a given value for a), the corresponding adjoint variable λ_{x0} is free, that is it does not appear as a boundary condition for optimality and it can assume any value; the same happens for a given value for x at the final time;

- if the initial value x_0 appears in neither the function φ nor in the boundary condition, the corresponding adjoint variable is null at the initial time ($\lambda_{x0} = 0$); the same happens for the analogous situation at the final time;

- if the state variable is continuous and its value is not explicitly set to a value at the internal boundary j ($\boldsymbol{\chi} = \mathbf{0}$ contains the equation $x_{j_+} - x_{j_-} = 0$), the corresponding adjoint variable is continuous ($\lambda_{x_{j+}} = \lambda_{x_{j-}}$);

- if the state variable is continuous and its value is explicitly set to a value at the internal boundary j ($\boldsymbol{\chi} = \mathbf{0}$ contains the equations $x_{j_+} - a = 0$ and $x_{j_-} - a = 0$), the corresponding adjoint variable has a free discontinuity, that is the value of $\lambda_{x_{j+}}$ is independent from that of $\lambda_{x_{j-}}$ and it has to be determined by the optimization procedure.

Analogously, if H does is not an explicit function of time, in some cases equations (5.12) and (5.13) yield particular boundary conditions for optimality:

- if the initial time t_0 appears explicitly in neither the boundary conditions nor the function φ, the Hamiltonian is null at the initial time ($H_0 = 0$); analogously, the Hamiltonian is null at the final time if t_f appears explicitly in neither $\boldsymbol{\chi}$ nor φ;

- if the intermediate time t_j does not explicitly appear in the function φ (it appears only in the boundary condition for the time continuity $t_{j_+} = t_{j_-}$), the Hamiltonian is continuous at the internal boundary j ($H_{j_-} = H_{j+}$);

- if the intermediate time t_j is explicitly assigned (it appears in the boundary conditions as $t_{j_+} - a = 0$ and $t_{j_-} - a = 0$), the Hamiltonian has a free discontinuity at the internal boundary j.

By canceling out the adjoint constants $\boldsymbol{\mu}$ from equations (5.10)-(5.13), the resulting boundary conditions for optimality and the boundary conditions on the state variables given by equation (5.2) can be collected in a vector:

$$\boldsymbol{\sigma}\left(\mathbf{x}_{(j-1)_+}, \mathbf{x}_{j_-}, \boldsymbol{\lambda}_{(j-1)_+}, \boldsymbol{\lambda}_{j_-}, t_{(j-1)_+}, t_{j_-}\right) = \mathbf{0} \tag{5.14}$$

Therefore, equations (5.1), (5.8), (5.9) and (5.14) define a multi-point boundary value problem (MPBVP).

5.2 Boundary value problem

The application of the theory of optimal control to the system (5.1) generally yields a MPBVP (in the case of one interval of integration a two-point boundary value problem). Equations (5.1) and (5.8) are the differential equations of the MPBVP and the controls are determined by equation (5.9). The solution to this problem is obtained by searching for the initial values of the unknown variables such that the integration of the differential equations satisfies the boundary conditions of equation (5.14). In particular, the interval of integration is split in sub-intervals and different sub-intervals can be characterized by different differential equations. Generally, the duration of each sub-interval is unknown and the boundary conditions may be non-linear. In addition, variables may be discontinuous at the internal boundaries and their values may be unknown after a discontinuity.

In order to deal with the indetermination of the duration of the sub-intervals of integration, a change of independent variable is introduced and, for each sub-interval j, time is replaced with

$$\varepsilon = j - 1 + \frac{t - t_{j-1}}{\tau_j} \tag{5.15}$$

where $\tau_j = t_j - t_{j-1}$ is the duration of the sub-interval. This way, the extremes of the integration sub-intervals are fixed and correspond to consecutive integer values of the new independent variable ε at the boundaries.

The description of the *shooting method* for the solution of the MPBVP is given by referring to the generic system

$$\frac{d\mathbf{y}}{dt} = \mathbf{f}^* (\mathbf{y}, t) \tag{5.16}$$

where the state variables and the adjoint variables are grouped in the vector $\mathbf{y} = (\mathbf{x}, \boldsymbol{\lambda})$ and assuming the substitution of the controls with the expressions defined by equation (5.9). Since the problem is also defined by constant parameters, such as the duration τ_j of the sub-intervals and the values of the variables after a discontinuity, it is convenient to introduce a new vector $\mathbf{z} = (\mathbf{y}, \mathbf{c})$ that groups together the state variable, the adjoint variables and the vector $\mathbf{c}$ of the constant parameters. The replacement of time with the new independent variable yields:

$$\frac{d\mathbf{z}}{d\varepsilon} = \mathbf{f} (\mathbf{z}, \varepsilon) \tag{5.17}$$

which can be explicitly written as

$$\frac{d\mathbf{z}}{d\varepsilon} = \left(\frac{d\mathbf{y}}{d\varepsilon}, \frac{d\mathbf{c}}{d\varepsilon}\right) \tag{5.18}$$

where

$$\frac{d\mathbf{y}}{d\varepsilon} = \tau_j \frac{d\mathbf{y}}{dt} \tag{5.19}$$

and

$$\frac{d\mathbf{c}}{d\varepsilon} = \mathbf{0} \tag{5.20}$$

The boundary conditions are generically expressed as

$$\mathbf{\Psi}\left(\mathbf{s}\right) = \mathbf{0} \tag{5.21}$$

where $\mathbf{s}$ is a vector that contains the values that the variables assume at the internal and external boundaries, as well as the unknown parameters:

$$\mathbf{s} = \left(\mathbf{y}_0, \mathbf{y}_1, \ldots, \mathbf{y}_n, \mathbf{c}\right) \tag{5.22}$$

Some of the initial values of the variables are unknown. As already stated, the solution to the MPBVP is obtained with an iterative numerical method that searches for the initial values such that the integration of the differential equations satisfies equation (5.21). The method is here described by assuming none of the initial values is known. The iteration r is started by integrating equations (5.17) with the initial values $\mathbf{p}^r$ yielded by the previous iteration. Namely:

$$\mathbf{z}\left(0\right) = \mathbf{p}^r \tag{5.23}$$

The integration is carried out for the entire trajectory (if r is the first iteration, tentative values for $\mathbf{p}^r$ are chosen). The values of the state variables are determined at each boundary and the errors on the boundary conditions are grouped in the vector $\mathbf{\Psi}^r$.

A variation $\Delta\mathbf{p}$ results in a variation of the errors on the boundary conditions equal to

$$\Delta\mathbf{\Psi} = \left[\frac{\partial\mathbf{\Psi}}{\partial\mathbf{p}}\right]\Delta\mathbf{p} \tag{5.24}$$

where higher than first order terms are neglected. Since the objective is bringing the errors to zero, the goal for each next iteration is achieving $\Delta\mathbf{\Psi} = -\mathbf{\Psi}^r$. In light of this observation, the initial values are corrected by a quantity equal to

$$\Delta\mathbf{p} = -\left[\frac{\partial\mathbf{\Psi}}{\partial\mathbf{p}}\right]^{-1}\mathbf{\Psi}^r \tag{5.25}$$

and iteration $r+1$ is started by integrating the differential equations with initial values

$$\mathbf{p}^{r+1} = \mathbf{p}^{r} + \Delta\mathbf{p} \tag{5.26}$$

The iterations are performed until the boundary conditions are satisfied with the wanted precision. The matrix from equation (5.25) can be calculated as

$$\left[\frac{\partial \boldsymbol{\Psi}}{\partial \mathbf{p}}\right] = \left[\frac{\partial \boldsymbol{\Psi}}{\partial \mathbf{s}}\right]\left[\frac{\partial \mathbf{s}}{\partial \mathbf{p}}\right] \tag{5.27}$$

The error gradient with respect to $\mathbf{s}$ is easily obtained by analytical derivation. On the other hand, the derivative of vector $\mathbf{s}$ with respect to vector $\mathbf{p}$ yields a matrix that contains the values assumed by matrix

$$[\mathbf{g}(\varepsilon)] = \left[\frac{\partial \mathbf{z}}{\partial \mathbf{p}}\right] \tag{5.28}$$

at the boundaries ($\varepsilon = 0, 1, \ldots, n$). Since taking the derivative of the system (5.17) with respect to the initial values yields

$$\left[\frac{\partial \mathbf{f}}{\partial \mathbf{p}}\right] = \left[\frac{d}{d\mathbf{p}}\left(\frac{\partial \mathbf{z}}{\partial \varepsilon}\right)\right] = \frac{d}{d\varepsilon}\left[\frac{\partial \mathbf{z}}{\partial \mathbf{p}}\right] = \left[\frac{d\mathbf{g}(\varepsilon)}{d\varepsilon}\right] \tag{5.29}$$

the matrix $[\partial\mathbf{s}/\partial\mathbf{p}]$ can be obtained by integrating

$$\left[\frac{d\mathbf{g}(\varepsilon)}{d\varepsilon}\right] = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{p}}\right] = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{z}}\right]\left[\frac{\partial \mathbf{z}}{\partial \mathbf{p}}\right] = \left[\frac{\partial \mathbf{f}}{\partial \mathbf{z}}\right][\mathbf{g}(\varepsilon)] \tag{5.30}$$

where the Jacobian matrix $[\partial\mathbf{f}/\partial\mathbf{z}]$ is obtained by analytical derivation. The initial values for the homogeneous system (5.30) are easily obtained by taking the derivative of (5.23) with respect to $\mathbf{p}$, and thus obtaining the identity matrix:

$$[\mathbf{g}(\varepsilon)] = \left[\frac{\partial \mathbf{z}(0)}{\partial \mathbf{p}}\right] = [\mathbf{I}] \tag{5.31}$$

This method allows to deal with variable discontinuities. As a matter of fact, if a discontinuity occurs at boundary j, the same relation $\mathbf{h}$ between the values of the variables before and after the discontinuity,

$$\mathbf{z}_{j_+} = \mathbf{h}\left(\mathbf{z}_{j_-}\right) \tag{5.32}$$

can be applied to matrix $\mathbf{g}$:

$$[\mathbf{g}_{j_+}] = \left[\frac{\partial \mathbf{h}}{\partial \mathbf{z}}\right][\mathbf{g}_{j_-}] \tag{5.33}$$

If some of the initial values of the variables are known, the same method can be applied and the vector $\mathbf{p}$ contains the unknown initial values of the variables, while vector $\mathbf{\Psi}$ only contains the boundary conditions that are not explicit at the initial time.

Since the procedure described above for the determination of the error gradient matrix $[\partial\mathbf{\Psi}/\partial\mathbf{p}]$ takes a heavy analytical effort and the computational time is rather long, another method can be easily adopted. As a matter of fact, the error gradient matrix can be evaluated numerically: row i of the matrix is obtained by perturbing the ith component of $\mathbf{p}$ by a small variation Δp and by then integrating the equations (5.17). By doing so for each component of $\mathbf{p}$ it is possible to evaluate the variation of the errors $\Delta\mathbf{\Psi}\,(\Delta p)$ and the corresponding ith row $\Delta\mathbf{\Psi}/\Delta p$. Empirical values for Δp are in the order of $10^{-6} \div 10^{-7}$. Although this other method is rather faster than the one described earlier, it may fail to provide convergence. Its implementation is therefore dependent on the complexity of the problem.

5.3 Approximate optimal LEO transfers between almost circular orbits

This work applies the procedure described above on a chaser S/C with a low-thrust electric propulsion system. The S/C is on a particular almost circular LEO and the optimization problem is formulated as the search for the optimal controls to achieve rendezvous with a target S/C in another circular LEO either in minimum time or, given a certain transfer time, with minimum propellant expenditure. In addition, the chaser is equipped with a drag sail; thus, the maximum frontal area is a control variable. As already mentioned in Section 4.2, since the transfer times are long and the orbits are almost circular, the equations for eccentricity and argument of periapsis can be neglected. In addition, the rendezvous maneuvers are not treated in this work and thus the true anomaly equation is neglected. Therefore, the state of the system is described by the semimajor axis, the inclination and the RAAN of the S/C that has to perform the transfer. The transfer is completed once the S/C has the same a, $\imath$ and Ω of the target. The differential equations that describe the physical system are given by equations 4.33:

$$\begin{aligned}
\frac{da}{dt} &= 2\sqrt{\frac{a^3}{\mu}}\left[\frac{T}{m}\cos\beta - (f_V)_{\text{drag}}\right] \\
\frac{d\imath}{dt} &= \sqrt{\frac{a}{\mu}}\cos\vartheta\left[\frac{T}{m}\sin\beta - (f_W)_{\text{drag}}\right] \\
\frac{d\Omega}{dt} &= \sqrt{\frac{a}{\mu}}\frac{\sin\vartheta}{\sin\imath}\left[\frac{T}{m}\sin\beta - (f_W)_{\text{drag}}\right] - \frac{3}{2}J_2\left(\frac{r_E}{a}\right)^2\sqrt{\frac{\mu}{a^3}}\cos\imath
\end{aligned} \tag{5.34}$$

A OCT approach is used to first determine the generic control law for a one-revolution transfer. The ratios of the changes of the orbital elements over one revolution to the time required to complete it are then used to approximate the time derivative of the orbital elements for the multiple-revolution transfer. The OCT approach is then repeated with the new differential equations.

5.3.1 One-revolution transfer

By using the argument of latitude as the independent variable, one has

$$\frac{d}{d\vartheta} = \frac{d}{dt}\frac{dt}{d\vartheta} = \sqrt{\frac{a^3}{\mu}}\frac{d}{dt} \tag{5.35}$$

and equations (5.34) become

$$\begin{aligned}
\frac{da}{d\vartheta} &= 2\frac{a^3}{\mu}\left[\frac{T}{m}\cos\beta - \frac{\rho}{2}\frac{SC_D}{m}\left(\sqrt{\frac{\mu}{a}} - \omega_\oplus a\cos\imath\right)v_{\mathrm{rel}}\right] \\
\frac{d\imath}{d\vartheta} &= \frac{a^2}{\mu}\cos\vartheta\left[\frac{T}{m}\sin\beta - \frac{\rho}{2}\frac{SC_D}{m}\left(\omega_\oplus a\sin\imath\cos\vartheta\right)v_{\mathrm{rel}}\right] \\
\frac{d\Omega}{d\vartheta} &= \frac{a^2}{\mu}\frac{\sin\vartheta}{\sin\imath}\left[\frac{T}{m}\sin\beta - \frac{\rho}{2}\frac{SC_D}{m}\left(\omega_\oplus a\sin\imath\cos\vartheta\right)v_{\mathrm{rel}}\right] - \frac{3}{2}J_2\left(\frac{r_E}{a}\right)^2\cos\imath
\end{aligned} \tag{5.36}$$

where $(f_V)_{\mathrm{drag}}$ and $(f_W)_{\mathrm{drag}}$ are substituted by equations (4.20) and (4.21), and the expression for v_{rel} is given by equation (4.19). Since low-thrust electric propulsion is assumed to be employed, the mass of the S/C can be treated as a constant during a single revolution. In addition, since thrust, drag and the J2 perturbation are small, semimajor axis, inclination and atmospheric density can be treated as constants in the right-hand side of equations (5.36).

In order to simplify the equation and obtain analytically integrable equations, the expression for v_{rel}

$$v_{\mathrm{rel}} = \sqrt{\left(\sqrt{\frac{\mu}{a}} - \omega_\oplus a\cos\imath\right)^2 + \left(\omega_\oplus a\sin\imath\cos\vartheta\right)^2} \tag{5.37}$$

has to be simplified. Since over one revolution a and $\imath$ are treated as constants, v_{rel} is only a function of ϑ in the form

$$v_{\mathrm{rel}} = \sqrt{x + \cos^2\vartheta} \tag{5.38}$$

which is a periodic function that oscillates between the values

$$v_{\mathrm{min}} = \left|\sqrt{\frac{\mu}{a}} - \omega_\oplus a\cos\imath\right| \qquad v_{\mathrm{max}} = \sqrt{\left(\sqrt{\frac{\mu}{a}} - \omega_\oplus a\cos\imath\right)^2 + \left(\omega_\oplus a\sin\imath\right)^2} \tag{5.39}$$

with period equal to π. Let us approximate v_{rel} with its mean value over one revolution:

$$v_{\mathrm{rel}} = \frac{v_{\mathrm{min}} + v_{\mathrm{max}}}{2} \tag{5.40}$$

Let us now apply OCT . The Hamiltonian, defined by equation (5.7), is

$$H = \lambda_a \frac{da}{d\vartheta} + \lambda_\imath \frac{d\imath}{d\vartheta} + \lambda_\Omega \frac{d\Omega}{d\vartheta} \tag{5.41}$$

where λ_x is the generic adjoint variable associated with variable x. Since a and $\imath$ are treated as constants, the Hamiltonian does not depend on the state variables and plugging equation (5.41) into the Euler-Lagrange equations yields

$$\begin{aligned} \frac{d\lambda_a}{d\vartheta} &= 0 \\ \frac{d\lambda_\imath}{d\vartheta} &= 0 \\ \frac{d\lambda_\Omega}{d\vartheta} &= 0 \end{aligned} \tag{5.42}$$

Equations (5.42) show that the adjoint variables are actually adjoint constants in the one-revolution problem. The optimal thrust angle β is obtained by nullifying the partial derivative of H with respect to β:

$$\tan\beta = \frac{\lambda_\imath \cos\vartheta + \frac{\lambda_\Omega}{\sin\imath}\sin\vartheta}{2\lambda_a a} \tag{5.43}$$

As already mentioned in Section 5.1, the generic form of the optimal control law does not formally depend on whether the performance index has to be maximized or minimized. Let us introduce the angle ϑ_0 such that

$$\tan\vartheta_0 = \frac{\lambda_\Omega / \sin\imath}{\lambda_\imath} \tag{5.44}$$

and the quantity

$$\Lambda = \sqrt{\lambda_\imath^2 + \left(\frac{\lambda_\Omega}{\sin\imath}\right)^2} \tag{5.45}$$

Let us now use equations (5.44) and (5.45) to write equation (5.43) as

$$\tan\beta = \frac{\Lambda}{2\lambda_a a}\cos(\vartheta - \vartheta_0) = K\cos(\vartheta - \vartheta_0) \tag{5.46}$$

By introducing the variable $\vartheta' = \vartheta - \vartheta_0$ one has

$$\cos\beta = \frac{1}{K'} \qquad \sin\beta = \frac{K\cos\vartheta'}{K'} \tag{5.47}$$

where $K' = \sqrt{1 + (K \cos \vartheta')^2}$. Plugging equations (5.47) into (5.36) yields

$$\begin{aligned}
\frac{da}{d\vartheta'} &= 2\frac{a^3}{\mu}\left[\frac{T}{m}\frac{1}{K'} - \frac{\rho}{2}\frac{SC_D}{m}\left(\sqrt{\frac{\mu}{a}} - \omega_{\oplus} a \cos \imath\right) v_{\text{rel}}\right] \\
\frac{d\imath}{d\vartheta'} &= \frac{a^2}{\mu}\cos\left(\vartheta' + \vartheta_0\right)\left[\frac{T}{m}\frac{K\cos\vartheta'}{K'} - \frac{\rho}{2}\frac{SC_D}{m}\left(\omega_{\oplus} a \sin \imath \cos\left(\vartheta' + \vartheta_0\right)\right) v_{\text{rel}}\right] \\
\frac{d\Omega}{d\vartheta'} &= \frac{a^2}{\mu}\frac{\sin\left(\vartheta' + \vartheta_0\right)}{\sin \imath}\left[\frac{T}{m}\frac{K\cos\vartheta'}{K'} - \frac{\rho}{2}\frac{SC_D}{m}\left(\omega_{\oplus} a \sin \imath \cos\left(\vartheta' + \vartheta_0\right)\right) v_{\text{rel}}\right] \\
&\quad -\frac{3}{2}J_2\left(\frac{r_E}{a}\right)^2 \cos \imath
\end{aligned} \tag{5.48}$$

The integration of these equation over one revolution doesn't have an analytical solution. When atmospheric drag and J2 are neglected, for $\vartheta_0 = 0$ these equations are the same as in Edelbaum's problem [8] for changes in a and $\imath$. Edelbaum showed that if instead of using the optimal control law from equation (5.46), which requires a continuously varying angle β, a sub-optimal law is adopted, the performance decrease in minimal and the equations become analytically integrable. Such sub-optimal control law is:

$$\begin{cases} \beta = \left|\overline{\beta}\right| & \cos\vartheta' > 0 \\ \beta = -\left|\overline{\beta}\right| & \cos\vartheta' < 0 \end{cases} \tag{5.49}$$

where $\overline{\beta}$ is a constant value. In practice, at $\vartheta = \vartheta_0 + \pi/2 + k\pi$ (for any integer k) the sign of the constant angle is switched. The integration over two half-revolutions with the sub-optimal law yields the changes of the orbital elements over one revolution:

$$\Delta a = 4\pi\frac{T}{m}\frac{a^3}{\mu}\cos\beta - 2\pi\frac{a^3}{\mu}\rho\frac{SC_D}{m}\left(\sqrt{\frac{\mu}{a}} - \omega_{\oplus} a \cos \imath\right) v_{\text{rel}} \tag{5.50}$$

$$\Delta \imath = 4\frac{T}{m}\frac{a^2}{\mu}\sin\beta\cos\vartheta_0 - \frac{\pi}{2}\frac{a^3}{\mu}\rho\frac{SC_D}{m}\omega_{\oplus}\sin\imath v_{\text{rel}} \tag{5.51}$$

$$\Delta\Omega = 4\frac{T}{m}\frac{a^2}{\mu}\frac{\sin\beta}{\sin\imath}\sin\vartheta_0 - 3\pi J_2\left(\frac{r_E}{a}\right)^2\cos\imath \tag{5.52}$$

$$\Delta t = 2\pi\sqrt{\frac{a^3}{\mu}} \tag{5.53}$$

If the Edelbaum's changes are denoted as Δa_E and $\Delta \imath_E$, one has

$$\Delta a = \Delta a_E - 2\pi\frac{a^3}{\mu}\rho\frac{SC_D}{m}\left(\sqrt{\frac{\mu}{a}} - \omega_{\oplus} a \cos \imath\right) v_{\text{rel}} \tag{5.54}$$

$$\Delta \imath = \Delta \imath_E \cos\vartheta_0 - \frac{\pi}{2}\frac{a^3}{\mu}\rho\frac{SC_D}{m}\omega_{\oplus}\sin\imath v_{\text{rel}} \tag{5.55}$$

$$\Delta\Omega = \Delta\imath_E\frac{\sin\vartheta_0}{\sin\imath} - 3\pi J_2\left(\frac{r_E}{a}\right)^2\cos\imath \tag{5.56}$$

In practice, the angle ϑ_0 splits the effect of the out-of-plane thrusting between inclination and RAAN change. In particular, if the sign switch of the thrust angle is performed at the nodes the out-of-plane thrusting only changes the inclination. On the other extreme, if it is performed at the antinodes the out-of-plane thrusting effort only changes the RAAN.

With the approximations introduced, equations (5.50)-5.52 show that drag produces a negative variation of semimajor axis and inclination ($\sin\imath$ is always positive as $0 < \imath < \pi$), while it has no effect on RAAN. The reason for this lies in the out-of-plane acceleration component of drag

$$(f_W)_{\text{drag}} = -\frac{\rho}{2}\frac{SC_D}{m}\omega_{\oplus} a \sin\imath \cos\vartheta v_{\text{rel}} \tag{5.57}$$

which is in the form

$$(f_W)_{\text{drag}} = -F\cos\vartheta \tag{5.58}$$

with F positive constant. For a half revolution, namely from $\vartheta = -\pi/2$ to $\vartheta = \pi/2$, the component is negative, while for the other half (from $\vartheta = \pi/2$ to $\vartheta = 3\pi/2$) is positive. Since the time derivative of $\imath$ is

$$\dot{\imath} = \sqrt{\frac{a}{\mu}}\cos\vartheta f_W \tag{5.59}$$

during the first half revolution f_W is negative and produces a negative change of inclination ($\cos\vartheta$ is positive for $\vartheta = -\pi/2$ to $\vartheta = \pi/2$), while during the second half f_W is positive but it still produces a negative change since $\cos\vartheta$ is negative. In particular,

$$(\Delta\imath)_{\text{drag}} = -F\sqrt{\frac{a}{\mu}}\int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}}\cos^2\vartheta d\vartheta = -\pi F\sqrt{\frac{a}{\mu}} \tag{5.60}$$

On the other hand, the time derivative of Ω is given by

$$\dot{\Omega} = \sqrt{\frac{a}{\mu}}\frac{\sin\vartheta}{\sin\imath}f_w \tag{5.61}$$

Therefore, during the first half revolution f_W is negative and produces a positive change of RAAN during the first quarter ($\sin\vartheta$ is negative) and a negative change during the second quarter ($\sin\vartheta$ is positive). During the second half f_W is positive and produces a positive change of RAAN during the third quarter ($\sin\vartheta$ is positive) and a negative change during the fourth quarter ($\sin\vartheta$ is negative). The overall effect is null. Rigorously:

$$(\Delta\Omega)_{\text{drag}} = -\frac{F}{\sin\imath}\sqrt{\frac{a}{\mu}}\int_{-\frac{\pi}{2}}^{\frac{3\pi}{2}}\cos\vartheta\sin\vartheta d\vartheta = 0 \tag{5.62}$$

5.3.2 Multiple-revolution transfer

Let us use the ratios of the changes of the orbital elements over one revolution to the time required to complete it to approximate the time derivative of the orbital elements for the multiple-revolution transfer. Therefore, the physical system is described by

$$\frac{da}{dt} = 2\frac{T}{m}\sqrt{\frac{a^3}{\mu}}\cos\beta + \rho\frac{SC_D}{m}v_{\rm rel}a\left(\sqrt{\frac{a^3}{\mu}}\omega_\oplus\cos\imath - 1\right) \tag{5.63}$$

$$\frac{d\imath}{dt} = \frac{2}{\pi}\frac{T}{m}\sqrt{\frac{a}{\mu}}\sin\beta\cos\vartheta_0 - \frac{1}{4}\sqrt{\frac{a^3}{\mu}}\rho\frac{SC_D}{m}\omega_\oplus\sin\imath v_{\rm rel} \tag{5.64}$$

$$\frac{d\Omega}{dt} = \frac{2}{\pi}\frac{T}{m}\sqrt{\frac{a}{\mu}}\frac{\sin\beta}{\sin\imath}\sin\vartheta_0 - \frac{3}{2}J_2\left(\frac{r_E}{a}\right)^2\sqrt{\frac{\mu}{a^3}}\cos\imath \tag{5.65}$$

$$\frac{dm}{dt} = -\frac{T}{c} \tag{5.66}$$

where $c = g_0 I_{sp}$ is the effective exhaust velocity. Equation (5.66) is added to account for the propellant consumption over multiple revolutions. The angle β varies from one revolution to another, while it is constant during each revolution and a sign switch occurs depending on ϑ_0.

Let us apply OCT to this system as well. The Hamiltonian is

$$\begin{aligned} H = {} & \lambda_a a\left[2\frac{T}{m}\sqrt{\frac{a}{\mu}}\cos\beta + \rho\frac{SC_D}{m}v_{\rm rel}\left(\frac{\omega_\oplus}{n_{\rm S/C}}\cos\imath - 1\right)\right] \\ & +\lambda_\imath\left[\frac{2}{\pi}\frac{T}{m}\sqrt{\frac{a}{\mu}}\sin\beta\cos\vartheta_0 - \frac{1}{4}\rho\frac{SC_D}{m}\frac{\omega_\oplus}{n_{\rm S/C}}\sin\imath v_{\rm rel}\right] \\ & +\lambda_\Omega\left[\frac{2}{\pi}\frac{T}{m}\sqrt{\frac{a}{\mu}}\frac{\sin\beta}{\sin\imath}\sin\vartheta_0 - \frac{3}{2}J_2\left(\frac{r_E}{a}\right)^2 n_{\rm S/C}\cos\imath\right] \\ & -\lambda_m\frac{T}{c} \end{aligned} \tag{5.67}$$

where the mean motion of the spacecraft $n_{\rm S/C} = \sqrt{\mu/a^3}$ has been introduced.

The Euler-Lagrange equations are

$$\begin{aligned}
\frac{d\lambda_a}{dt} &= -\lambda_a 3\frac{T}{m}\sqrt{\frac{a}{\mu}}\cos\beta \\
&+\lambda_a\frac{SC_D}{m}\left(\rho v_{\mathrm{rel}}+a\frac{d\rho}{da}v_{\mathrm{rel}}+a\rho\frac{dv_{\mathrm{rel}}}{da}\right) \\
&-\lambda_a\frac{SC_D}{m}\frac{\omega_\oplus}{n_{\mathrm{S/C}}}\cos\imath\left(\frac{5}{2}\rho v_{\mathrm{rel}}+a\frac{d\rho}{da}v_{\mathrm{rel}}+a\rho\frac{dv_{\mathrm{rel}}}{da}\right) \\
&-\lambda_\imath\frac{1}{\pi}\frac{T}{m}\frac{1}{\sqrt{a\mu}}\sin\beta\cos\vartheta_0 \\
&+\frac{\lambda_\imath}{4}\frac{SC_D}{m}\frac{\omega_\oplus}{n_{\mathrm{S/C}}}\sin\imath\left(\frac{3}{2}\frac{\rho}{a}v_{\mathrm{rel}}+\frac{d\rho}{da}v_{\mathrm{rel}}+\rho\frac{dv_{\mathrm{rel}}}{da}\right) \\
&-\lambda_\Omega\left[\frac{1}{\pi}\frac{T}{m}\frac{1}{\sqrt{a\mu}}\frac{\sin\beta}{\sin\imath}\sin\vartheta_0+\frac{21}{4}J_2\left(\frac{r_E}{a}\right)^2\frac{n_{\mathrm{S/C}}}{a}\cos\imath\right] \qquad (5.68)
\end{aligned}$$

$$\begin{aligned}
\frac{d\lambda_\imath}{dt} &= \lambda_a a\left[\rho\frac{SC_D}{m}\frac{dv_{\mathrm{rel}}}{d\imath}-\rho\frac{SC_D}{m}\frac{\omega_\oplus}{n_{\mathrm{S/C}}}\left(\cos\imath\frac{dv_{\mathrm{rel}}}{d\imath}-\sin\imath v_{\mathrm{rel}}\right)\right] \\
&+\frac{\lambda_\imath}{4}\rho\frac{SC_D}{m}\frac{\omega_\oplus}{n_{\mathrm{S/C}}}\left(\cos\imath v_{\mathrm{rel}}+\sin\imath\frac{dv_{\mathrm{rel}}}{d\imath}\right) \\
&+\lambda_\Omega\left[\frac{2}{\pi}\frac{T}{m}\sqrt{\frac{a}{\mu}}\sin\beta\frac{\sin\vartheta_0}{\sin^2\imath}\cos\imath-\frac{3}{2}J_2\left(\frac{r_E}{a}\right)^2 n_{\mathrm{S/C}}\sin\imath\right] \qquad (5.69)
\end{aligned}$$

$$\frac{d\lambda_\Omega}{dt} = 0 \qquad (5.70)$$

$$\begin{aligned}
\frac{d\lambda_m}{dt} &= \lambda_a a\left[2\frac{T}{m^2}\sqrt{\frac{a}{\mu}}\cos\beta+\rho\frac{SC_D}{m^2}v_{\mathrm{rel}}\left(\frac{\omega_\oplus}{n_{\mathrm{S/C}}}\cos\imath-1\right)\right] \\
&+\lambda_\imath\left[\frac{2}{\pi}\frac{T}{m^2}\sqrt{\frac{a}{\mu}}\sin\beta\cos\vartheta_0-\frac{1}{4}\rho\frac{SC_D}{m^2}\frac{\omega_\oplus}{n_{\mathrm{S/C}}}\sin\imath v_{\mathrm{rel}}\right] \\
&+\lambda_\Omega\frac{2}{\pi}\frac{T}{m^2}\sqrt{\frac{a}{\mu}}\frac{\sin\beta}{\sin\imath}\sin\vartheta_0 \qquad (5.71)
\end{aligned}$$

The expressions for dv_{rel}/da, $dv_{\mathrm{rel}}/d\imath$ and $d\rho/da$ are given by:

$$\begin{aligned}
\frac{dv_{\mathrm{rel}}}{da} &= -\frac{1}{4}n_{\mathrm{S/C}}-\frac{\omega_\oplus}{2}\cos\imath+ \\
&-\frac{\left(n_{\mathrm{S/C}}-\omega_\oplus\cos\imath\right)\left(\frac{1}{2}n_{\mathrm{S/C}}+\omega_\oplus\cos\imath\right)-\left(\omega_\oplus\sin\imath\right)^2}{2\sqrt{\left(n_{\mathrm{S/C}}-\omega_\oplus\cos\imath\right)^2+\left(\omega_\oplus\sin\imath\right)^2}} \qquad (5.72)
\end{aligned}$$

$$\frac{dv_{\mathrm{rel}}}{d\imath} = \frac{\omega_\oplus a\sin\imath}{2}\left[1+\frac{n_{\mathrm{S/C}}}{\sqrt{\left(n_{\mathrm{S/C}}-\omega_\oplus\cos\imath\right)^2+\left(\omega_\oplus\sin\imath\right)^2}}\right] \qquad (5.73)$$

$$\frac{d\rho}{da} = \rho\left[4A\left(a-r_E\right)^3+3B\left(a-r_E\right)^2+C\left(a-r_E\right)+D\right] \qquad (5.74)$$

where r_E is the radius of the Earth and the coefficients A, B, C and D depend on the altitude as described in Section (4.1).

The control variables are β, ϑ_0, T and S. Nullifying the partial derivative of H with respect to ϑ_0 yields

$$\tan \vartheta_0 = \frac{\lambda_\Omega}{\lambda_\imath \sin \imath} \tag{5.75}$$

whereas nullifying the partial derivative of H with respect to β yields

$$\tan \beta = \frac{\lambda_\imath \cos \vartheta_0 + \frac{\lambda_\Omega}{\sin \imath} \sin \vartheta_0}{\pi a \lambda_a} \tag{5.76}$$

Therefore, one has

$$\sin \beta = \pm \frac{\Lambda}{\sqrt{\Lambda^2 + (\pi a \lambda_a)^2}} \tag{5.77}$$

$$\cos \beta = \pm \frac{\pi a \lambda_a}{\sqrt{\Lambda^2 + (\pi a \lambda_a)^2}} \tag{5.78}$$

$$\sin \vartheta_0 = \pm \frac{\lambda_\Omega}{\Lambda \sin \imath} \tag{5.79}$$

$$\cos \vartheta_0 = \pm \frac{\lambda_\imath}{\Lambda} \tag{5.80}$$

where Λ is given by equation (5.45). The optimization problem is here formulated as a maximization problem. Therefore, in agreement with the Pontryagin's maximum principle the optimal controls must maximize the Hamiltonian. In order to satisfy the principle, from equation (5.67) one notes that $\cos \beta$ must have the same sign as λ_a and $\sin \beta$ must have the same sign as $\lambda_\imath \cos \vartheta_0 + \lambda_\Omega \sin \vartheta_0 / \sin \imath$. By arbitrarily selecting $\sin \beta > 0$, one has that $\cos \vartheta_0$ must have the same sign as $\lambda_\imath$ and $\sin \vartheta_0$ must have the same sign as $\lambda_\Omega / \sin \imath$. Therefore, selecting the correct quadrants yields

$$\sin \beta = \frac{\Lambda}{\sqrt{\Lambda^2 + (\pi a \lambda_a)^2}} \tag{5.81}$$

$$\cos \beta = \frac{\pi a \lambda_a}{\sqrt{\Lambda^2 + (\pi a \lambda_a)^2}} \tag{5.82}$$

$$\sin \vartheta_0 = \frac{\lambda_\Omega}{\Lambda \sin \imath} \tag{5.83}$$

$$\cos \vartheta_0 = \frac{\lambda_\imath}{\Lambda} \tag{5.84}$$

The Hamiltonian is linear with respect to the thrust magnitude and can be written as

$$\begin{aligned} H = {} & S_{FT} T + \rho \frac{S C_D}{m} v_{\mathrm{rel}} \left[\lambda_a a \left(\frac{\omega_\oplus}{n_{\mathrm{S/C}}} \cos \imath - 1 \right) - \lambda_\imath \frac{1}{4} \frac{\omega_\oplus}{n_{\mathrm{S/C}}} \sin \imath \right] \\ & - \lambda_\Omega \frac{3}{2} J_2 \left(\frac{r_E}{a} \right)^2 n_{\mathrm{S/C}} \cos \imath \end{aligned} \tag{5.85}$$

where S_{FT} denotes the thrust switching function:

$$S_{FT} = \frac{2}{\pi}\frac{1}{m}\sqrt{\frac{a}{\mu}}\sqrt{\Lambda^2 + (\pi a \lambda_a)^2} - \frac{\lambda_m}{c} \tag{5.86}$$

In agreement with the Pontryagin's maximum principle, the thrust magnitude must assume its maximum value when $S_{FT} > 0$ and its minimum value, that is the thruster must be turned off, when $S_{FT} > 0$.

Analogously, the Hamiltonian is linear with respect to the frontal area of the satellite and can be written as

$$H = S_{FT}T + S_{FS}S - \lambda_\Omega \frac{3}{2} J_2 \left(\frac{r_E}{a}\right)^2 n_{\mathrm{S/C}} \cos \imath \tag{5.87}$$

where S_{FS} denotes the area switching function:

$$S_{FS} = \rho \frac{C_D}{m} v_{\mathrm{rel}} \left[\lambda_a a \left(\frac{\omega_\oplus}{n_{\mathrm{S/C}}} \cos \imath - 1\right) - \lambda_\imath \frac{1}{4} \frac{\omega_\oplus}{n_{\mathrm{S/C}}} \sin \imath\right] \tag{5.88}$$

The frontal area must assume its maximum value, that is the drag sail is deployed, when $S_{FS} > 0$. On the other hand, the frontal area must assume its minimum value, that is the drag sail is furled, when $S_{FS} < 0$.

5.3.3 Problem formulation

The optimal control law depends on the cost function and on the boundary conditions on the variables. The optimization problem is here formulated by imposing the initial orbital elements a_0, $\imath_0$ and Ω_0 and the target orbital elements at the initial time $t_0 = 0$. The initial RAAN is fixed to zero ($\Omega_0 = 0$) by properly selecting the coordinate system; if the initial orbit is defined with a certain RAAN $\Omega_0 = \Omega_p$ with respect to the geocentric-equatorial coordinate system, the simple rotation through an angle Ω_p of such system about K yields a new coordinate system in which $\Omega_0 = 0$. The orbital elements of the target S/C are perturbed by J2, which causes the target RAAN to change with a rate $\left(\dot{\Omega}_{J2}\right)_T$ function of a_T and $\imath_T$ only. The boundary conditions at the final time are:

$$a_f = a_T \tag{5.89}$$

$$\imath_f = \imath_T \tag{5.90}$$

$$\Omega_f = \Omega_{T0} + \left(\dot{\Omega}_{J2}\right)_T t_f \tag{5.91}$$

It goes without saying that the target orbit should be perturbed with aerodynamic drag as well. However, the case study object of this work assumes that the target S/C has the ability to maintain the semimajor axis and inclination of the orbit unchanged through station keeping maneuvers. The same optimization method proposed here is readily applicable in other circumstances, as in the case of a passive target satellite perturbed by J2 and drag.

The theory of optimal control provides the boundary conditions for optimality, which depend on the cost function. In the case of minimum-time with free final mass, the cost function to be maximized is

$$J = -t_f \tag{5.92}$$

and the optimal boundary conditions for optimality are

$$H_f - \lambda_\Omega \left(\dot{\Omega}_{J2}\right)_T = 1 \tag{5.93}$$

$$\lambda_{mf} = 0 \tag{5.94}$$

In this case the thrust switching function is always positive and the engine is always on at maximum thrust. The five boundary conditions at the final time (equations (5.89), (5.90), (5.91), (5.93) and (5.94)) determine the five unknown of the two point boundary value problem (2PBVP): t_f, λ_{a0}, $\lambda_{\imath 0}$, λ_Ω, λ_{m0} (λ_Ω is constant and thus it is not denoted with the subscript 0). The boundary condition (5.93) on the final value of the Hamiltonian can be replaced by specifying one of the initial values of an adjoint variable, such as λ_Ω. As a matter of fact, the problem is homogeneous in the adjoint variables and they appear only as fractions in the algebraic equations for the control variables; by specifying λ_Ω, the adjoint variables and the Hamiltonian are scaled by a scale factor with respect to the solution that yields $H_f - \lambda_\Omega \left(\dot{\Omega}_{J2}\right)_T = 1$, while the optimal controls are the same (the scale factor is canceled out in the fractions). Therefore, when the values for t_f, λ_{a0}, $\lambda_{\imath 0}$ and λ_{m0} that satisfy equations equations (5.89), (5.90), (5.91) and (5.94) are found, one has

$$H_f - \lambda_\Omega \left(\dot{\Omega}_{J2}\right)_T = \text{scale factor} \tag{5.95}$$

Finally, one may divide λ_{a0}, $\lambda_{\imath 0}$, λ_Ω and λ_{m0} by the scale factor and obtain $H_f - \lambda_\Omega \left(\dot{\Omega}_{J2}\right)_T = 1$. However, this last step would be useless, as the optimal control law would be exactly the same. Consequently, the problem can be reduced to four unknowns and four boundary conditions. However, one has to note that the proper sign for λ_Ω must be selected in order to avoid negative time-of-flight solutions.

In the case of minimum-propellant, the cost function to be maximized is

$$J = m_f \tag{5.96}$$

and the optimality boundary condition are

$$\lambda_{mf} = 1 \tag{5.97}$$

and either

$$t_f = k \tag{5.98}$$

for specified final time or

$$H_f - \lambda_\Omega \left(\dot{\Omega}_{J2}\right)_T = 0 \tag{5.99}$$

for free final time. The boundary condition (5.97) can be replaced by specifying λ_Ω. The thrust switching function can now become negative and the engine must be turned off when this occurs.

In some cases, the optimal trajectory requires a decrease of the orbit altitude in order to increase the effect of J2. However, the S/C may be required to reach such a low altitude that drag would make it impossible to increase its altitude again. Particularly, there is a limit altitude for which this happens. However, the exact value depends on the local properties of the atmosphere and it would be impossible to calculate it in advance. For this reason, a safe altitude $h_{\text{lim}} = 200\,\text{km}$ is considered as a minimum constraint. When such constraint is introduced, a three-arc structure becomes optimal; the satellite follows the optimal controls from t_0 to t_1 and from t_2 to t_f, whereas it flies at the minimum altitude during the intermediate arc. In order to make this happen, the thrust angle must have a value such that $da/dt = 0$. That is

$$\cos\beta = \sqrt{\frac{\mu}{a}}\frac{\rho S C_D}{2T}\overline{v}_{\text{rel}}\left(\frac{\omega_\oplus}{n_{\text{S/C}}}\cos\imath - 1\right) \tag{5.100}$$

The minimum altitude boundary is introduced with the additional boundary condition

$$a_2 = r_E + h_{\text{lim}} \tag{5.101}$$

at the end of the intermediate arc. From the optimal boundary conditions one has that H and the adjoint variables must be continuous at the internal boundaries, except for λ_a which has a free discontinuity at point 2. As a result, $\cos\beta$ must be continuous at t_1 and t_2 and two boundary conditions for optimality are added:

$$\sqrt{\frac{\mu}{a_1}}\frac{\rho S C_D}{2T}v_{\text{rel1}}\left(\frac{\omega_\oplus}{n_{\text{S/C1}}}\cos\imath_1 - 1\right) = \frac{\pi a_1 \lambda_{a1}}{\sqrt{\Lambda_1^2 + (\pi a_1 \lambda_{a1})^2}} \tag{5.102}$$

$$\sqrt{\frac{\mu}{a_2}}\frac{\rho S C_D}{2T}v_{\text{rel2}}\left(\frac{\omega_\oplus}{n_{\text{S/C2}}}\cos\imath_2 - 1\right) = \frac{\pi a_2 \lambda_{a2+}}{\sqrt{\Lambda_2^2 + (\pi a_2 \lambda_{a2+})^2}} \tag{5.103}$$

and the problem has three additional unknowns: t_1, t_2 and λ_{a2-}.

Chapter 6

Results

This chapter presents the results for a case study concerning a small 15-kg spacecraft in an initial orbit similar to that of the International Space Station. The altitude of the initial orbit is fixed at 400 km and the initial inclination is 51 degrees. Since we are dealing with direct orbits ($\imath < 90°$), the line of nodes regresses as previously described in Section 2.4.2.2. Particularly, the RAAN regression rate increases as the altitude is decreased or as the inclination is decreased, as it can be seen from equation (4.30). The electric propulsion system is supposed to be able to provide 10 mN of thrust with a specific impulse of 2500 s. The frontal area of the S/C is supposed to be equal to $0.04\,\mathrm{m}^2$ and this value is assumed to correspond to the minimum. Three drag sail scenarios are considered; the first scenario assumes that the S/C is not equipped with a drag sail and $S_{\max}$ is always equal to the minimum value; the second scenario assumes a maximum drag sail area of $4\,\mathrm{m}^2$, whereas the value for the third scenario is $400\,\mathrm{m}^2$. The frontal area value of the third scenario is unrealistically large. Even though such a sail has been developed and tested [10], the mass-to-area ratio and the packed size in order to be mounted on a small satellite would have to be far lower than current technologies can achieve. However, this currently unrealistic scenario has been chosen to be compared to the second scenario, which is characterized by a frontal area value which would make the sail very light with current technologies [24]. The value of the drag coefficient is fixed to 2.5, although a more rigorous analysis could be carried out by considering its variation at different altitudes and at epochs between solar maximum and minimum.

The numerical results are obtained by following the procedure described in Section 5.3 and by integrating the differential equations with a variable-step variable-order method based on the Adams-Moulton method, as described by Shampine and Gordon [21]. Since much easier convergence can be obtained if the variables have the similar orders of magnitude, variables normalization has been adopted. The wanted precision for the boundary conditions is set to 10^{-7}. That is, the maximum error $E_{\max} = \max \Psi_i$ has to be lower than such value. According to the procedure the initial tentative solution is corrected by a quantity $\Delta\mathbf{p}$, computed by neglecting higher than first order terms. Such linearization could yield values for $\Delta\mathbf{p}$ that would make the error increase. In order to avoid this, a relaxation factor K_1 has been introduced, such that

$$\mathbf{p}^{r+1} = \mathbf{p}^r + K_1 \Delta\mathbf{p} \tag{6.1}$$

where K_1 ranges from 0.1 to 1. Therefore, such relaxation factor reduces the theoretical correction of the procedure. In addition, if the maximum error $E_{\max}^{r+1}$ of the iteration $r+1$ is lower than a multiple of the maximum error $E_{\max}^{r}$ of the iteration r, the code proceeds with a new iteration. Namely, a new iteration is started if $E_{\max}^{r+1} < K_2 E_{\max}^{r}$ with $K_2 = 2 \div 3$. Since the boundary conditions error can grow after the first iterations, the introduction of this factor allows this behavior but stops the procedure if the error is growing too much. On the other hand, if the error is too big with respect to the previous iteration ($E_{\max}^{r+1} > K_2 E_{\max}^{r}$) a bisection of the correction is performed:

$$\mathbf{p}^{r+1} = \mathbf{p}^r + K_1 \frac{\Delta\mathbf{p}}{2} \tag{6.2}$$

The iteration is then repeated and the new maximum error is compared to that of the previous iteration. If necessary, the iteration is repeat for a maximum of five times and then the procedure is stopped.

6.1 Minimum-time solutions

As described in Section 5.3.3, the transfer is characterized by the chaser's initial orbital elements a_0, $\imath_0$ and Ω_0 and the target's orbital elements a_T, $\imath_T$ and Ω_{T0} at the initial time $t_0 = 0$. The initial RAAN of the chaser is fixed to zero by properly selecting the coordinate system. Thus, the initial RAAN phase angle is $\Delta\Omega_0 = \Omega_{T0}$. Since a_T and $\imath_T$ are unchanged during the transfer, the boundary conditions at the final time are given by equations (5.89), (5.90) and (5.91). Namely, at the final time Δa, $\Delta\imath$ and $\Delta\Omega$ must be zero. For every combination of a_0, $\imath_0$, a_T and $\imath_T$, the minimum-time problem is characterized by a value of initial RAAN phase angle $\Delta\Omega^*$ such that the time-of-flight reaches a global minimum. In this case, the chaser achieves the target's RAAN thanks to the effect of J2 only, while thrust is only used to change the semimajor axis and the inclination of the orbit. When $\Delta\Omega_0 \neq \Delta\Omega^*$ part of the thrusting effort is used to change Ω and the trip time is greater than the global minimum. Since minimum-time transfers require the engine to be always on at maximum thrust, shorter transfers correspond to smaller propellant consumption.

6.1.1 Negative change of altitude

The results for transfers with an altitude change of $-200\,\mathrm{km}$ are shown in figure 6.1 and displayed in table 6.1. Figures 6.1a, 6.1b and 6.1c show the results for different inclination changes for the no drag sail scenario, the $4\,\mathrm{m}^2$ scenario and the $400\,\mathrm{m}^2$ scenario, respectively.

In all cases, the line of nodes of the target orbit is initially regressing at a higher rate. Therefore, at the initial time the J2 perturbation moves the orbit plane of the target closer to the orbit plane of the chaser if $\Delta\Omega_0 > 0$. Vice versa, the two orbit planes initially drift away from each other if $\Delta\Omega_0 < 0$, as the target orbit plane regresses faster and $\Delta\Omega$ initially increases. The value of $\Delta\Omega^*$ is between 0 and 1 degree in all cases. When $\Delta\Omega_0$ is less than $\Delta\Omega^*$ by a sufficiently large margin, the optimal transfer with a positive inclination change requires less time. On the other hand, when $\Delta\Omega_0$ is greater than $\Delta\Omega^*$, the optimal transfer with a negative inclination change is slightly faster. The reason for this can be explained with figure 6.2, which shows the semimajor axis and inclination time histories for a $\Delta\Omega_0 = -10°$ transfer (figure 6.2a) and for a $\Delta\Omega_0 = 10°$ transfer (figure 6.2b). Since for all sail scenarios the behavior of the trajectories is the same, the $4\,\mathrm{m}^2$ scenario is used as an example.

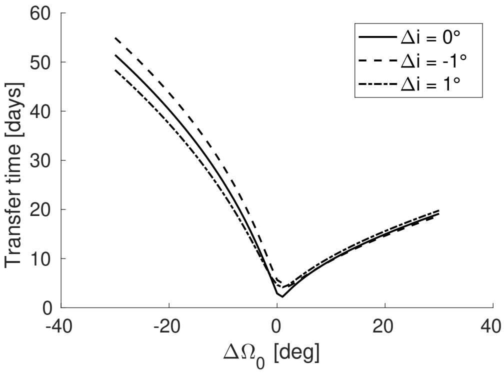
In the $\Delta\Omega_0 = -10°$ case, the orbit planes of chaser and target are initially moved away by the effect of J2. The optimal transfer requires the chaser to initially decrease its altitude and reach the limit of 200 km, which is also the altitude of the target orbit. From this point on, the RAAN change can only be achieved by means of out-of-plane thrusting. Since the cost of changing Ω can be reduced by decreasing the inclination of an orbit (as seen in Section 2.4.1.1), in all $\Delta\imath_0$ cases the inclination is initially reduced. However, the greater part of the out-of-plane thrusting effort for the optimal $\Delta\imath_0 = -1°$ trajectory is initially used to decrease the inclination of the orbit rather than to attain the required RAAN change. Although $\imath$ reaches a lower value than the other two trajectories and the cost of changing Ω is also lower, the initial RAAN rate of change is slower. The overall effect is that the trajectory ends up being longer. On the other hand, in the $\Delta\Omega_0 = 10°$ case the two orbit planes are initially moved closer. In order speed up this process, the optimal transfer requires the chaser to initially increase its altitude in order to reduce the effect of the J2 perturbation. Even though a positive inclination change would further reduce the perturbation, a negative change is beneficial in terms of cost of changing Ω. Between the two effects, a negative inclination change results to be more efficient.

Figures 6.1d, 6.1e and 6.1f show the same results but highlight the effect of maximum frontal area. As a matter of fact, each one of the figures displays the same transfer for the three different scenarios. As it can be seen from the figures, the benefit of having a drag sail are more important when $\Delta\Omega_0$ is greater than $\Delta\Omega^*$. Since the effect of $S_{\max}$ on the trajectories is analogous for all $\Delta\imath_0$ cases, figure 6.3 shows the semimajor axis and inclination time histories for the $\Delta\imath_0 = 1°$ case, used as an example. The dashed line denotes the time intervals of the trajectories when the sail is not deployed, while the solid line corresponds to the time intervals when the sail is deployed.

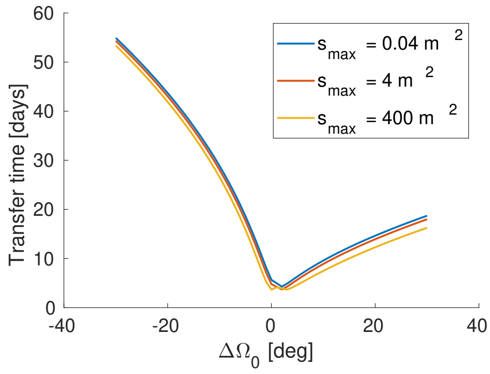
In the $\Delta\Omega_0 = -10°$ case (figure 6.3a), the sail is used to speed up the initial altitude decrease, while the rest of the transfer is carried out with minimum frontal area. Therefore, the advantage of increased drag can only be exploited for a limited amount of time and the time-of-flight improvement is slightly less than that of the $\Delta\Omega_0 = 10°$ case. As a matter of fact, in this other configuration the chaser S/C can exploit the drag sail for a much longer time. In addition, as the maximum frontal area increases the chaser can reach higher altitudes. Not only this allows for a faster convergence of the two orbital planes, but the higher value of S_{max} can then be exploited for also decreasing the altitude at a much higher rate.

Table 6.1: Minimum transfer time and corresponding propellant consumption for transfers with an altitude change of $-200\,\text{km}$

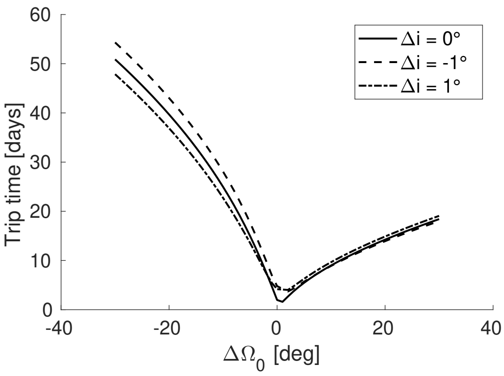
		No drag sail		$S_{\text{max}} = 4\,\text{m}^2$		$S_{\text{max}} = 400\,\text{m}^2$	
$\Delta\imath$	$\Delta\Omega_0$	t_f	m_p [kg]	t_f	m_p [kg]	t_f	m_p [kg]
	-30°	54 d 22h	1.936	54 d 8h	1.914	53 d 9h	1.880
	-20°	43 d 17h	1.540	43 d 1h	1.516	41 d 21h	1.476
	-10°	28 d 21h	1.017	28 d 1h	0.989	26 d 14h	0.936
-1°	0°	5 d 16h	0.199	4 d 19h	0.169	3 d 16h	0.129
	10°	9 d 14h	0.338	8 d 22h	0.314	7 d 12h	0.264
	20°	14 d 15h	0.516	13 d 22h	0.491	12 d 7h	0.433
	30°	18 d 17h	0.660	17 d 24h	0.634	16 d 6h	0.573
	-30°	51 d 11h	1.813	50 d 21h	1.793	49 d 23h	1.761
	-20°	40 d 7h	1.420	39 d 16h	1.398	38 d 15h	1.360
	-10°	25 d 16h	0.905	24 d 23h	0.880	23 d 14h	0.832
0°	0°	2 d 21h	0.102	1 d 23h	0.070	0 d 3h	0.004
	10°	9 d 16h	0.341	8 d 24h	0.317	7 d 14h	0.267
	20°	14 d 23h	0.527	14 d 6h	0.502	12 d 14h	0.444
	30°	19 d 3h	0.674	18 d 10h	0.649	16 d 16h	0.587
	-30°	48 d 9h	1.705	47 d 21h	1.687	46 d 24h	1.656
	-20°	37 d 10h	1.319	36 d 21h	1.300	35 d 21h	1.265
	-10°	23 d 8h	0.822	22 d 17h	0.801	21 d 12h	0.758
1°	0°	4 d 16h	0.164	4 d 4h	0.147	-	-
	10°	10 d 9h	0.365	9 d 16h	0.341	8 d 5h	0.290
	20°	15 d 14h	0.550	14 d 21h	0.524	13 d 5h	0.466
	30°	19 d 19h	0.697	19 d 2h	0.672	17 d 8h	0.610



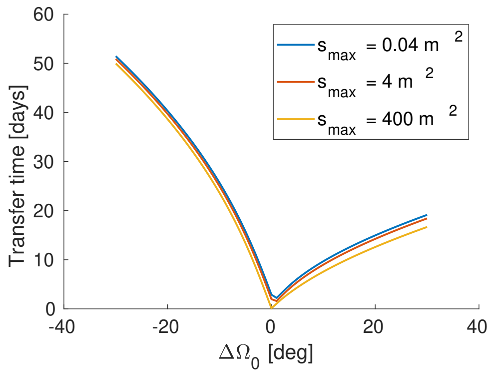
(a) No sail scenario



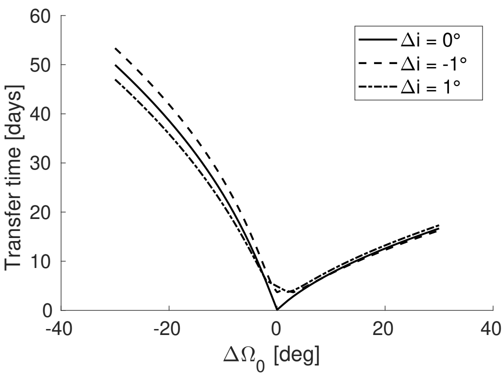
(d) Inclination change of −1 degree



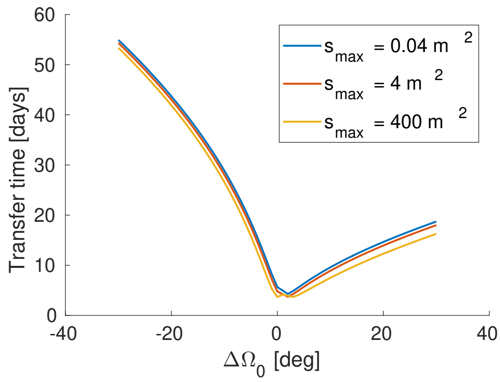
(b) Maximum drag sail area equal to $4\,\mathrm{m}^2$



(e) No inclination change

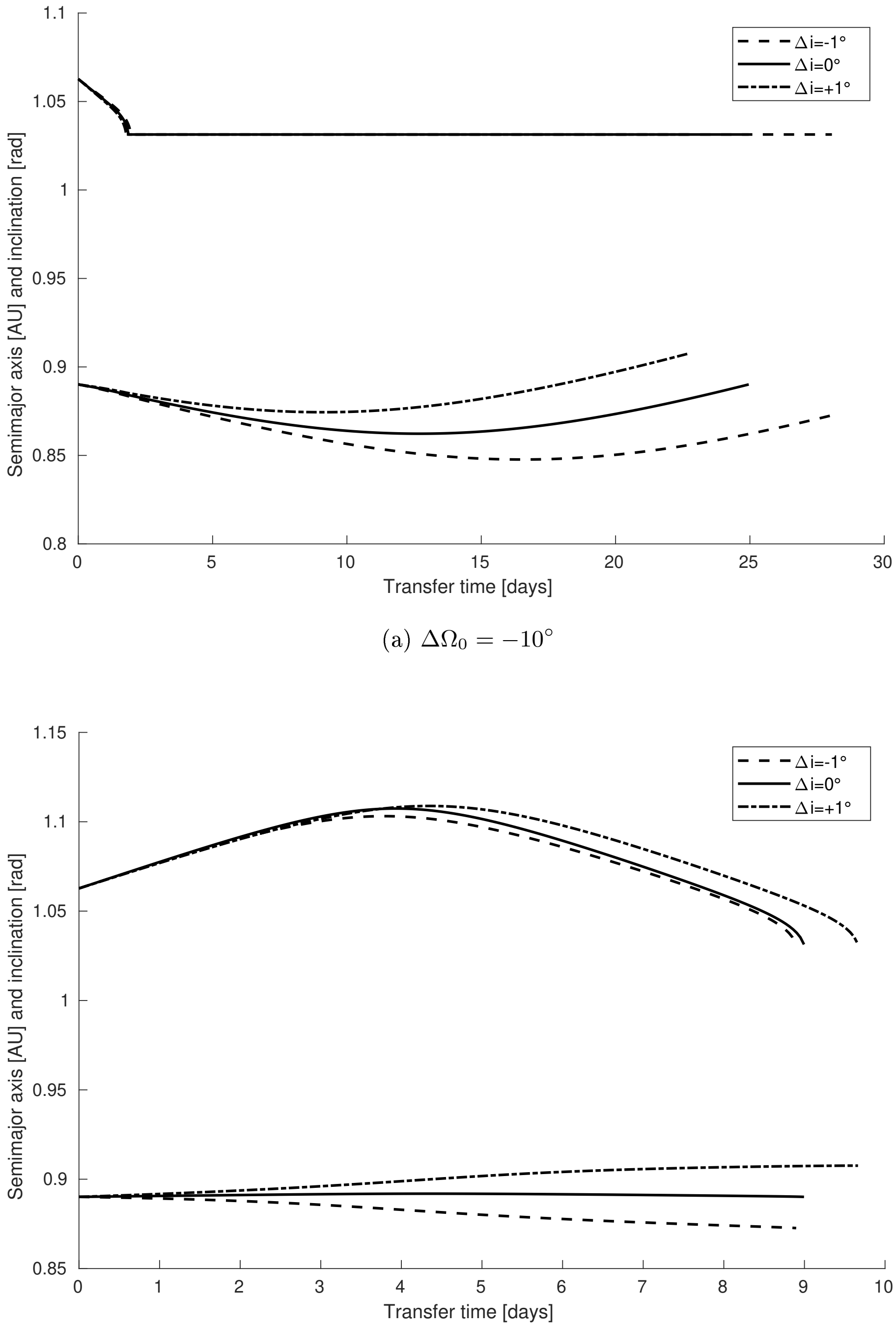


(c) Maximum drag sail area equal to $400\,\mathrm{m}^2$



(f) Inclination change of 1 degree

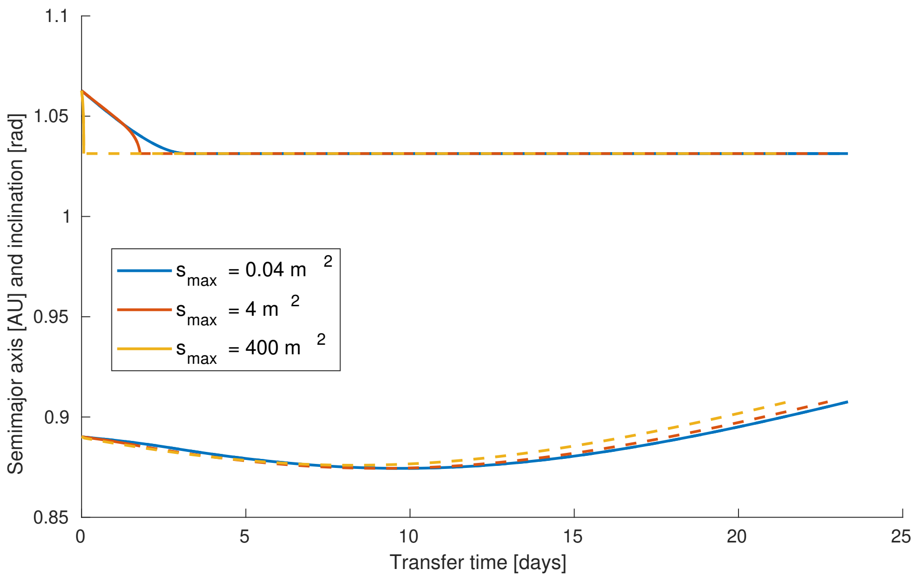
Figure 6.1: Minimum transfer time as a function of the target's initial RAAN for transfers with an altitude change of $-200\,\mathrm{km}$



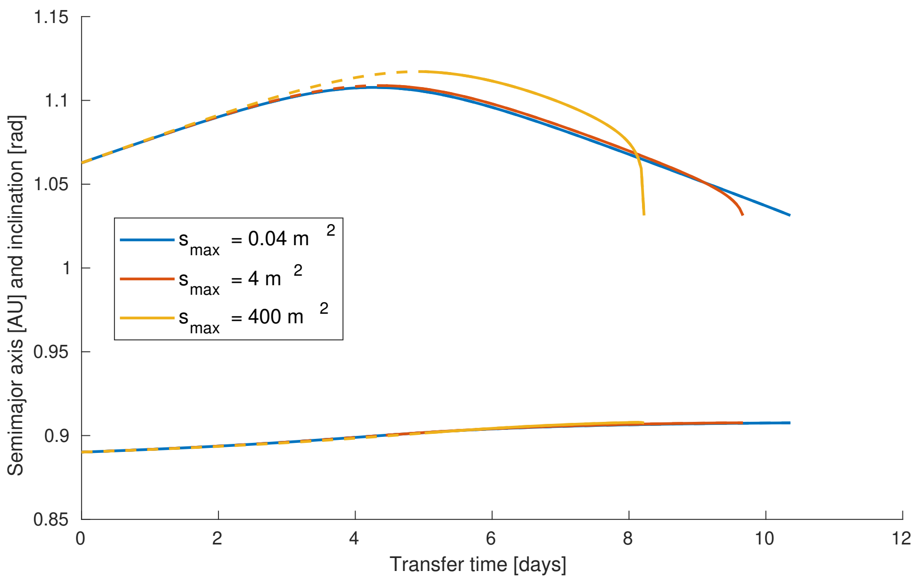
(a) $\Delta\Omega_0 = -10°$

(b) $\Delta\Omega_0 = +10°$

Figure 6.2: Minimum-time trajectories for $\Delta a_0 = -200\,\mathrm{km}$ and $S_{\max} = 4\,\mathrm{m}^2$



(a) $\Delta\Omega_0 = -10°$



(b) $\Delta\Omega_0 = +10°$

Figure 6.3: Minimum-time trajectories for $\Delta a_0 = -200\,\text{km}$ and $\Delta\imath_0 = 1°$

6.1.2 Positive change of altitude

As opposed to the negative Δa_0 case, if the target orbit is at a higher altitude, its line of nodes initially regresses at a slower rate than that of the chaser. Thus, at the initial time the two orbit planes drift away from each other if $\Delta\Omega_0 > 0$ and move closer to each other if $\Delta\Omega_0 < 0$.

The results for transfers with an altitude change of +200 km exhibit an analogous qualitative behavior to those of the $\Delta a_0 = -200$ km transfer, as shown in figure 6.4. This time, the value of $\Delta\Omega^*$ is between 0 and -1 degree in all cases. When $\Delta\Omega_0$ is less than $\Delta\Omega^*$ the optimal transfer with a positive inclination change takes less time to be completed and, vice versa, when $\Delta\Omega_0$ is greater than $\Delta\Omega^*$ the optimal transfer with a negative inclination change is faster. Using the 4 m^2 scenario as an example, the semimajor axis and inclination time histories for a $\Delta\Omega_0 = -30°$ transfer and for a $\Delta\Omega_0 = 30°$ transfer are displayed in figures 6.5a and 6.5b, respectively.

In the $\Delta\Omega_0 = -30°$ case, the orbit planes of chaser and target are initially moved closer to each other by the effect of J2. Even though the target altitude is higher than that of the chaser, the optimal transfer requires the S/C to initially decrease its altitude and reach the altitude limit. This way, the two orbit planes converge at an even faster rate and the chaser waits for the target orbit plane to get sufficiently close before increasing its altitude again. For the same reason of the negative Δa_0 transfer, in all $\Delta\imath_0$ cases the inclination is initially reduced. As already discussed for the negative change of altitude case, the greater part of the out-of-plane thrusting effort for the optimal $\Delta\imath_0 = -1°$ trajectory is initially used to decrease the inclination of the orbit rather than to attain the required RAAN change. Since the initial RAAN rate of change is slower, in the negative $\Delta\imath_0$ case the chaser spends more time changing Ω at the altitude limit and the time-of-flight is longer than that of the other two cases.

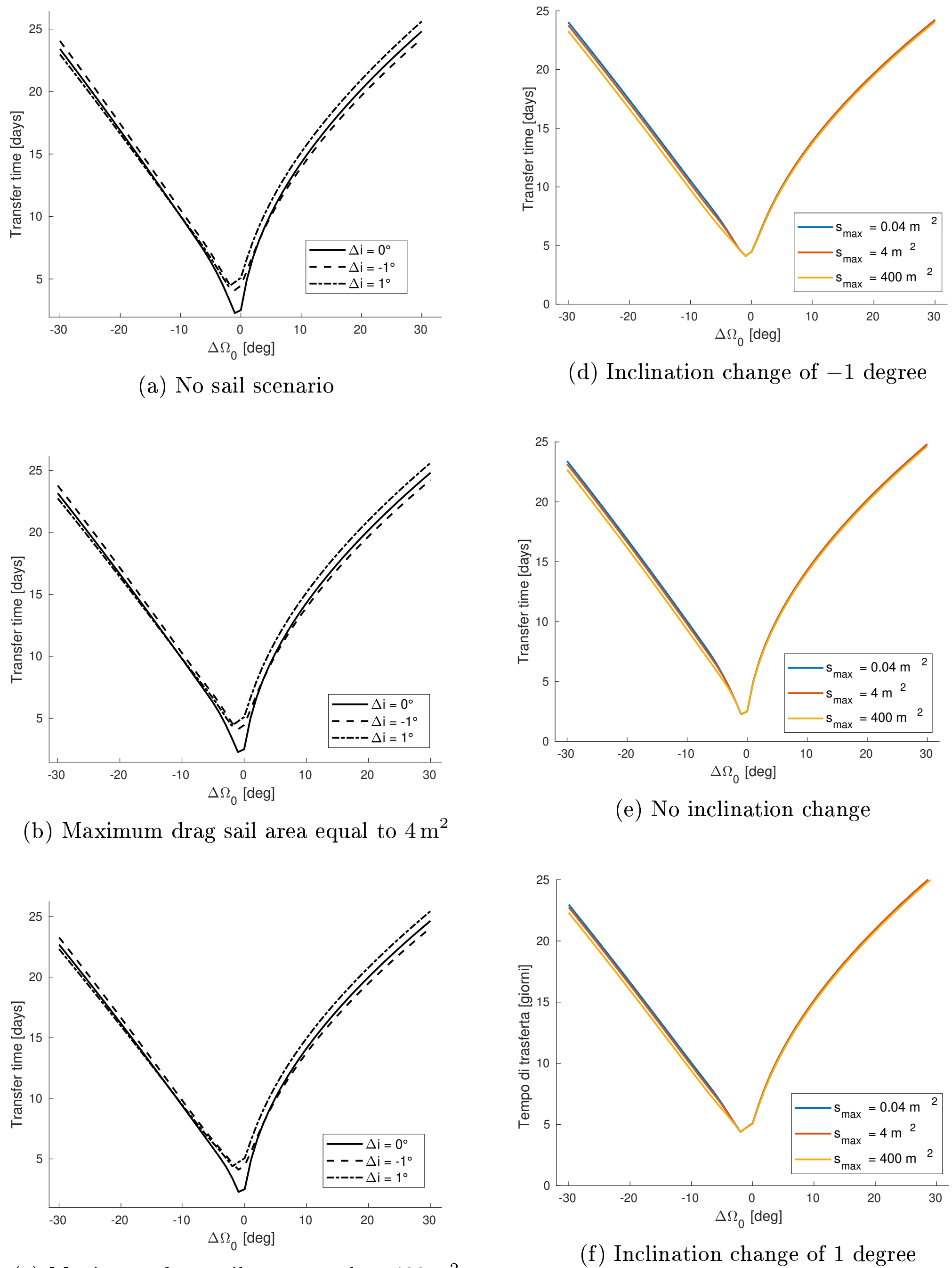
On the other hand, in the $\Delta\Omega_0 = 30°$ case the two orbit planes are initially drifting away from each other, as the chaser orbit plane regresses at a faster rate. In order to reverse this process, the altitude of the chaser is increased beyond the target altitude. As a consequence, in this new configuration the line of nodes of the chaser regresses at a slower rate and the target orbit can catch up. If the required $\Delta\imath_0$ is negative, while the S/C is increasing its altitude Ω is changed more efficiently and $\Delta\Omega$ decreases at a faster rate; the S/C can reach a lower maximum altitude than that of the other $\Delta\imath_0$ cases and thus can start earlier its descent.

From figures 6.4d, 6.4e and 6.4f it can be readily seen that the use of a drag sail doesn't bring much improvement to these transfers. Table 6.2 shows that when $\Delta\Omega_0 < \Delta\Omega^*$ the time-of-flight of the 4 m^2 scenario is on average six hours less than that of the no-sail scenario; the 400 m^2 scenario further improves the trip time by an average of twelve hours. On the other hand, when $\Delta\Omega_0 > \Delta\Omega^*$ there is almost no difference of minimum-time between the no-sail scenario and the $4\,\mathrm{m}^2$ scenario, while the $400\,\mathrm{m}^2$ scenario is barely better by an average of four hours.

The semimajor axis and inclination time histories for the $\Delta\imath_0 = 1°$ case are shown in figure 6.6. In the $\Delta\Omega_0 = -30°$ case (figure 6.6a), the small improvement brought by the sail is due to the brief initial altitude decrease. In the $\Delta\Omega_0 = 10°$ case, the atmospheric density at the altitudes of the transfers is so low that the deployment of the sail has a negligible impact on the trajectories, even for the $400\,\mathrm{m}^2$ scenario.

Table 6.2: Minimum transfer time and corresponding propellant consumption for transfers with an altitude change of $+200\,\mathrm{km}$

		No drag sail		$S_{\max} = 4\,\mathrm{m}^2$		$S_{\max} = 400\,\mathrm{m}^2$	
$\Delta\imath$	$\Delta\Omega_0$	t_f	m_p [kg]	t_f	m_p [kg]	t_f	m_p [kg]
-1°	-30°	24 d 1h	0.847	23 d 19h	0.838	23 d 7h	0.820
	-20°	17 d 10h	0.614	17 d 4h	0.604	16 d 15h	0.587
	-10°	10 d 13h	0.371	10 d 6h	0.362	9 d 18h	0.344
	0°	4 d 11h	0.157	4 d 11h	0.157	4 d 11h	0.157
	10°	13 d 22h	0.491	13 d 22h	0.491	13 d 18h	0.485
	20°	19 d 16h	0.693	19 d 16h	0.693	19 d 12h	0.687
	30°	24 d 6h	0.854	24 d 6h	0.854	24 d 1h	0.848
0°	-30°	23 d 10h	0.825	23 d 4h	0.816	22 d 16h	0.800
	-20°	16 d 21h	0.595	16 d 15h	0.586	16 d 4h	0.570
	-10°	10 d 1h	0.353	9 d 19h	0.345	9 d 8h	0.329
	0°	2 d 12h	0.088	2 d 12h	0.088	2 d 12h	0.088
	10°	14 d 7h	0.504	14 d 7h	0.504	14 d 3h	0.498
	20°	20 d 4h	0.711	20 d 4h	0.711	19 d 24h	0.705
	30°	24 d 20h	0.874	24 d 19h	0.874	24 d 15h	0.868
1°	-30°	22 d 23h	0.809	22 d 18h	0.802	22 d 7h	0.786
	-20°	16 d 15h	0.585	16 d 10h	0.578	15 d 23h	0.562
	-10°	10 d 1h	0.354	9 d 20h	0.346	9 d 10h	0.331
	0°	5 d 2h	0.180	5 d 2h	0.180	5 d 2h	0.178
	10°	15 d 3h	0.533	15 d 3h	0.533	14 d 23h	0.527
	20°	20 d 23h	0.739	20 d 23h	0.739	20 d 19h	0.733
	30°	25 d 15h	0.903	25 d 15h	0.902	25 d 10h	0.896



(a) No sail scenario

(d) Inclination change of -1 degree

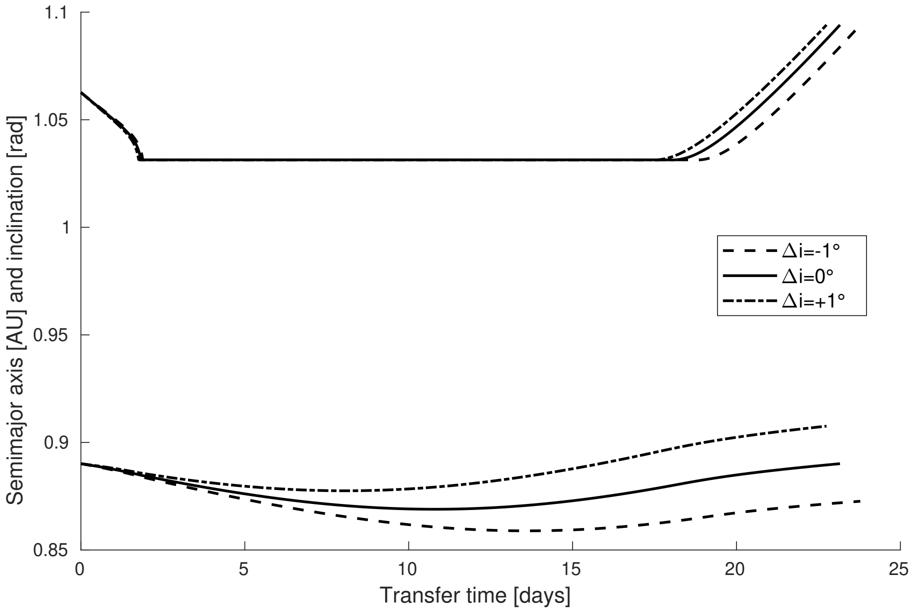
(b) Maximum drag sail area equal to $4\,\mathrm{m}^2$

(e) No inclination change

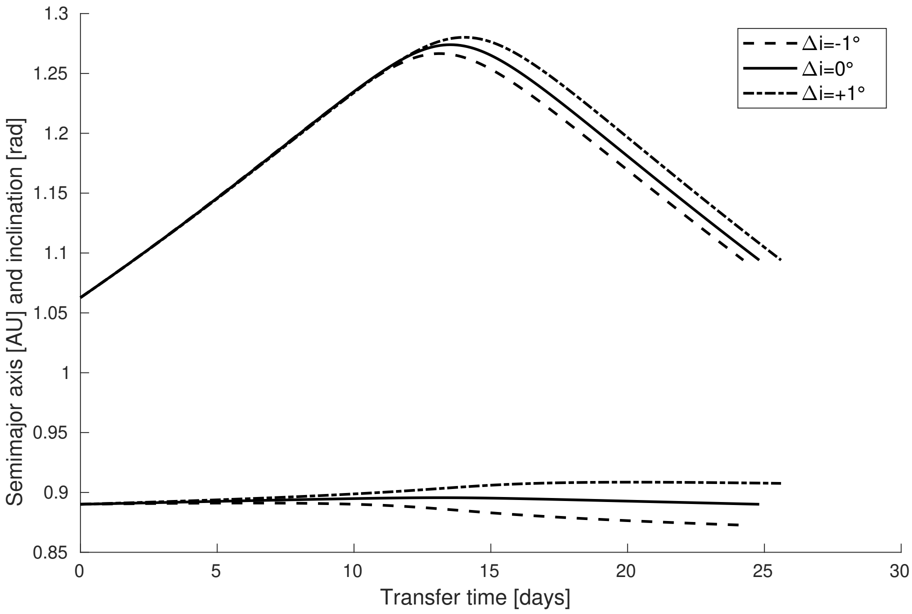
(c) Maximum drag sail area equal to $400\,\mathrm{m}^2$

(f) Inclination change of 1 degree

Figure 6.4: Minimum transfer time as a function of the target's initial RAAN for transfers with an altitude change of $+200\,\mathrm{km}$

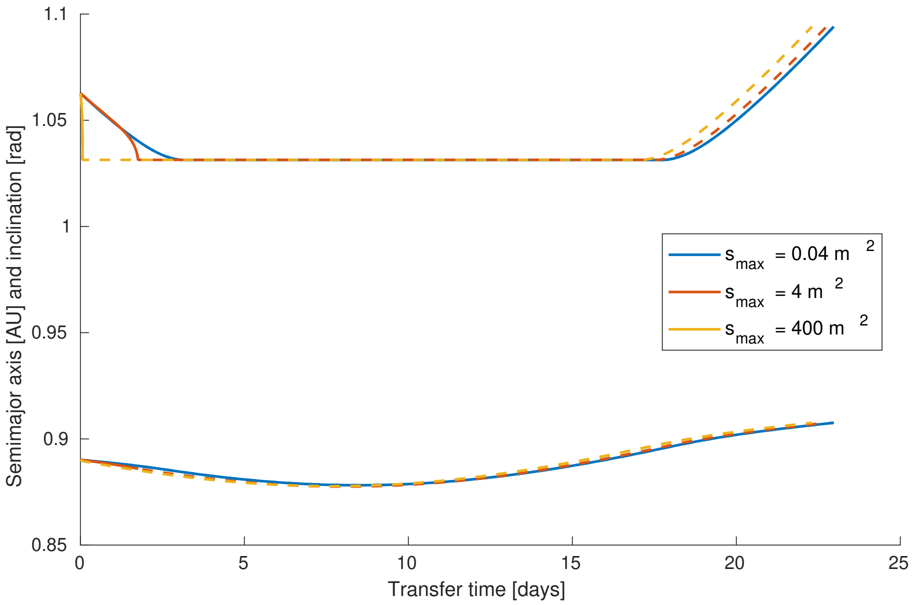


(a) $\Delta\Omega_0 = -30°$

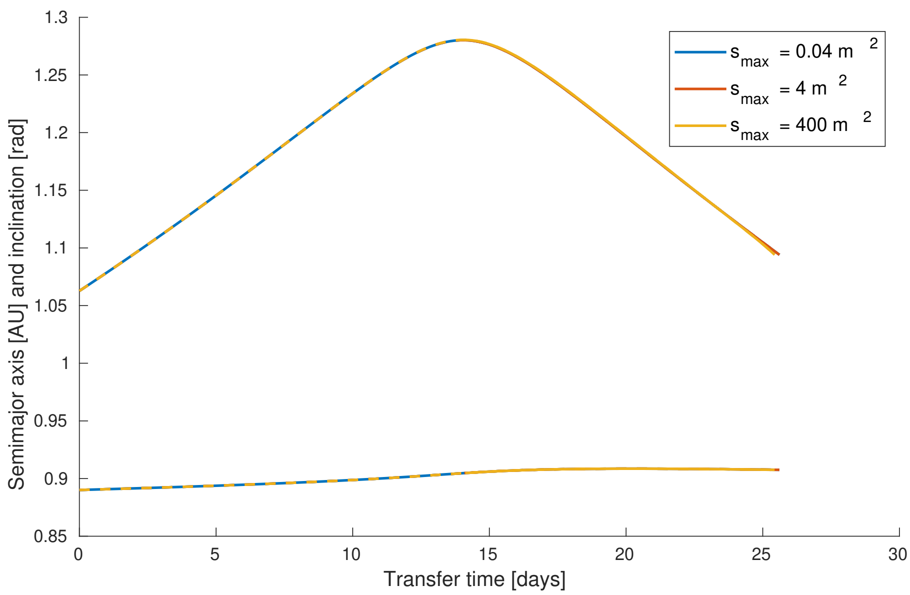


(b) $\Delta\Omega_0 = +30°$

Figure 6.5: Minimum-time trajectories for $\Delta a_0 = +200\,\mathrm{km}$ and $S_{\max} = 4\,\mathrm{m}^2$



(a) $\Delta\Omega_0 = -30°$



(b) $\Delta\Omega_0 = +30°$

Figure 6.6: Minimum-time trajectories for $\Delta a_0 = +200\,\mathrm{km}$ and $\Delta\imath_0 = 1°$

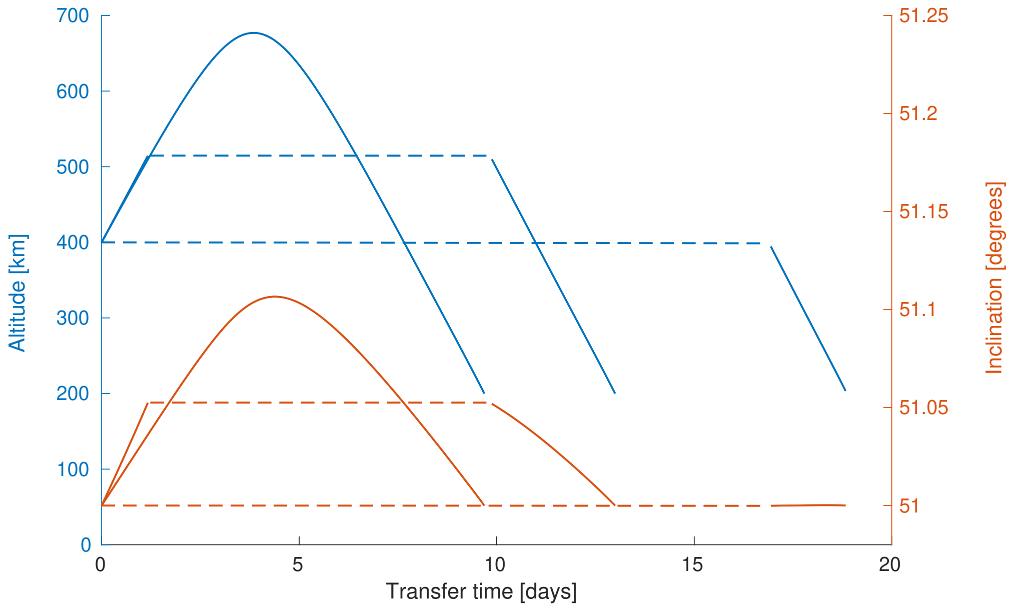


Figure 6.7: Minimum-propellant trajectories for different trip times in the no-sail scenario ($\Delta a_0 = -200\,\mathrm{km}$, $\Delta \imath_0 = 0°$ and $\Delta\Omega_0 = 10°$)

6.2 Minimum-propellant solutions

When a given transfer is carried out over a longer period of time than the minimum, coasting arcs can appear and propellant consumption is reduced. Depending on whether the two orbit planes are initially moving closer to each other or drifting apart, the minimum propellant solutions are driven by two different strategies.

The case of a $-200\,\mathrm{km}$ change of semimajor axis, same inclination and $\Delta\Omega_0 = 10°$ is used as an example to show the first strategy, that is when the two orbit planes are initially moving closer to each other. Figure 6.7 shows the semimajor axis and inclination time histories in the no-sail scenario. Three different trip times are shown: the minimum-time solution, the global-optimal solution for propellant consumption and a third solution between the two. If the trip time is increased from the minimum-time solution, the trajectory is characterized by a coasting arc that separates two burns. The S/C can reach lower maximum altitude and maximum inclination than those of the minimum-time trajectory. After the thruster is switched off (dashed line in the figure), the S/C waits can wait for the J2 perturbation to bring the orbit planes closer (a 13-day trip is used as an example in fig. 6.7) and then perform a second burn to reach the target orbit. In the presented case, for a time of flight close to 20 days, the initial burn disappears and the waiting orbit becomes the initial one. In this case, the global-optimum trajectory can be performed when the target orbit is close enough. For comparison, figure 6.8 shows the trajectories for the $4\,\mathrm{m}^2$ scenario. The periods of time when the sail is deployed are denoted by a bigger width of the

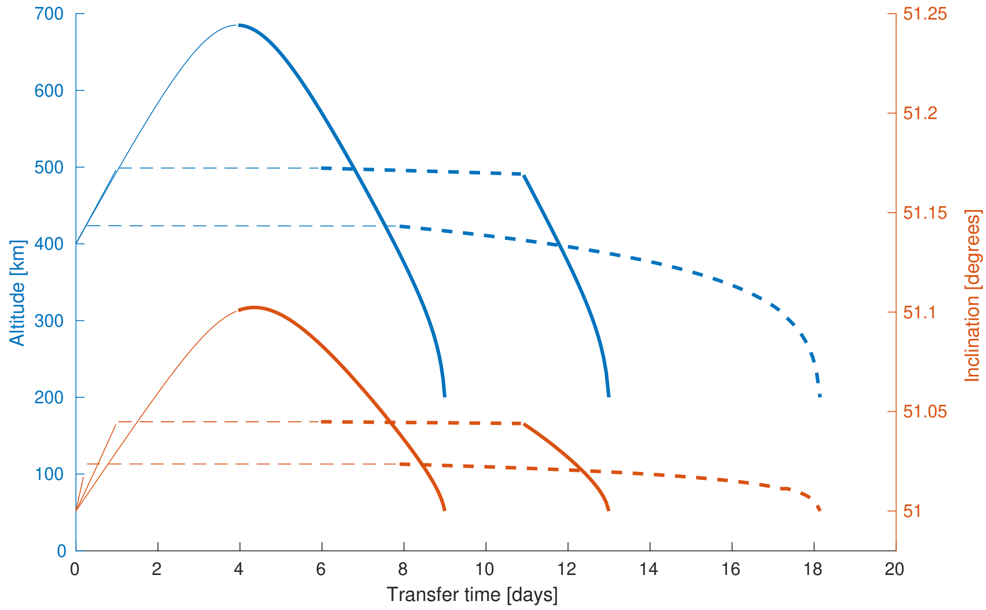


Figure 6.8: Minimum-propellant trajectories for different trip times in the $4\,\mathrm{m}^2$ scenario ($\Delta a_0 = -200\,\mathrm{km}$, $\Delta \imath_0 = 0°$ and $\Delta\Omega_0 = 10°$)

line, whereas coasting arcs are still denoted by a dashed line. The structure of the solutions is quite similar and when coasting arcs appear the sail is deployed before the second burn. As opposed to the previous scenario, the deployment of the sail enables the S/C to perform an aeroassisted descent without the use of thrust. This was not possible for no-sail scenario because the orbital decay of the S/C to the target orbit would have required more time than what it takes for J2 to nullify $\Delta\Omega$. Therefore, in the $4\,\mathrm{m}^2$ scenario if the available time is increased beyond a certain value the second burn disappears and the descent is performed thanks to the effect of drag only. It is interesting to note that the first burn doesn't disappear for the global-minimum trajectory. Since drag decreases the inclination of the orbit (as described in Chapter 4), $\imath$ must be initially increased in order to account for the subsequent decrease after the sail deployment. In addition, the global-minimum trajectory also prescribes a small initial increase of semimajor axis during the first burn.

The same strategy of waiting for the target orbit to come closer, spending less propellant by giving up some transfer time, can be performed in the case of a positive change of semimajor axis and negative $\Delta\Omega_0$. For example, a case of $+200\,\mathrm{km}$ change of semimajor axis, same inclination and $\Delta\Omega_0 = -10°$ is shown in figure 6.9. The minimum propellant trajectory requires the S/C to reach the altitude limit in order to increase the J2 perturbation as much as possible before performing the final ascent. The inclination is also slightly decreased for the same reason. If sufficient time is available, the descent to the altitude limit can be performed by deploying the sail

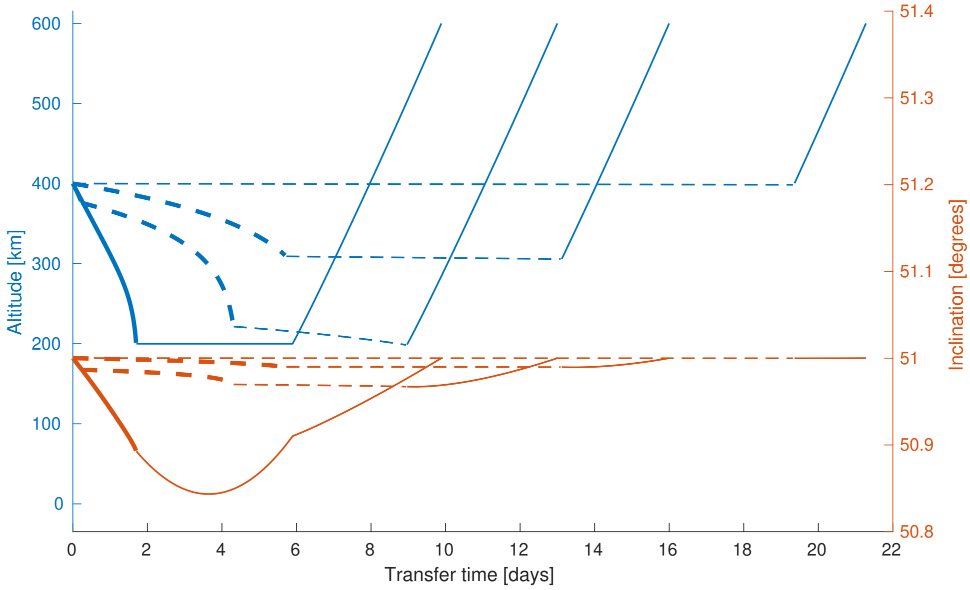


Figure 6.9: Minimum-propellant trajectories for different trip times in the $4\,\mathrm{m}^2$ scenario ($\Delta a_0 = +200\,\mathrm{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = -10°$)

and initially firing the thruster, and then turning off the thruster while still having the sail deployed. Beyond a certain trip time value, the initial burn disappears, the sail is initially deployed and the altitude limit is not reached. The global minimum is reached when the sail is never deployed and the S/C slowly decays before the final burn. Since the inclination is also slightly decreased during the coasting arc, some thrusting effort is spent to change $\imath$ during the final maneuver.

The second strategy is implemented when the two orbit planes are drifting apart. The case of $+200\,\mathrm{km}$ change of semimajor axis, same inclination and $\Delta\Omega_0 = 10°$ is used as an example to present the structure of the solutions. Figure 6.7 shows the semimajor axis and inclination time histories in the $4\,\mathrm{m}^2$ scenario. As already discussed for the minimum-time solutions, the semimajor axis must be initially increased beyond the target value in order to reverse the relative motion of the two orbits. However, if more time is available to complete the transfer, the thruster can be switched off before the maximum altitude is reached and the chaser can wait at a lower altitude for J2 to move the two orbit planes closer. The lower the altitude at which the coasting arc starts, the smaller the value of the orbit nodes relative motion. If the effects of aerodynamic drag were not accounted for, the theoretic trip time for the global optimum solution would grow to infinity, as the waiting orbit would be infinitesimally higher than the target. In reality, the global optimum solution requires the S/C to reach an altitude of approximately 606 km and then turn off the engine. After more than 3 years the altitude reaches 602 km and the sail is deployed for 6 days in order to bring semimajor axis and inclination at target values.

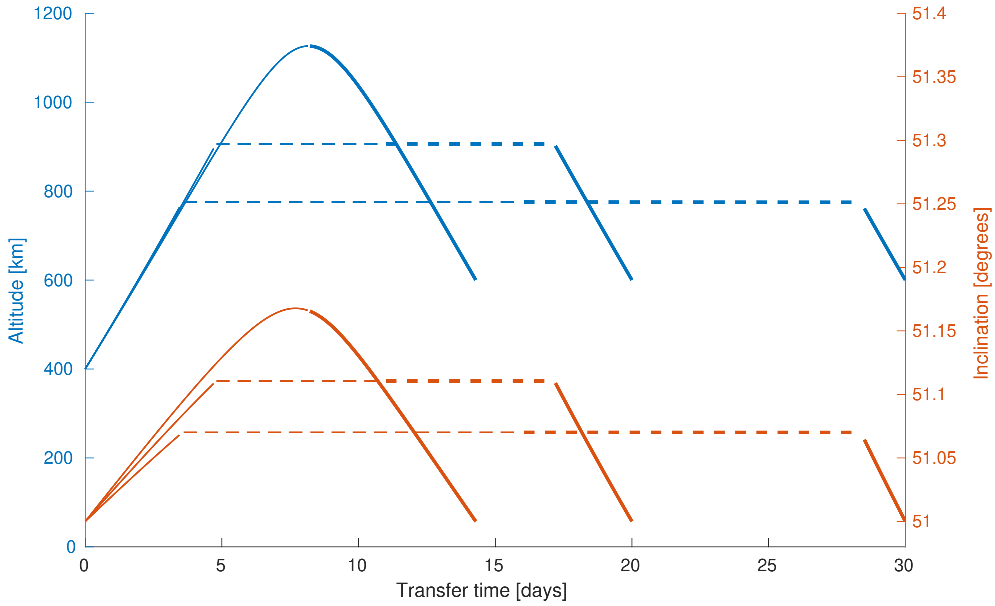


Figure 6.10: Minimum-propellant trajectories for different trip times in the $4\,\mathrm{m}^2$ scenario ($\Delta a_0 = +200\,\mathrm{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = 10°$)

This maneuver requires a total of approximately 1215 days to be completed, with a propellant consumption of 0.069 kg.

An alternative strategy in case of orbits that are drifting apart is favoring the initial relative node motion. Namely, instead of bringing $\Delta\Omega$ to zero, one can increase it to 2π and the structure of the trajectories become similar to that of figure 6.9. The minimum propellant solution for this strategy is shown in figure 6.11: the sail is never deployed and the chaser coasts for almost 600 days before turning on the engine to reach the target altitude and inclination. It has to be noted that this strategy can't reach the global optimum solution. As a matter of fact, the required change of semimajor axis and inclination after the long coasting arc is much bigger than the one of the previous strategy. For the case shown the propellant consumption is roughly 0.103 kg. This value can be compared to the one of the previous strategy for the same trip time (598 days), which however is about 0.070 kg. Therefore, the alternative strategy performs worse and it would be meaningless to adopt it.

6.2.1 Propellant consumption analysis

In order to draw useful conclusions on the use of drag sails for minimum propellant trajectories, let us now refer to the presented cases and study how aeroassisted maneuvers can improve the solutions. The case of orbits with lines of nodes initially moving closer is presented first. Figure 6.12a shows the propellant consumption for the three scenarios as a function of transfer time in the case of $\Delta a_0 = -200\,\mathrm{km}$,

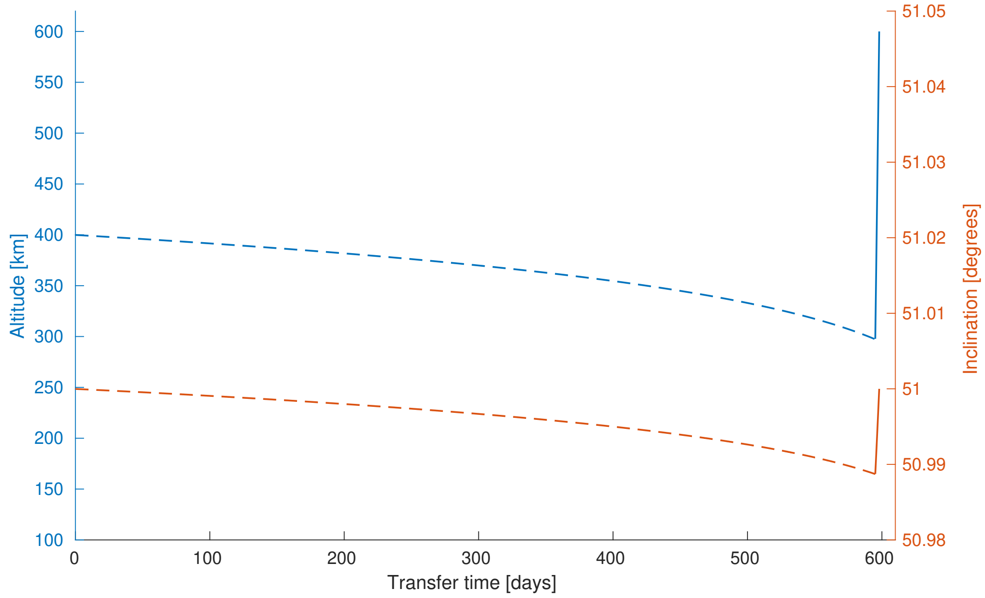
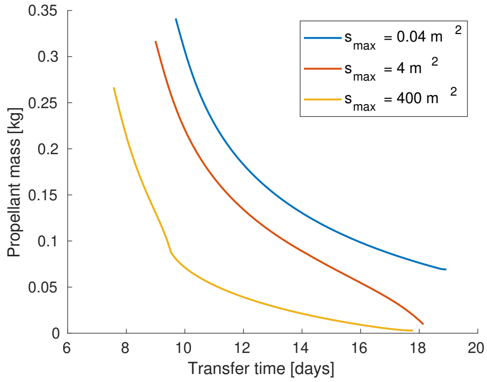


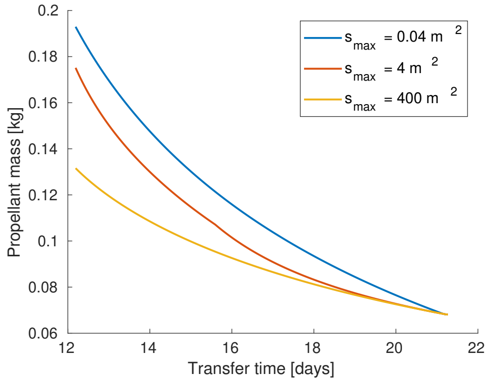
Figure 6.11: Minimum propellant trajectory for $\Delta\Omega_f = 2\pi$ in the $4\,\mathrm{m}^2$ scenario ($\Delta a_0 = +200\,\mathrm{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = 10°$)

$\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = 10°$, while figure 6.12b presents the analogous plot in the case of $\Delta a_0 = +200\,\mathrm{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = -10°$. In addition, table C.1 from Appendix C displays some values from the plots.

In the case of negative change of semimajor axis, the $4\,\mathrm{m}^2$ scenario requires from 30 to 60 percent less propellant than the no sail scenario for a given transfer time, whereas for the $400\,\mathrm{m}^2$ scenario the savings grow from 80 to 90 percent. In addition, both sail scenarios have a global minimum around zero, as the propellant is only used for the initial inclination increase. Particularly, the $4\,\mathrm{m}^2$ scenario is characterized by a global optimum of 0.009 kg of propellant consumed over a transfer time of 18 days and 3 hours. The $400\,\mathrm{m}^2$ scenario takes only 2 grams less of propellant to perform the global optimum transfer over 17 days and 18 hours. On the other hand, the global minimum for the no sail scenario is 0.069 kg and the corresponding trip time is 18 days and 22 hours. The technologically feasible scenario, i.e. $4\,\mathrm{m}^2$ of drag sail area, not only improves the trip time of the propellant global minimum trajectory by 19 hours with respect to the no sail scenario, but it also improves consumption by almost 90%. It has to be noted that the real advantage brought by the sail appears when the transfer can be completed without a second burn. Using the presented $4\,\mathrm{m}^2$ scenario as an example, this happens roughly from a 17-day trip up to the global optimum. Since the time minimum to complete the transfer is approximately 9 days with a propellant consumption of 0.317 kg, depending on the mission it could make sense to perform the propellant global minimum maneuver. This way, the transfer time would double from 9 to 18 days, but the propellant consumption would



(a) $\Delta a_0 = -200\,\mathrm{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = 10°$



(b) $\Delta a_0 = +200\,\mathrm{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = -10°$

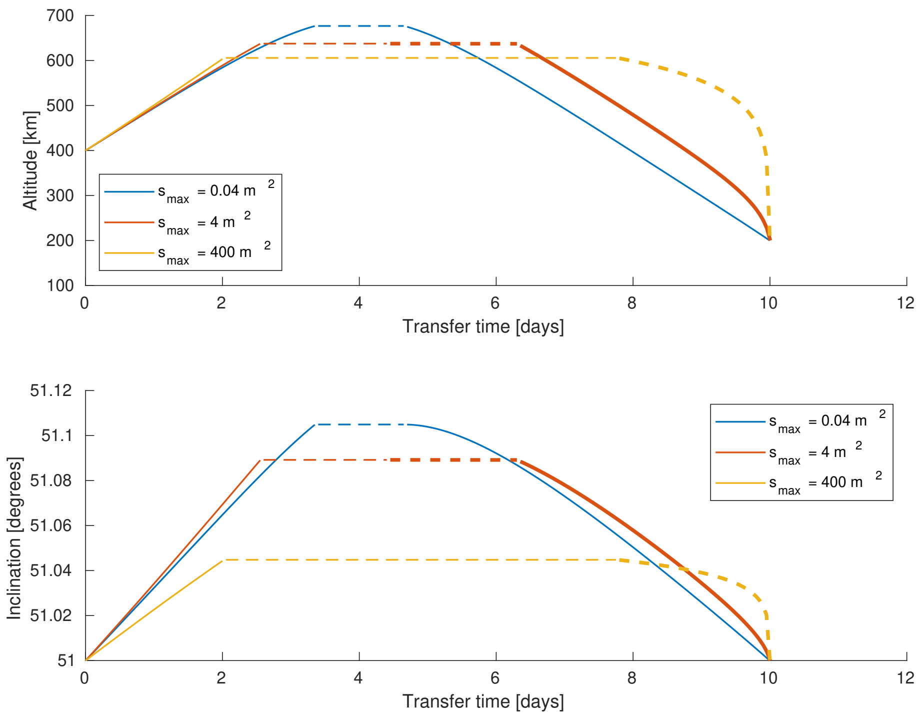
Figure 6.12: Minimum-propellant as a function of transfer time

drop almost completely. This could be for example the case of an ADR mission, which would require to collect as much debris as possible given a certain amount of propellant.

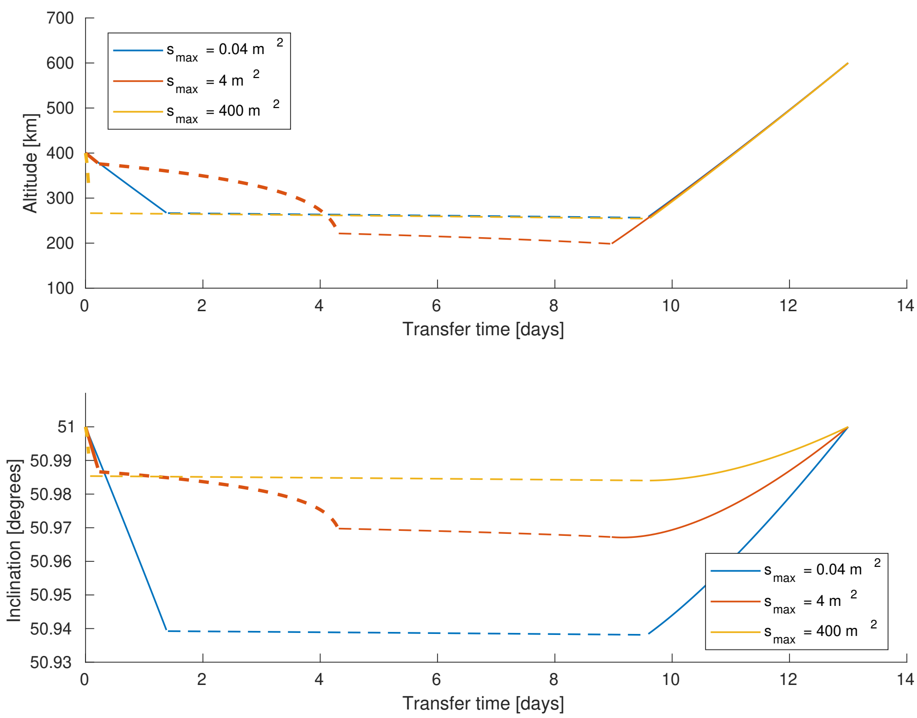
Opposite conclusions can be drawn for the case of positive change of semimajor axis. In fact, for both sail scenarios the performance improvement in terms of propellant consumption decreases as the transfer time increases. Furthermore, the curves converge to the same global optimum of 0.069 kg, as the corresponding solutions require the sail to be never deployed and the trajectories of all scenarios are identical. Given a certain trip time, the propellant consumption improvement with respect to the no sail scenario is around 10% in the 4 m^2 scenario and reaches a maximum of almost 40% in the 400 m^2 scenario. Once again, depending on the mission it could make sense to perform the global optimum trajectory, as for ADR. On the other hand, for missions that require trade-offs between transfer time and propellant consumption, it has to be kept into account that the improvements brought by a drag sail decrease as the trip time approaches the one of the global optimum trajectory. This could be for example the case of on-orbit servicing.

Figure 6.13 highlights how increasing the maximum sail area value improves the propellant consumption for a given trip time. Let us discuss the negative altitude change first (figure 6.13a). To begin with, the descent is aeroassisted and requires less propellant. In addition, for higher values of $S_{\max}$ it can also be performed at a faster rate. This means that the coasting arc can be extended over a longer period of time, and moved at a lower altitude and at a lower inclination,. Consequently, the first and second burn are shorter. In the presented case, i.e. a 10-day trip, a S/C equipped with a 4 m^2 sail would deploy it before the second burn, while a 400 m^2 value would be enough to not require a second burn at all.

As for the positive altitude change case, a 13-day trip is presented. In the 4 m^2 scenario, the S/C initially fires the thruster and deploys the sail at the same time.



(a) $\Delta a_0 = -200\,\mathrm{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = 10°$



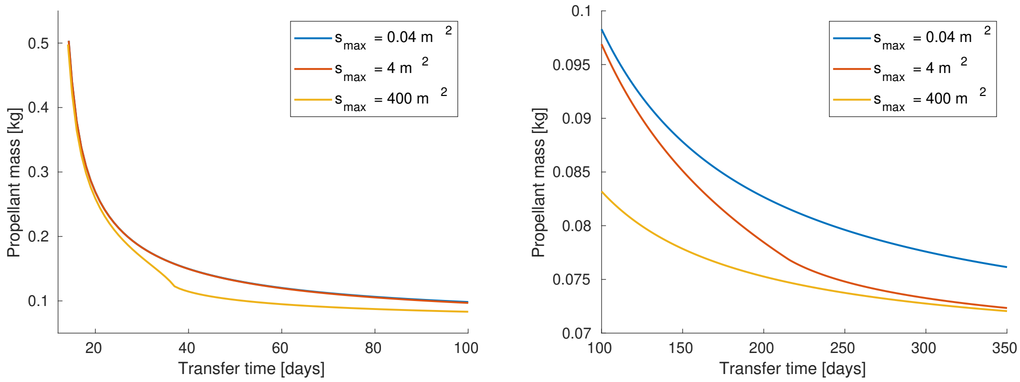
(b) $\Delta a_0 = +200\,\mathrm{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = -10°$

Figure 6.13: Minimum-propellant trajectories for a given trip time

The descent is similar to that of the no-sail scenario; however, the S/C turns off the thruster at an earlier time and at a higher altitude. Nevertheless, the descent is continued to a lower altitude and the sail is then furled. Although the second burn is longer, the propellant saved during the descent results in an overall lower expenditure. Interestingly, in the $400\,\mathrm{m}^2$ scenario the initial descent is performed without the use of thrust and the S/C coasts at approximately the same altitude of the no sail scenario. Since there's no initial burn and the final one takes roughly the same amount of time, the propellant consumption is null for the first and just about the same for the latter.

Let us now study the case of orbits with lines of nodes initially drifting apart. Figure 6.14 presents the propellant consumption for the three scenarios as a function of transfer time, in the case of $\Delta a_0 = +200\,\mathrm{km}$, $\Delta \imath_0 = 0°$ and $\Delta\Omega_0 = 10°$. The faint aerodynamic effects at the altitudes of the transfer and the brief time over which the descent is performed result in modest improvement of the two drag sails scenarios. For trip times under 100 days (figure 6.14a) small improvements can be seen only for the $400\,\mathrm{m}^2$ scenario, whereas the $4\,\mathrm{m}^2$ is almost identical to the no sail scenario. The trajectories for a trip time of 30 days and for a trip time of 40 days are shown in figure 6.15. While the trajectory for the first two scenarios is almost the same, the $400\,\mathrm{m}^2$ value is large enough to be briefly exploited during the final descent; the inclination of the waiting orbit can be slightly lower and the altitude slightly higher. For the 30-day transfer the second burn is somewhat shorter, while the 40-day maneuver is long enough to perform the descent without the use of thrust. For trip times from 100 to 350 days (figure 6.14b) the $4\,\mathrm{m}^2$ scenario goes from being similar to the no-sail scenario to being almost equivalent to the $400\,\mathrm{m}^2$ scenario. In fact, as the trip time increases, the orbital decay in the $4\,\mathrm{m}^2$ scenario becomes more and more pronounced than that of the no-sail scenario. The second burn can then be slightly shorter, analogously to the 30-day trip in the $400\,\mathrm{m}^2$ scenario. For even longer transfers (over 200 days), the second burn disappears in the $4\,\mathrm{m}^2$ scenario as well. However, the semimajor axis decrease happens at a much slower rate than that of the $400\,\mathrm{m}^2$ scenario. As a consequence, the S/C spends more time at lower altitudes, where the effect of J2 is higher. In order to account for this, the waiting orbit is higher, hence the difference of propellant consumption. For trip times over 350 days the descents of the two drag sail scenarios become very brief and similar, and so does the propellant expenditure. From a certain point on, the curves begin to converge at the global optimum of 0.069 kg, common to all three cases since the trajectory doesn't require the deployment of the sail. Table C.2 from Appendix C shows the results in detail.

It has to be noted that even for long mission such as ADR, a single debris to debris transfer can't be extended over prohibitively long times; an acceptable transfer could be in the order of 60 days. In the case presented, having a $4\,\mathrm{m}^2$ drag sail would be



(a) Detail from minimum-time to 100 days

(b) Detail from 100 to 350 days

Figure 6.14: Minimum-propellant as a function of transfer time ($\Delta a_0 = +200\,\text{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = 10°$)

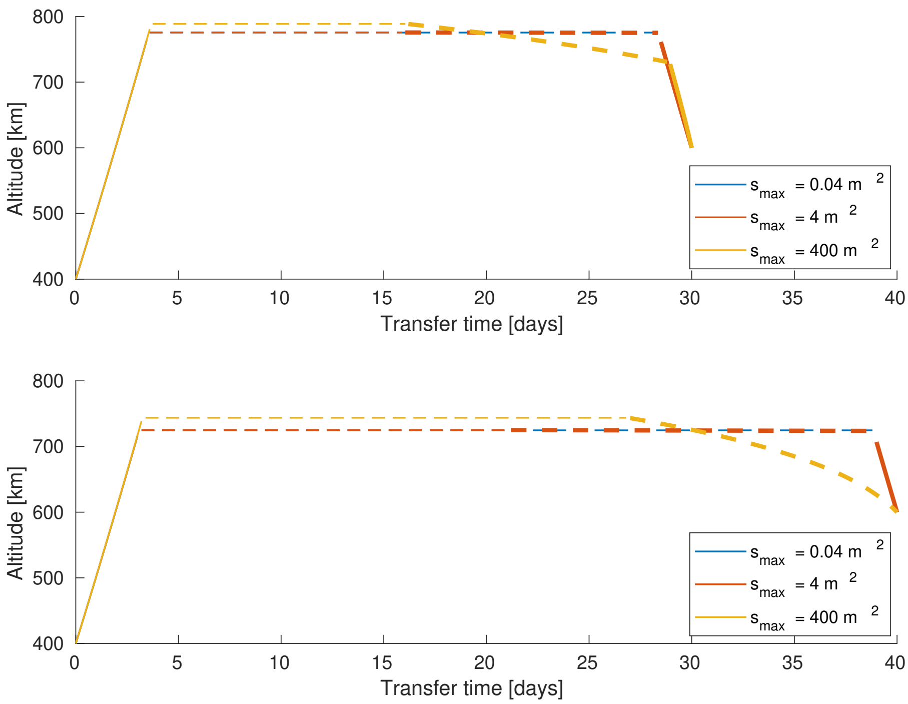


Figure 6.15: Minimum-propellant trajectories for trip times of 30 and 40 days ($\Delta a_0 = +200\,\text{km}$, $\Delta\imath_0 = 0°$ and $\Delta\Omega_0 = 10°$)

almost completely useless for such trip time. However, the propellant consumption improvement highly depends on the initial and final orbit, as flying through thicker density values enhances the aeroassisted maneuvers. As a general trend, in the case positive altitude change transfers, the use of drag sails is beneficial when the line of nodes of the target orbit is initially moving closer to the one of the chaser or, when the lines of nodes are drifting apart, if the transfer is performed between orbits below 400 km of altitude. On the other hand, it has been shown that negative altitude change transfers can be carried out with nearly-zero-consumption maneuvers. As for minimum-time trajectories, negative altitude change transfers are improved by a few hours when a $4\,\mathrm{m}^2$ drag sail is employed and by a a few days when $S_{\max} = 400\,\mathrm{m}^2$. In addition, increasing the maximum frontal area doesn't improve the minimum time of positive altitude change transfers by much. In light of this, the use of a drag sail wouldn't make much sense for missions that require the S/C to reach the target orbit as fast as possible (e.g. a satellite that must reach its operative orbit). As a matter of fact, even a light sail would take a toll on the cost of the mission and, at the same time, take up some volume and mass that would be otherwise assigned to other subsystems. On the flip side, OOS and ADR missions could exploit drag sails to a much greater extent. Even though the scope of this thesis was not to give a direct example of such missions, the results obtained show that a drag-enhancement system should be investigated during mission analysis.

Conclusions

The optimization of space trajectories in low Earth orbit is becoming of extreme interest due to several new mission concepts, such as active debris removal, on-orbit servicing and small-sat deployment from orbiting platforms. This thesis provides a fast and accurate method for the evalutation of these transfers. In particular, the effects of the Earth's gravitational harmonic J2 coefficient and aerodynamic drag are kept into account in order to obtain meaningful results.

After showing how to build a mathematical model to approximate the aforementioned perturbations, an indirect optimization approach has been applied. This method transforms the optimization problem into a boundary value problem, which is solved by means of shooting procedures. The use of drag sails is investigated by treating the frontal area of the spacecraft as a control variable, in addition to thrust magnitude and direction. The analysis of a case study has displayed the structure of the solutions to the minimum-time and minimum-propellant problems. Depending on the parameters that define the transfer, i.e. relative orbital elements and available trip time, the enhancement of aerodynamic drag may be highly beneficial in terms of propellant expenditure. On the flip side, time of flight is improved significantly only for relatively large values of maximum frontal area. Therefore, this thesis offers a framework for fast trade-off studies on the use of drag sails for future missions.

Although the atmospheric model used in this work is not particularly accurate, the method proposed is readily extendible to deal with more sophisticated models. Proper phasing of the chaser spacecraft to rendezvous with the target has not been accounted for. Moreover, the changes of semimajor axis and inclination of the target have been neglected, as the case study assumed it performed maneuvers in opposition to the effects of aerodynamic drag. In light of this, future studies may use the described method to define optimal strategies for a specific active debris removal or on-orbit servicing mission. A more detailed atmospheric model may be easily plugged in and the rendezvous phase may be accounted for. Finally, further complexity may be added by removing the almost-circular orbit approximation and considering the eccentricity equation and the effect of J2 on the line of apses.

Bibliography

[1] Roger R Bate, Donald D Mueller, Jerry E White, and William W Saylor. *Fundamentals of astrodynamics.* Courier Dover Publications, 2020.

[2] Bruce Bowman, W. Kent Tobiska, Frank Marcos, Cheryl Huang, Chin Lin, and William Burke. *A New Empirical Thermospheric Density Model JB2008 Using New Solar and Geomagnetic Indices.*

[3] Arthur Earl Bryson. *Applied optimal control: optimization, estimation and control.* CRC Press, 1975.

[4] Lorenzo Casalino and Guido Colasurdo. *Improved Edelbaum's Approach to Optimize LEO-GEO Low-Thrust Transfers.*

[5] Guido Colasurdo and Lorenzo Casalino. *Indirect Methods for the Optimization of Spacecraft Trajectories*, pages 141–158. Springer New York, New York, NY, 2013.

[6] Guido Colasurdo and Lorenzo Casalino. *Tentative Solutions for Indirect Optimization of Spacecraft Trajectories*, pages 87–102. Springer International Publishing, Cham, 2016.

[7] Bruce A. Conway. *The Problem of Spacecraft Trajectory Optimization*, pages 1–15. Cambridge Aerospace Series. Cambridge University Press, 2010.

[8] Theodore N. Edelbaum. Propulsion requirements for controllable satellites. *ARS Journal*, 31(8):1079–1089, 1961.

[9] Theodore N. Edelbaum. Theory of maxima and minima. In George Leitmann, editor, *Optimization Techniques*, volume 5 of *Mathematics in Science and Engineering*, pages 1–32. Elsevier, 1962.

[10] L. Johnson, D. Alhorn, Mark Boudreaux, J. Casas, D. Stetson, and R. Young. Solar and drag sail propulsion: From theory to mission implementation. 2014.

[11] Jean A. Kechichian. Optimal altitude-constrained low-thrust transfer between inclined circular orbits. *The Journal of the Astronautical Sciences*, 54(3):485–503, December 2006.

[12] Jean A. Kechichian. Optimal low-thrust transfer in general circular orbit using analytic averaging of the system dynamics. *The Journal of the Astronautical Sciences*, 57(1):369–392, January 2009.

[13] Jean Albert Kechichian. Reformulation of edelbaum's low-thrust transfer problem using optimal control theory. *Journal of Guidance, Control, and Dynamics*, 20(5):988–994, 1997.

[14] Donald J. Kessler and Burton G. Cour-Palais. Collision frequency of artificial satellites: The creation of a debris belt. *Journal of Geophysical Research: Space Physics*, 83(A6):2637–2646, 1978.

[15] Donald E Kirk. *Optimal control theory: an introduction*. Courier Corporation, 2004.

[16] Louis Komzsik. *Applied calculus of variations for engineers*. CRC Press, 2019.

[17] Joerg Kreisel. On-orbit servicing (oos): Issues & commercial implications. 09 2003.

[18] Dazhong Lao and Shanshan Zhao. *Fundamental Theories and Their Applications of the Calculus of Variations*. Springer, 2020.

[19] J.-C. Liou and N. L. Johnson. Risks in space from orbiting debris. *Science*, 311(5759):340–341, 2006.

[20] United States. National Oceanic, Atmospheric Administration, and United States. Air Force. *US standard atmosphere, 1976*, volume 76. National Oceanic and Atmospheric Administration, 1976.

[21] L.F. Shampine, M.K. Gordon, L.F. Gordon, M.K. Gordon, W. H. Freeman, and Company. *Computer Solution of Ordinary Differential Equations: The Initial Value Problem*. W. H. Freeman, 1975.

[22] Daniel Tabak. Introduction to control theory including optimal control, d. n. burghes and a. graham, ellis horwood ltd. *Optimal Control Applications and Methods*, 2(2):203–204, 1981.

[23] Andrew Tatsch, Norman Fitz-Coy, and Svetlana Gladun. On-orbit servicing: A brief survey. In *Proceedings of the IEEE International Workshop on Safety, Security, and Rescue Robotics*, pages 276–281, 2006.

[24] Craig Underwood, Andrew Viquerat, Mark Schenk, Ben Taylor, Chiara Massimiani, Richard Duke, Brian Stewart, Simon Fellowes, Chris Bridges, Guglielmo Aglietti, Berry Sanders, Davide Masutti, and Amandine Denis. Inflatesail de-orbit flight demonstration results and follow-on drag-sail applications. *Acta Astronautica*, 162:344–358, 2019.

[25] David A Vallado. *Fundamentals of astrodynamics and applications*, volume 12, chapter 8, pages 559–569. Springer Science & Business Media, 2001.

[26] Ting Wang. Analysis of debris from the collision of the cosmos 2251 and the iridium 33 satellites. *Science & Global Security*, 18(2):87–118, 2010.

[27] William E. Wiesel and Salvatore Alfano. Optimal many-revolution orbit transfer. *Journal of Guidance, Control, and Dynamics*, 8(1):155–157, 1985.

Appendix A

Vector and matrix operations

This appendix lists useful vector and matrix operations that will be used in this thesis.

A.1 Vector operations

Scalar product The scalar product of two vectors $\mathbf{a}$ and $\mathbf{b}$, denoted by $\mathbf{a} \cdot \mathbf{b}$, is defined as

$$\mathbf{a} \cdot \mathbf{b} = ab \cos \vartheta \tag{A.1}$$

where ϑ is the angle between the two vectors. Following this definition, one has

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{b} \cdot \mathbf{a} \tag{A.2}$$

$$\mathbf{a} \cdot (\mathbf{b} + \mathbf{c}) = \mathbf{a} \cdot \mathbf{b} + \mathbf{a} \cdot \mathbf{c} \tag{A.3}$$

$$\alpha \left(\mathbf{a} \cdot \mathbf{b}\right) = \alpha \mathbf{a} \cdot \mathbf{b} = \mathbf{a} \cdot \alpha \mathbf{b} \tag{A.4}$$

$$\mathbf{a} \cdot \mathbf{a} = a^2 \tag{A.5}$$

$$\mathbf{a} \cdot \dot{\mathbf{a}} = a\dot{a} \tag{A.6}$$

If $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are unit orthogonal vectors, $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}$, one has

$$\mathbf{i} \cdot \mathbf{j} = \mathbf{j} \cdot \mathbf{k} = \mathbf{i} \cdot \mathbf{k} = 0 \tag{A.7}$$

$$\mathbf{a} \cdot \mathbf{b} = (a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}) \cdot (b_1\mathbf{i} + b_2\mathbf{j} + b_3\mathbf{k}) \tag{A.8}$$

Using equations (A.3) and (A.7), equation (A.8) rewrites as

$$\mathbf{a} \cdot \mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3 \tag{A.9}$$

Vector product The vector product of two vectors $\mathbf{a}$ and $\mathbf{b}$, denoted by $\mathbf{a} \times \mathbf{b}$, is a vector $\mathbf{c}$ such that

$$c = ab \sin \vartheta \tag{A.10}$$

where ϑ is the angle between the two vectors and its direction is perpendicular to both $\mathbf{a}$ and $\mathbf{b}$. Following this definition, one has

$$\mathbf{a} \times \mathbf{b} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{pmatrix} \tag{A.11}$$

whete $\mathbf{i}$, $\mathbf{j}$ and $\mathbf{k}$ are unit orthogonal vectors, $\mathbf{a} = a_1\mathbf{i}+a_2\mathbf{j}+a_3\mathbf{k}$ and $\mathbf{b} = b_1\mathbf{i}+b_2\mathbf{j}+b_3\mathbf{k}$. In addition,

$$\mathbf{a} \times \mathbf{a} = \mathbf{0} \tag{A.12}$$

$$\mathbf{a} \times \mathbf{b} = -\left(\mathbf{b} \times \mathbf{a}\right) \tag{A.13}$$

$$\mathbf{a} \times \left(\mathbf{b} + \mathbf{c}\right) = \mathbf{a} \times \mathbf{b} + \mathbf{a} \times \mathbf{c} \tag{A.14}$$

$$\alpha\left(\mathbf{a} \times \mathbf{b}\right) = \alpha\mathbf{a} \times \mathbf{b} = \mathbf{a} \times \alpha\mathbf{b} \tag{A.15}$$

$$\mathbf{a} \times \left(\mathbf{b} \times \mathbf{c}\right) = \left(\mathbf{a} \cdot \mathbf{c}\right)\mathbf{b} - \left(\mathbf{a} \cdot \mathbf{b}\right)\mathbf{c} \tag{A.16}$$

$$\left(\mathbf{a} \times \mathbf{b}\right) \times \mathbf{c} = \left(\mathbf{a} \cdot \mathbf{c}\right)\mathbf{b} - \left(\mathbf{b} \cdot \mathbf{c}\right)\mathbf{a} \tag{A.17}$$

$$\mathbf{a} \cdot \left(\mathbf{b} \times \mathbf{c}\right) = \mathbf{c} \cdot \left(\mathbf{a} \times \mathbf{b}\right) = \mathbf{b} \cdot \left(\mathbf{c} \times \mathbf{a}\right) \tag{A.18}$$

Equation (A.18) is readily proved by expanding equation (A.18) by using equations (A.11) and (A.9), which give

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = \begin{vmatrix} c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix} = \begin{vmatrix} b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \\ a_1 & a_2 & a_3 \end{vmatrix}$$

which is easy to verify.

A.2 Vector notation

This work assumes a bold character denotes a column vector:

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{A.19}$$

A row vector is written as $\mathbf{a}^T$, where the letter T denotes the transpose matrix. This way, the scalar product $\mathbf{a} \cdot \mathbf{b}$ can be written as

$$\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b} = \mathbf{b}^T \mathbf{a} \tag{A.20}$$

The time derivative of a column vector yields a column vector whose components are the time derivatives of each one of the components of the vector:

$$\frac{d\mathbf{a}}{dt} = \begin{pmatrix} {da_1}/{dt} \\ {da_2}/{dt} \\ \vdots \\ {da_n}/{dt} \end{pmatrix} \tag{A.21}$$

The derivative of a scalar quantity with respect to a column vector yields a row vector whose components are the derivatives of the scalar quantity with respect to the components of the vector:

$$\frac{d\varphi}{d\mathbf{a}} = \left(\frac{d\varphi}{da_1}, \frac{d\varphi}{da_2}, \dots, \frac{d\varphi}{da_n} \right)$$

The derivative of a vector $\mathbf{a}$ with n components with respect to another vector $\mathbf{b}$ with m components yields a matrix with n rows and m columns. Each column containes the derivatives of the components of $\mathbf{a}$ with respect to a single component of $\mathbf{b}$, whereas each row containes the derivatives of a single components of $\mathbf{a}$ with respect to the components of $\mathbf{b}$:

$$\left[\frac{d\mathbf{a}}{d\mathbf{b}} \right] = \begin{bmatrix} {da_1}/{db_1} & \cdots & {da_1}/{db_m} \\ \vdots & \ddots & \vdots \\ {da_n}/{db_1} & \cdots & {da_n}/{db_m} \end{bmatrix}$$

Appendix B

Coordinate transformations

A physical vector is a mathematical quantity that has three dimensions and only possesses two properties: magnitude and direction. The concept of physical vector and all of the operations associated with physical vectors are completely indipendent of a coordinate system. However, frames of reference are needed for the formulation of kinematics and dynamics. The same physical vector can be expressed in any suitable coordinate system. In the case of rectangular coordinate systems, any physical vector can be expressed as a linear combination of the three unit vectors that make up the basis of the system. Given a coordinate system, a coordinate transformation only changes the basis of a vector. That mathematical properties that define the vector (magnitude and direction) remain unchanged.

Rotation matrices Suppose we have the coordinates of a vector in a certain coordinate system XYZ

$$\mathbf{a} = a_X \mathbf{X} + a_Y \mathbf{Y} + a_Z \mathbf{Z}$$

and we want to calculate its coordinates in another system UVW

$$\mathbf{a} = a_U \mathbf{U} + a_V \mathbf{V} + a_W \mathbf{W}$$

The coordinates are simply given by

$$a_U = \mathbf{a} \cdot \mathbf{U}$$

$$a_V = \mathbf{a} \cdot \mathbf{V}$$

$$a_W = \mathbf{a} \cdot \mathbf{W}$$

Therefore, by using equations (A.3) and (A.4) we can write

$$\begin{aligned} a_U &= a_X\left(\mathbf{X}\cdot\mathbf{U}\right) + a_Y\left(\mathbf{Y}\cdot\mathbf{U}\right) + a_Z\left(\mathbf{Z}\cdot\mathbf{U}\right) \\ a_V &= a_X\left(\mathbf{X}\cdot\mathbf{V}\right) + a_Y\left(\mathbf{Y}\cdot\mathbf{V}\right) + a_Z\left(\mathbf{Z}\cdot\mathbf{V}\right) \\ a_W &= a_X\left(\mathbf{X}\cdot\mathbf{W}\right) + a_Y\left(\mathbf{Y}\cdot\mathbf{W}\right) + a_Z\left(\mathbf{Z}\cdot\mathbf{W}\right) \end{aligned} \tag{B.1}$$

By using the notation introduced in Appendix A, we can write

$$\mathbf{a}_{UVW} = \begin{pmatrix} a_U \\ a_V \\ a_W \end{pmatrix}$$

$$\mathbf{a}_{XYZ} = \begin{pmatrix} a_X \\ a_Y \\ a_Z \end{pmatrix}$$

where the subscripts identify the coordinate system. By introducing the rotation matrix

$$\mathbf{L} = \begin{bmatrix} \mathbf{X}\cdot\mathbf{U} & \mathbf{Y}\cdot\mathbf{U} & \mathbf{Z}\cdot\mathbf{U} \\ \mathbf{X}\cdot\mathbf{V} & \mathbf{Y}\cdot\mathbf{V} & \mathbf{Z}\cdot\mathbf{V} \\ \mathbf{X}\cdot\mathbf{W} & \mathbf{Y}\cdot\mathbf{W} & \mathbf{Z}\cdot\mathbf{W} \end{bmatrix} \tag{B.2}$$

we can write

$$\mathbf{a}_{UVW} = \mathbf{L}\mathbf{a}_{XYZ} \tag{B.3}$$

which, by applying the rules of matrix multiplication, yields the set of equations (B.1).

The columns of matrix (B.2) are the unit vectos of XYZ expressed in UVW:

$$\mathbf{L} = \begin{bmatrix} \mathbf{X}_{UVW} & \mathbf{Y}_{UVW} & \mathbf{Z}_{UVW} \end{bmatrix}$$

The multiplication between the transpose of $\mathbf{L}$ and itself yields

$$\begin{aligned} \mathbf{L}^T \cdot \mathbf{L} &= \begin{bmatrix} \mathbf{X}^T_{UVW} \\ \mathbf{Y}^T_{UVW} \\ \mathbf{Z}^T_{UVW} \end{bmatrix} \begin{bmatrix} \mathbf{X}_{UVW} & \mathbf{Y}_{UVW} & \mathbf{Z}_{UVW} \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \end{aligned}$$

This shows that $\mathbf{L}^{-1} = \mathbf{L}^T$ and the rotation matrix is an orthonormal matrix. This implies that we can invert the transformation in equation (B.3) and write

$$\mathbf{a}_{XYZ} = \mathbf{L}^T \mathbf{a}_{UVW}$$

Let us now consider three coordinate systems, denoted as 1, 2 and 3. Let $\mathbf{L}_{21}$ be the rotation matrix from coordinate system 1 to 2 and $\mathbf{L}_{32}$ the transformation matrix from 2 to 3. Therefore, we can write

$$\mathbf{a}_3 = \mathbf{L}_{32}\mathbf{a}_2 = \mathbf{L}_{32}\mathbf{L}_{21}\mathbf{a}_1$$

If $\mathbf{L}_{31}$ is the rotation matrix from coordinate system 1 to 3, we have

$$\mathbf{L}_{31} = \mathbf{L}_{32}\mathbf{L}_{21} \tag{B.4}$$

Therefore, successive rotations can be combined to transform the coordinates of a vector.

Principal rotations If two coordinate systems are related by a simple rotation through a positive angle α about an axis, the rotation matrix assumes a simple form. Let us assume coordinate system UVW is obtained by rotating XYZ through an angle α about the Z-axis, such that $\mathbf{Z} \equiv \mathbf{W}$. The rotation matrix from equation (B.2) becomes

$$\mathbf{L} = \begin{bmatrix} \cos\alpha & \sin\alpha & 0 \\ -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

In the same manner, for a rotation about the Y-axis,

$$\mathbf{L} = \begin{bmatrix} \cos\alpha & 0 & -\sin\alpha \\ 0 & 1 & 0 \\ \sin\alpha & 0 & \cos\alpha \end{bmatrix}$$

whereas for a rotation about the X-axis,

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha \\ 0 & -\sin\alpha & \cos\alpha \end{bmatrix}$$

Transformation from the Geocentric-Equatorial to the Perifocal coordinate system The perifocal coordinate system can be obtained from the geocentric-equatorial system through successive rotations. As a matter of fact, by rotating IJK through an angle equalt to the RAAN Ω about K, the transformed x-axis lies along the line of nodes. A further rotation through the inclination angle $\imath$ about the new x-axis, transforms the z-axis in w (the z-axis of the perifocal frame). Finally, a rotation through an angle equal to the argument of perigee ω about w completes the transformation. By applying the rule of successive rotations as described by equation (B.4), we have

$$\mathbf{L}_{\mathrm{pqwIJK}} = \begin{bmatrix} \cos\omega & \sin\omega & 0 \\ -\sin\omega & \cos\omega & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\imath & \sin\imath \\ 0 & -\sin\imath & \cos\imath \end{bmatrix} \begin{bmatrix} \cos\Omega & \sin\Omega & 0 \\ -\sin\Omega & \cos\Omega & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Multiplicating the matrixes yields

$$\mathbf{L}_{\mathrm{pqwIJK}} = \begin{bmatrix} \cos\Omega\cos\omega - \sin\Omega\sin\omega\cos\imath & \sin\Omega\cos\omega + \cos\Omega\sin\omega\cos\imath & \sin\omega\sin\imath \\ -\cos\Omega\sin\omega - \sin\Omega\cos\omega\cos\imath & -\sin\Omega\sin\omega + \cos\Omega\cos\omega\cos\imath & \cos\omega\sin\imath \\ \sin\Omega\sin\imath & -\cos\Omega\sin\imath & \cos\imath \end{bmatrix} \tag{B.5}$$

Therefore, we have

$$\mathbf{a}_{\mathrm{pqw}} = \mathbf{L}_{\mathrm{pqwIJK}}\mathbf{a}_{\mathrm{IJK}} \tag{B.6}$$

and

$$\mathbf{a}_{\mathrm{IJK}} = \mathbf{L}^{T}_{\mathrm{pqwIJK}}\mathbf{a}_{\mathrm{pqw}} \tag{B.7}$$

Transformation from the Perifocal to the RTW coordinate system As described in Section 2.4.1.1, the RTW coordinate system is the rotation of the perifocal coordinate system through an angle equal to the true anomaly ν about $\mathbf{w}$. The rotation matrix is

$$\mathbf{L}_{\mathrm{RTWpqw}} = \begin{bmatrix} \cos\nu & \sin\nu & 0 \\ -\sin\nu & \cos\nu & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{B.8}$$

Appendix C

Minimum-propellant tables

Table C.1: Minimum propellant consumption for transfer with lines of nodes initially moving closer

	$\Delta a_0 = -200\,\mathrm{km}$ $\Delta \imath_0 = 0°$ $\Delta\Omega_0 = 10°$			$\Delta a_0 = +200\,\mathrm{km}$ $\Delta \imath_0 = 0°$ $\Delta\Omega_0 = -10°$		
	No drag sail	$S_{\max} = 4\,\mathrm{m}^2$	$S_{\max} = 400\,\mathrm{m}^2$	No drag sail	$S_{\max} = 4\,\mathrm{m}^2$	$S_{\max} = 400\,\mathrm{m}^2$
t_f[days]	m_p [kg]					
12	0.183	0.134	0.039	0.198	0.184	0.126
12.5	0.167	0.121	0.034	0.183	0.164	0.120
13	0.153	0.109	0.029	0.170	0.150	0.114
13.5	0.141	0.098	0.025	0.158	0.139	0.109
14	0.130	0.089	0.021	0.148	0.130	0.104
14.5	0.121	0.080	0.018	0.138	0.122	0.100
15	0.113	0.071	0.015	0.130	0.115	0.096
15.5	0.105	0.063	0.012	0.123	0.108	0.093
16	0.099	0.055	0.009	0.116	0.102	0.089
16.5	0.092	0.046	0.007	0.110	0.096	0.087
17	0.086	0.036	0.005	0.104	0.091	0.084
Global optimum	0.069	0.009	0.002	0.069	0.069	0.069

Table C.2: Minimum propellant consumption for transfer with lines of nodes initially drifting apart ($\Delta a_0 = +200\,\mathrm{km}$, $\Delta \imath_0 = 0°$ and $\Delta\Omega_0 = 10°$)

	No drag sail	$S_{\max} = 4\,\mathrm{m}^2$	$S_{\max} = 400\,\mathrm{m}^2$
t_f[days]		m_p [kg]	
20	0.269	0.269	0.259
30	0.183	0.183	0.167
40	0.150	0.150	0.114
50	0.132	0.131	0.102
60	0.120	0.120	0.095
70	0.112	0.111	0.090
80	0.106	0.105	0.087
90	0.102	0.101	0.085
100	0.098	0.097	0.083
110	0.095	0.094	0.082
120	0.093	0.091	0.080
130	0.091	0.089	0.079
140	0.089	0.087	0.079
150	0.088	0.085	0.078
160	0.087	0.084	0.077
170	0.085	0.082	0.077
180	0.084	0.081	0.076
190	0.083	0.080	0.076
200	0.083	0.078	0.075
210	0.082	0.077	0.075
220	0.081	0.076	0.075
230	0.081	0.076	0.074
240	0.080	0.075	0.074
250	0.080	0.075	0.074
260	0.079	0.074	0.074
270	0.079	0.074	0.073
280	0.078	0.074	0.073
290	0.078	0.074	0.073
300	0.078	0.073	0.073
310	0.077	0.073	0.073
320	0.077	0.073	0.072
330	0.077	0.073	0.072
340	0.076	0.072	0.072
350	0.076	0.072	0.072

	No drag sail	$S_{\max} = 4\,\mathrm{m}^2$	$S_{\max} = 400\,\mathrm{m}^2$
t_f[days]		m_p [kg]	
360	0.076	0.072	0.072
370	0.076	0.072	0.072
380	0.075	0.072	0.072
390	0.075	0.072	0.072
400	0.075	0.072	0.072
410	0.075	0.072	0.071
420	0.075	0.072	0.071
430	0.075	0.071	0.071
440	0.074	0.071	0.071
450	0.074	0.071	0.071
460	0.074	0.071	0.071
470	0.074	0.071	0.071
480	0.074	0.071	0.071
490	0.074	0.071	0.071
500	0.074	0.071	0.071
510	0.073	0.071	0.071
520	0.073	0.071	0.071
530	0.073	0.071	0.071
540	0.073	0.071	0.071
550	0.073	0.071	0.071
560	0.073	0.071	0.071
570	0.073	0.071	0.070
580	0.073	0.070	0.070
590	0.073	0.070	0.070
600	0.073	0.070	0.070
610	0.072	0.070	0.070
620	0.072	0.070	0.070
630	0.072	0.070	0.070
640	0.072	0.070	0.070
650	0.072	0.070	0.070
660	0.072	0.070	0.070
670	0.072	0.070	0.070
680	0.072	0.070	0.070
690	0.072	0.070	0.070
700	0.072	0.070	0.070
710	0.072	0.070	0.070
Global optimum	0.069	0.069	0.069