Politecnico di Torino

Master's Degree Course in Aerospace Engineering
A.Y. 2022/2023

Downrange Analysis for Optimal Lunar Soft Precision Landing

Supervisor
Prof. Lorenzo Casalino

Candidate
Daniele Amore

Abstract

From the moment Galileo first laid eyes on the Moon's craters to the historic Apollo mission that saw the first footsteps on its surface, lunar exploration has always been a fascinating and ambitious endeavor for humankind. In the near future, the world is poised to return to the Moon with NASA's Artemis program, which aims to establish sustainable human presence on another celestial body. Due to its proximity, the Moon is the ideal place to test and develop the technologies and systems required for deep space exploration, as well as offering a unique environment to provide relevant contributions in different fields of scientific research. The study of lunar geology could provide valuable insights into astrophysics, expanding our point of view on the formation and evolution of the solar system. Additionally, examining the effects of microgravity and radiation on living organisms could lead to significant advancements in space medicine, with resulting benefits for life on Earth. Finally, a return to the Moon could have a significant social and cultural impact by inspiring the next generation of scientists, engineers, and explorers. However, lunar exploration poses technical, environmental, and logistical challenges that demand innovative solutions. Among these is the need to develop a precise and robust Guidance, Navigation, and Control (GNC) system to achieve a successful landing on the surface. This system must be capable of guiding a spacecraft through the harsh lunar environment, providing accuracy in navigation, and ensuring accurate autonomous control of the vehicle during descent and landing. A reliable GNC system can significantly reduce the risk of mission failures and increase the safety of the astronauts and payloads involved. This thesis work investigates a guidance algorithm for lunar powered descent and soft-landing trajectories. Taking inspiration from the concept of operations of the Apollo Lunar Module, it focuses on a 3-degree-of-freedom lander, considering only the pitch dynamics and neglecting the cross-range motion. The model also involves optimal control theory to maximize the final mass of the vehicle. This approach can considerably reduce the necessary propellant mass for the mission, allowing for the opportunity to carry more payload onboard and reduce mission-related costs. Furthermore, to account for the possibility of unknown hazards that could become clearer as the lander approaches the lunar surface, a downrange analysis was performed. This allows for real-time adjustments of the landing site to ensure safe and efficient touchdown, also evaluating the potential for propulsion system shutdown. The optimal guidance laws are found using an indirect optimization method based on Pontryagin's Maximum Principle, which achieves high numerical precision and reduces computation time. This mathematical theorem states that the optimal control strategy is the one that maximizes a certain cost functional, that, in this case, is defined in terms of propellant consumption. The resulting equations lead to a Boundary Value Problem (BVP), where constraints on the initial, final and intermediate state, and bounded control variables must be considered and rigorously fulfilled.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

This thesis proposes a guidance algorithm for optimizing the trajectory of lunar soft-landing. Precise and systematic definition of spacecraft maneuvers is a critical step in the early stages of space mission design. In this regard, the primary objective of the work is to identify optimal landing trajectories that allow for maximum final lander mass, i.e., that minimize propellant consumption, while ensuring high precision and safety during landing. To achieve this goal, a 3-degree-of-freedom lander model was implemented, drawing inspiration from the characteristics and mission architecture of the Lunar Module used by NASA in the Apollo missions. For this reason, the first chapter will provide a detailed introduction to the Apollo program, trying to give an idea of the historical and political context that led to the first human landing on the Moon. Furthermore, the analysis of past missions can provide valuable insights to improve our capabilities for future lunar exploration. The chapter will continue with a review of the primary goals and interests of returning to the Moon, with a focus on current and future Artemis missions, and providing a detailed overview of NASA's approach to Moon-to-Mars exploration.

Subsequently, in Chapter 3 the main characteristics of a guidance, navigation and control (GN&C) system will be presented. To ensure a soft and precise landing on the lunar surface requires a highly efficient and robust GN&C system. It must be capable of detecting and avoiding hazards, via the Hazard Detection and Avoidance system, which should reliably find an alternate landing site and command a re-targeting maneuver if the default landing site it is not safe. In addition, a vision-based navigation system is also required to achieve high-precision position and orientation control. It must consider motion and attitude constraints, as well as time and altitude margin, to allow the sensors to see the target area. To introduce these concepts, in the first paragraph of the chapter, the configuration of the Lunar Module's GN&C and a brief description of the guidance law adopted for its descent and ascent phases will be presented. The fundamental aspects of nowdays GN&C systems for landing missions on celestial bodies will then be detailed.

At this point, in order to find a trajectory that minimizes fuel consumption, it is necessary to introduce the fundamental concepts of optimal control theory. This theory provides a mathematical framework for determining the control inputs required to optimize the performance of a system, subject to constraints and uncertainties. Chapter 4, starting with the formulation of the optimal control problem, will initially present an overview of the numerical methods used to search for optimal trajectories. Among all these categories, the indirect approach is widely used in space engineering due to its high level of precision

and accuracy. For this reason, the solution algorithm shall employ an integrator based on Adams method. Before presenting the code adopted for the lunar landing trajectory, the variational approach to the optimal control problem will be introduced. In this way it will be possible to derive the necessary conditions for optimality, that will then be utilized to determine the optimal control law. Additionally, we shall derive Pontryagin's Maximum Principle (PMP) as a generalization of the fundamental theorem of the calculus of variations. Finally, adopting this theorem a Boundary Value Problem (BVP) will be obtained, and the shooting method to solve it will be presented.

Once the mathematical formulation of the problem has been given, it is necessary to define the model examined in this thesis work. Specifically, Chapter 5 delineates the simplifications assumed for the system's dynamics. Starting with the presentation of mission scenario and showing the main phases that the lander will have to follow, then a brief description of the Moon will be provided. Following, the reference frames adopted to derive the motion equations will be explained, also providing an outline of the boundary conditions for the landing maneuver. Additionally, the main propulsion system's for translational dynamics and the reaction control system for attitude control will be specified in detail, highlighting how the vehicle's orientation influences the path to follow and explaining their limitations.

After, the foundation for understanding the adopted model is laid, it will be possible to address trajectory optimization for powered descent and soft-landing. The choice of a fuel-minimizing strategy is a fundamental aspect of the lunar landing mission design, as it has important implications for the vehicle's mass budget. Therefore, Chapter 6 presents the formulation of the Boundary Value Problem that allows us to maximize the mass of the lander at the end of the trajectory, obtained by applying the Pontryagin's Maximum Principle to the optimal control problem. Initially, the primer vector introduced by D. F. Lawden, which indicates the optimal thrust direction, will be defined. Then, we will continue with the presentation of the Bang-Bang control, adopeted for both motion and attitude control systems, and the constraints imposed on the maneuver to ensure its efficiency and safety. To this end, the subdivision into sub-intervals introduced in the trajectory will be explained in detail. Finally, the algorithm implemented to solve this optimization problem will be discussed.

Ultimately, in Chapter 7 a downrange analysis will be conducted, in order to evaluate the response of the lander to the variation of the thrust and the position of the landing site. Starting from a reference case, the primary objective was to look for better solutions in terms of propellant saving. To this end, the effect of increasing the aviable thrust was evaluated in the first part. Additionally, it will be shown that further improvements can be obtained with an initial coasting arc turning off the propulsion system. On the other hand, in order to ensure a good safety margin for the mission, a real-time change of the landing site was evaluated, lengthening or shortening the downrange. Therefore, the second part will present the results for this scenario, where a trajectory correction is necessary.

Chapter 2

Lunar Mission

For centuries, the idea of reaching the Moon seemed impossible, but it became a reality four decades ago. During the historic Apollo 11 mission, Neil Armstrong became the first human to step on the lunar surface. His words

"It is one small step for man, one giant leap for mankind"

inspired all mankind on July 20, 1969. Several events led to this success. Therefore, in the first part of this chapter, previous lunar missions will be analyzed in chronological order, trying to give the reader an idea of the historical and political background that led humans to the Moon for the first time. Specifically, it will begin with a brief description of the soviet Luna program. It began in 1959 with the launch of the Luna 1 probe, the first human project to leave Earth and approach the Moon. This mission was followed by many others, some of which failed and led to the program's definitive interruption in 1976. A few years after the start of the Luna missions, the United States launched the Surveyor program. Between 1966 and 1968, seven probes were sent to the lunar surface with the main objective of testing landing techniques and collecting scientific information about the Moon, in preparation for future Apollo missions. Finally, the Apollo program will be presented in detail, focusing on the characteristics of the Lunar Module and introducing part of the mission scenario from which this thesis work takes inspiration.

After the Apollo 17 mission, interest in the Moon gradually declined. However, scientific research has been carried out over the years by orbiters and robotic missions, in order to deepen the study of the Moon. We also started living in space aboard the International Space Station (ISS). Today, scientific and technological progress since then, the knowledge and experience gained from space missions, and the innate desire of humans to explore new horizons will bring humans back to tread on the lunar surface. Thus, the chapter will continue with a look at the future objectives and main interests of lunar exploration, focusing mainly on current and upcoming Artemis missions and presenting in detail NASA's approach to Moon. Artemis program is designed to advance humanity's exploration and prepare for the next significant leap, which is the exploration of Mars. Sending human explorers to the Moon, which is 250,000 miles away, requires a bold vision, effective program management, funding for modern systems development and mission operations, as well as the support and collaboration with commercial and international partners worldwide.

2.1 Historical Background

When designing a new mission, it is essential to analyze previous missions to gain insights and develop improved techniques and technologies. Reviewing past mission outcomes and challenges can enhance our technological knowledge about landing on the Moon and improve our understanding of its physical characteristics. Lunar missions from the past, including Luna, Surveyor, and Apollo programs, have been instrumental in gathering data on the Moon, which can now be used to model its surface. In particular, a detailed gravity model and surface map are critical for designing a successful descent mission. Therefore, by examining past missions that have landed on the Moon, we can advance our capabilities for future lunar exploration. It is important to note that this paragraph will only focus on lander missions and not orbiters.

2.1.1 Luna Program

The soviet space program Luna, initiated in the 1950s, was the first program to send a probe to the Moon. Historically, it was influenced by the success of the first satellite launched into orbit, Sputnik, in 1957, which sparked a space race between the United States and the Soviet Union fueled by the political rivalry of the Cold War. Consequently, the Soviet Union was interested in demonstrating its technological and scientific superiority over the United States to the international community, and the Luna program was an important part of this political strategy. The program mainly consisted of robotic missions aimed at studying the lunar surface, its composition, and radiation. This was not only done for scientific purposes but also as a preparation for future missions, including manned missions to the Moon. Additionally, the program was also seen as a way to explore and develop the lunar space resources such as helium-3 and other substances useful for nuclear energy production, which could have led to significant economic benefits.

The opening mission was Luna 1, launched on January 2, 1959, and was the first probe to reach the Moon, although it was not able to enter into lunar orbit. Luna 1 passed about 5,995 kilometers from the lunar surface, transmitting scientific data about its structure and gravity. Although not a completely successful mission, it was still a great success for the Soviet Union, demonstrating their ability to send a spacecraft towards the Moon and obtain important scientific information about our natural satellite. Shortly after, in 1959, Luna 2 became the first spacecraft to touch the Moon's surface. It crashed into the Mare Imbrium, an immense lava plain situated in the Imbrium Basin that is one of the larger craters in the Solar System. Only a few weeks later, a third mission, Luna 3, was the first to enter into lunar orbit and capture the primary, even though blurry, images of the more remote side of the Moon, which made it possible to observe that the rough highland terrain is strongly different from the smoother basins on the side facing the Earth. Thus, the program saw several firsts, with fifteen successful missions. The greatest one was certainly the Luna 9 mission, which completed the first soft landing on the Moon in 1966. To give an impression of the spacecraft mass, Luna 9 had a launch dry mass of 1850 kg and a landing mass of 99 kg, whereas Luna 24 was heavier with values of 4800 and 1900 kg, respectively. Luna 9 began its trajectory in a low orbit around the Earth. Then, it started with a direct landing trajectory to the Moon. Luna 16, 17, 20, 21, and 24 used different trajectories. These missions also started their journey in a low orbit around the

Earth. However, from this orbit, they went into a translunar trajectory followed by an orbit around the Moon. From here, they started with the descent to the surface of the Moon. Only Luna 16, 20, and 24 returned to Earth after completing their mission on the Moon, carrying samples of lunar soil. The capsules returned to Earth with a direct trajectory and made a ballistic entry into the Earth's atmosphere. With parachutes, they landed on Soviet ground. However, the program also experienced numerous failures, including several failed launches and missions that never managed to land on the lunar surface due to technical issues. Nonetheless, the program was considered a success as it led to important scientific and technical developments.

2.1.2 Surveyor Program

Always in the context of the space race, in 1961, US President John F. Kennedy announced the goal of landing a man on the Moon by the end of the decade. The Surveyor program was the first step towards achieving this objective. Its primary goal was to develop enough technical knowledge to support the Apollo program, specifically by testing the technology for soft-landing and midcourse corrections. Additionally, the program was an important milestone in the exploration of the lunar surface, as it aimed to select suitable landing sites for the Apollo program and evaluate their suitability for human landing. To achieve this, each of the seven probes sent to the Moon was equipped with cameras to map the topography of the surface, spectrometry tools and magnets to determine the soil composition, as well as sensors to measure temperature and atmospheric pressure. Moreover, some of the spacecrafts were equipped with robotic arms to collect lunar soil samples and test its mechanics. In fact, before the Luna 9 mission and the Surveyor project, the depth of the lunar surface was unknown.

The Surveyor 1 mission was launched on May 30th, 1966. Despite landing on the Moon about four months after the Russian Luna 9, it was a significant achievement for the Americans. Surveyor 1 utilized a direct transfer between Earth and the Moon, and Vernier engines were used to slow down the spacecraft during descent and achive a soft landing. Afterwards, the program continued with the sending of new probes until 1968. Overall, five of the spacecraft were successful. Surveyor 2 crashed at high velocity after a failed correction, and Surveyor 4 lost contact, possibly exploding, 2.5 minutes before its scheduled touch-down. All the other missions were successful and are still on the Moon as none of these missions involved a return to Earth. Parts of Surveyor 3 were later brought back to Earth by the crew of Apollo 12, which landed near it in 1969. To give an idea, the launch mass of the spacecrafts was about 1000 kg, while the landing mass was around 300 Kg. Surveyor missions began their journeys in an Earth parking orbit and landed on the Moon via lunar orbit. The Surveyor missions collected information about landing sites and determined whether they were safe enough for Apollo. Four of the Surveyor spacecraft landed in mare regions near the equator, which were considered relatively safe. Surveyor 7 was the only exception, with its landing site chosen on a scientific basis, on the outer rim of the Tycho crater, as enough information had been gathered for the Apollo program.

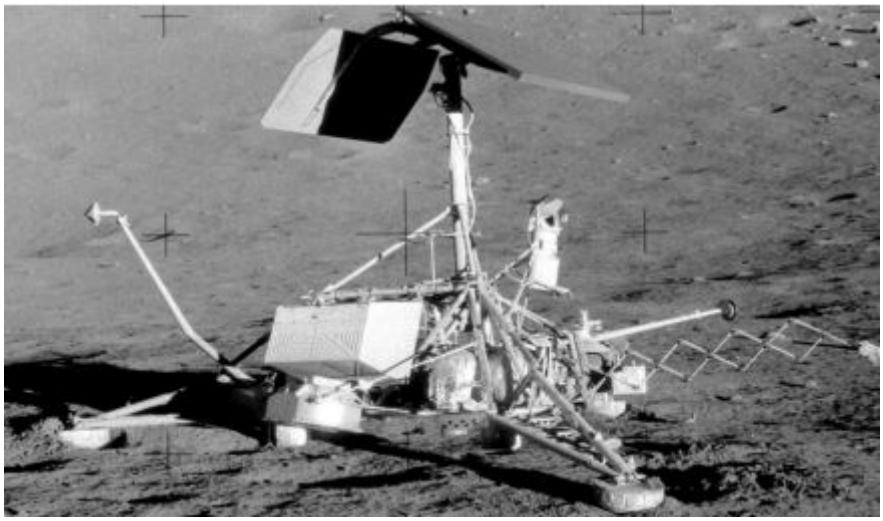


Figure 2.1.1: Surveyor 3 photographed by the Apollo 12 crew

Total Mass [Kg]	**16,471**
Stages	**2**
Descent Stage Mass [Kg]	**11,666**
Ascent Stage Mass [Kg]	**4,805**
Pressurized Volume [m^3]	**16,471**
Landing Site Capability	**16,471**
Descent Stage Engine	$T = 44.1kN$, $I_{sp} = 311s$
Ascent Stage Engine	$T = 15.6kN$, $I_{sp} = 311s$
Reaction Control System	$T = 411N$, $I_{sp} = 311s$

Table 2.1.1: Apollo's Lunar Module specifications

2.1.3 Apollo Porgram

Following the success of the Surveyor missions, the United States was ready to conduct crewed missions to the Moon. The Apollo Program proved to be a remarkable achievement with Neil Armstrong becoming the first human to set foot on the lunar surface during the Apollo 11 mission on July 20, 1969, followed by five more successful landings. The only unsuccessful mission was Apollo 13, which failed to reach its destination due to an oxygen tank explosion.

The spacecraft used in the Apollo program shared several common features. They were all comprised of two major components: the Command Service Module (CSM) and the Lunar Module (LM). The CSM had two sections, the Command Module (CM) which served as a living quarters and control center for the astronauts, and the module that was used for re-entry at the end of the mission. The second part was the Service Module (SM), which provided the primary propulsion and maneuverability for the spacecraft. The LM was made up of a descent and ascent stage. Specifications about LM and its main and Reaction Control System thrusters are shown in Table 2.1.1. Additionally, the Apollo 15, 16 and 17 missions included a rover with a mass of 210 kg, which allowed astronauts to travel further away from the landing site and collect more samples. During

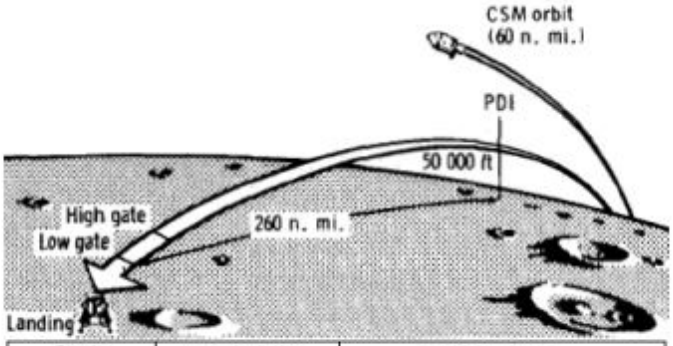
Phase	**Mission Time** [hh:mm:ss]	**Altitude** [Km]
Lunar orbit insertion	75:49:50.37	160.6
Lunar orbit circularization	080:11:36.75	114.5
CSM/LM undocked	100:12:00	116.5
CSM/LM separation	100:39:52.90	116.1
LM descent orbit insertion	101:36:14.00	104.5
LM powered descent initiation	102:33:05.01	11.9

Table 2.1.2: Apollo 11 Descent Phases

the Apollo Program, three distinct mission profiles were evaluated: direct ascent, Earth Orbit Rendezvous (EOR), and Lunar Orbit Rendezvous (LOR). NASA ultimately selected the LOR approach for the Apollo missions. A Saturn V rocket was employed to launch both the Command Service Module (CSM) and the Lunar Module (LM) into Earth orbit. The third stage engine of the rocket was then utilized to place the spacecraft on a trajectory toward the Moon. Once in lunar orbit, the LM separated from the CSM, with one astronaut remaining in the CSM in orbit while the other two astronauts descended to the Moon's surface in the LM. The Descent Module (DM) of the LM was left behind on the Moon as a launch platform for the ascent stage, while the ascent stage functioned as a habitation environment for the astronauts and transported them back to the CSM after their lunar exploration. Prior to re-entry into Earth's atmosphere, the Service Module (SM) was jettisoned from the CM, and the astronauts returned to Earth in the CM, splashing down in the ocean.

Following the de-coupling of the LM from the CSM, a Descent Orbit Insertion (DOI) maneuver was performed to bring the LM to a lower altitude and start the powered descent. This maneuver consisted of three operational phases, each serving different requirements. These phases were the braking phase, which reduced the orbital velocity and maximized the efficient use of propellant; the approach phase, where the pilot used visual monitoring to approach the landing site; and the landing phase, during which the pilot took control from the automatic system for the final touchdown. The powered descent phase began with the ignition of the Descent Propulsion System (DPS) at an altitude of 15.2 km, approximately 480 km from the landing site. Performing the descent at higher altitudes would be inefficient, while at lower ones would exceed safety margins in case of failed powered descent initiation or terrain clearance. The powered descent duration was approximately 12 minutes, during which the throttle setting was at maximum for most of the braking phase. In the last 2 minutes of the braking phase, the DPS was throttled for guidance control of thrust and trajectory dispersions. The approach phase commenced at an altitude of 2.1 km and a ground range of 8.3 km from the landing site. During this phase, the crew monitored the external environment and could still alter the target landing site. The vertical landing phase began at an altitude of approximately 50 m above the landing site, with an altitude rate of 0.9 m/s. The powered descent terminated at touchdown with a vertical velocity less than 1 m/s. Notably, the LM was designed to land on the Moon's equator and could not land on the lunar South Pole due to the Free Return option, that ensures the ability to abort the mission and return to Earth safely. Despite this limitation, a soft landing was achieved, and the crew looked outside the LM during the

landing to detect potential hazards, such as craters and boulders. More information on the main phases of the Apollo 11 mission can be found in Table 2.1.2. Project Apollo, ended



Phase	Initial event	Design criteria
Braking	PDI	Minimize propellant usage
Approach	High gate	Crew visibility
Landing	Low gate	Manual control

Figure 2.1.2: Powered Descent trajectory followed by Lunar Module

in 1972, culminating in a total of nine missions. The most iconic of these missions was Apollo 11, during which humans first set foot on another world. On July 20, 1969, astronaut Neil Armstrong became the first person to step onto the moon's surface. Armstrong and fellow astronaut Buzz Aldrin spent 21 hours and 36 minutes on the Moon before rendezvousing with Michael Collins in the command module and returning to Earth.

Subsequent Apollo missions set new landmarks in space travel and lunar exploration. Just four months after Apollo 11, Apollo 12 touched down on the moon and executed an even more precise landing. Apollo 13 narrowly avoided catastrophe when on-board oxygen tanks exploded in April 1970, forcing the crew to abort their mission and miss the third planned moon landing, but fortunately all three astronauts survived. During the next mission, Apollo 14 in January 1971, commander Alan Shepard set a record for the farthest distance traveled on the moon surface during two planned walks, covering a total distance of about 3 kilometers. Shepard even demonstrated the light gravity of the moon by throwing and soaring two golf balls for miles with the help of a makeshift 6-iron club. Apollo 15, launched in July 1971, was the first of the last three missions designed for extended stays on the Moon. During three days on the surface, the crew collected several kilograms of lunar samples and drove more than 27 kilometers in the first piloted lunar buggy. Apollo 16 and 17 in 1972 were the last manned missions to the Moon. Following the incredible achievements of the 1960s and 1970s, major space organizations shifted their focus elsewhere for many years. To date, only 12 humans (all Americans and all men) have set foot on the Moon.

2.2 Current and Future Perspective

Four decades after the last human landing on the Moon, all major nations on the planet are planning to return to the Moon, indicating a significant step forward in human space exploration with the establishment of a permanent lunar human occupation. The renewed

Figure 2.1.3: John W. Young in the "Grand Prix" run during the first Apollo 16 spacewalk

interest in exploring the Moon will reignite the dream of space exploration, facilitating human flights to Mars and beyond. Lunar exploration is not only about conquering an unspoiled world but also has benefits in several areas. Firstly, it benefits our scientific knowledge. Sample collection during previous lunar explorations provided vital data about the geology and formation of Earth's natural satellite. Future lunar explorations will offer the opportunity to gain further insight into the Earth's history and the evolution of the Solar System by studying the lunar surface and the numerous impacts that have characterized its history. In addition, lunar exploration would allow testing of new technologies, both for improving spacecraft technical capabilities in interplanetary environments and for soft-landing technologies, as well as in the field of space medicine to preserve the astronauts' lives. From this point of view, lunar missions are also essential in planning and developing future more challenging missions, such as manned missions to the Martian surface. Furthermore, our next lunar landing will pave the way for a new and sustainable lunar economy, where international partners and companies will benefit from and build on what we learn. Also the possibility of utilizing lunar resources such as helium-3, titanium, and iron to sustain human life and extract energy could be a significant advantage. Moreover, colonizing another world would mean safeguarding the species and life. In this regard, a team of researchers from the University of Florida demonstrated for the first time in 2022 that plants can successfully sprout and grow in lunar soil. The experiment was conducted by cultivating Arabidopsis seeds on samples of soil brought to Earth by Apollo missions. Ultimately, returning to the Moon could have the same social impact as it had in the 1970s, inspiring a new generation of astronauts, engineers, and scientists. Mentioning American astronaut and lunar geologist during the Apollo 17 mission, Harrison Schmitt

"The Moon is a world that represents our ambition, our creativity and our ability to reach new horizons."

For the reasons stated above, next generation of astronauts will step foot where no human has ever been before: the Moon's South Pole. As previously mentioned, for safety reasons, human lunar missions have only landed in areas around the equator. However, for future missions, the interest of various international agencies has focused on the lunar

South Pole, which is of significant scientific interest. In particular, the region of greatest interest is the South Pole-Aitken basin, a large impact crater on the far side of the Moon. The basin's bottom is considerably darker than the surrounding highlands and contains more iron and titanium. Moreover, the ejecta and rim of the basin contain material from the lower crust and possibly the upper mantle. Investigation of the upper mantle is particularly important because it is compositionally different from the lunar crust and could provide essential information about the origin of the Moon. Futhermore, some areas at the lunar South Pole are highly illuminated, providingpermanent communication with Earth. However, the primary reason for exploring the lunar South Pole, instead of other areas, is the presence of water. Water is crucial for establishing a permanent Moon base because it supports life and can be used for chemical processing and as a basis for chemical rocket propulsion. Although earlier missions, such as Clementine and Lunar Prospector, suggested the presence of water on the lunar poles, they could not prove it. It was only recently that NASA's Lunar Crater Observation and Sensing Satellite (LCROSS) mission successfully detected water in the permanently shadowed region of the crater Cabeus near the lunar South Pole on October 9, 2009. For this purpose, several nations have sent various missions in recent years to explore this area.

United States leads in space exploration and now with the Artemis program they are ready to return to the Moon. However, as more countries and companies take aim at the Moon, while maintaining leadership in exploration, America will build a global alliance and explore deep space for the benefit of all. The Moon exploration plan has a dual objective: the first is to achieve the initial human landing by 2025 with acceptable technical risks, while the second is to work towards sustainable lunar exploration in the mid- to late 2020s, preparing for a historic first manned mission to Mars. Artemis was the twin sister of Apollo and goddess of the Moon in Greek mythology. Now, she embodies the path to the Moon, as NASA aims to land the first woman and the first person of color on the Moon using innovative technologies to explore more of the lunar surface than ever before. The ultimate objective is to establish the first long-term presence on the Moon. To accomplish this, several years in orbit and on the surface are required to build operational confidence and support life away from Earth. NASA's deep space transportation system forms the foundation for human return to the Moon. The first component is the Orion spacecraft, which is powered by a service module provided by the European Space Agency (ESA). It has been specifically designed for deep space human operations for up to four crew members. The fully assembled Orion crew module was successfully tested during the Artemis I mission, including the deployment of 11 parachutes at almost five miles in altitude during the capsule's Earth re-entry. In addition, in 2019 NASA conducted a successful test known as Ascent Abort-2, which tested the Orion launch abort system that sits atop Orion at launch and during ascent. This system pulls Orion and its crew away from the rocket to land safely in the Atlantic Ocean in case of an emergency during launch. The other important component is the Space Launch System (SLS), a heavy-lift rocket designed to launch Orion and send it on missions to the Moon and, in the future, to Mars. SLS engines consist of four RS-25 liquid rocket engines, two solid rocket boosters, the massive core stage, and the interim cryogenic propulsion stage that provides Orion's final push toward the Moon. To carry out launch operations, the Exploration Ground System (EGS), including a modernized spaceport, was established. The EGS team responsible for launch operations has performed training simulations inside the firing room to certify the team is

ready for launch and can work through any type of issue in real time. Finally, in the first missions of the program a Human Landing System (HLS) will be necessary to transfer astronauts from a lunar orbit to the surface of the Moon and back.

On later Artemis missions, crew will arrive at the Gateway aboard Orion, which will serve as a staging point for human and robotic lunar missions, supporting longer expeditions on the Moon, and potentially multiple trips to the surface during a single Artemis mission. Additionally, the $Gateway\ to-surface$ operational system is also analogous to how a human Mars mission may work, thus it is crucial to gain operational confidence in this system at the Moon before the first human missions to Mars. The Gateway is designed to operate autonomously and with internationally agreed-upon interoperability standards. It will provide a unique platform to conduct science investigations in deep space and outside the protection of the Earth's Van Allen radiation belts. Once in lunar orbit, the Gateway will enter a period of scientific operations. The international science community has identified heliophysics, radiation, and space weather as high-priority investigations to fly on the Gateway. To accomplish long-term operations, the Gateway will have two modules, the Power and Propulsion Element (PPE) and the Habitation and Logistics Outpost (HALO). The PPE and HALO will be integrated on the ground and launched together on a single rocket in 2024. The PPE's solar electric propulsion system is three times more powerful than current systems, and provides Gateway with electrical power, control, thrust, and communication capabilities. It also provides accommodations for science and technology demonstration payloads. HALO, on the other hand, will be the initial crew cabin for astronauts visiting the Gateway. Its primary purpose is to provide basic life support needs for the visiting astronauts after they arrive in the Orion and prepare for their trip to the lunar surface. HALO will provide command, control, and data handling capabilities, energy storage and power distribution, thermal control, communications and tracking capabilities, as well as environmental control and life support systems to augment the Orion spacecraft and support crew members. It also will have several docking ports for visiting vehicles and future modules, as well as space for science and stowage. The first two payloads to be flown on the Gateway are a radiation instrument package, European Radiation Sensors Array (ERSA), provided by ESA, and a space weather instrument suite, Heliophysics Environmental and Radiation Measurement Experiment Suite(HERMES), provided by NASA. Ultimately, at the lunar South Pole, NASA and its partners will develop an Artemis Base Camp to support longer expeditions on the lunar surface.

2.2.1 Artemis Program

It has been nearly half a century since astronauts last set foot on the lunar surface during the Apollo program, and since then, robotic exploration of deep space has undergone decades of technological advancements and scientific discoveries. Over the past two decades, humans have been living and working continuously aboard the International Space Station, located 250 miles above Earth, in preparation for future missions to explore deeper into the Solar System. Recently, the United States has ushered in a new era of human space exploration with the issuance of *Space Policy Directive-1* by former President Donald Trump on December 11, 2017. An excerpt from the directive reads:

"Lead an innovative and sustainable program of exploration with com-

Figure 2.2.4: Artist's image of the PPE and the HALO in lunar orbit

> *mercial and international parterners to enable human expansion across the Solar System and to bring back to Earth new knowledge and opportunities. Beginning with missions beyond low Earth orbit, the united States will lead the return of humans to the Moon for long-term exploration and utilization, followed by human missions to Mars and other destinations."*

In response to the President's call, NASA was ready to return humans to the Moon with architectural and hardware solutions to leverage the core deep space transportation systems, including the Space Launch System (SLS) rocket, the Orion spacecraft, and the supporting Exploration Ground Systems (EGS). Moreover, through partnerships with U.S. industry, NASA was already developing deep space habitation capabilities and investing in lunar lander technologies. Also upgrades to spacesuits have been made. They have resulted in a lunar surface suit that provides improved safety features, custom fitting, simplified maintenance, and better communications, which enables more frequent and longer spacewalks. To achieve its goals, NASA will collaborate with commercial and international partners to build a sustainable human presence on the Moon, with the ultimate goal of sending astronauts to Mars. This requires the development of new technologies, such as in-situ resource utilization, and the establishment of a lunar economy that will enable NASA and its partners to explore and conduct scientific research more efficiently and effectively. To this end, in 2018 NASA created the Commercial Lunar Payload Services (CLPS) initiative, which encouraged the U.S. commercial space industry to develop new lander technologies to transport NASA and commercial payloads to the Moon's surface. Currently, the American space agency solicits bids from these companies as needed and awards surface task orders for lunar surface deliveries. Multiple deliveries have already been assigned to flights, with the first of two sent in 2022. At least two science instruments and technology experiments are planned to be sent to the Moon's surface per year on CLPS flights. Among these, the most important is certainly Volatiles Investigating Polar Exploration Rover (VIPER) designed by Astrobotic of Pittsburgh. VIPER is planned for 2023 and it is a robotic mission being developed to explore the South Pole of the Moon and search for water and other resources. The rover will travel across the lunar surface, using a suite of instruments to detect and map the location and distribution of water ice and other volatile compounds. Water ice is a valuable resource for future human explo-

ration as it can be used to create rocket fuel, drinking water, and breathable air. VIPER's data will also inform future missions to the Moon. The rover is equipped with a drill that can penetrate up to one meter deep into the lunar regolith to collect samples for analysis. It will explore multiple locations within a 100-day mission, collecting data on the geology and composition of the lunar surface, as well as measuring temperature, lighting, and other environmental factors.

As the first payloads are delivered to the lunar surface aboard CLPS provider landers, the first true step for human exploration began on November 16, 2022, with the launch of **Artemis I**, formerly $Exploration\ Mission-1$ ($EM-1$). The SLS rocket launched an uncrewed Orion into Earth orbit, placing it on a path toward a lunar distant retrograde orbit, where it traveled a total of about 280,000 miles from Earth before returning home. This crucial flight test demonstrated the performance of the SLS rocket on its maiden flight and gathered engineering data throughout before Orion returned on a high speed Earth reentry at Mach 32. The high-speed lunar velocity reentry was the top mission priority and a necessary test of the heat shield's performance as it enters Earth's atmosphere, heating to nearly 5,000 degrees Fahrenheit—about half as hot as the surface of the Sun. For this uncrewed configuration, engineering equipment flew in place of astronaut-essential elements. Instead of the cockpit displays and controls and life support systems that will fly on the first crewed flight, this first flight carried the data-gathering tools needed to validate performance and compare predictive models with actual flight data. This mission also deployed 13 CubeSats to conduct new scientific investigations and new technology demonstrations that will improve our knowledge of the deep space environment, while engaging a broader set of universities, international partners, and private companies in lunar exploration than ever before on a single mission. Over the course of around 25 days, the Orion spacecraft traveled more than 1.4 million miles, surpassing Apollo 13's record for distance traveled from Earth in a spacecraft designed for humans. Finally, the mission was successfully completed with the return to Earth on December 11, 2022, splashing down in the Pacific Ocean. Once recovered, retrieval and post-flight engineering assessment could proceed. With confidence based on the Artemis I mission and extensive prior flight and ground testing, NASA has scheduled **Artemis II** for November 2024, which will mark the first crewed flight of SLS and Orion, sending four astronauts to the lunar environment for the first time in over 50 years. The objective is to conduct in-space testing of every possible hardware, software, and operational system required for Artemis III. The crew will board Orion atop the SLS for an approximate 10-day mission. After launch, Orion will first make two orbits around Earth before committing to the trip to the Moon. The spacecraft will reach an initial elliptical insertion orbit, with a period of around 90 minutes. The perigee will be adjusted via the rocket's first firing of the interim cryogenic propulsion stage (ICPS). After the first orbit, the ICPS will provide the thrust to raise Orion into a high-Earth orbit (HEO) where it will fly in an ellipse for approximately 42 hours. After reaching HEO, Orion will separate from the ICPS. The expended stage will have one final use before it is disposed through Earth's atmosphere. The crew will use it as a target for a proximity operations demonstration. During these activities, the astronauts will pilot Orion's flight path and orientation in manual mode. The crew will use onboard cameras and the spacecraft's windows to line up with the ICPS as they approach and back away from the stage to assess Orion's handling characteristics. This demonstration will provide performance data and operational experience that cannot be

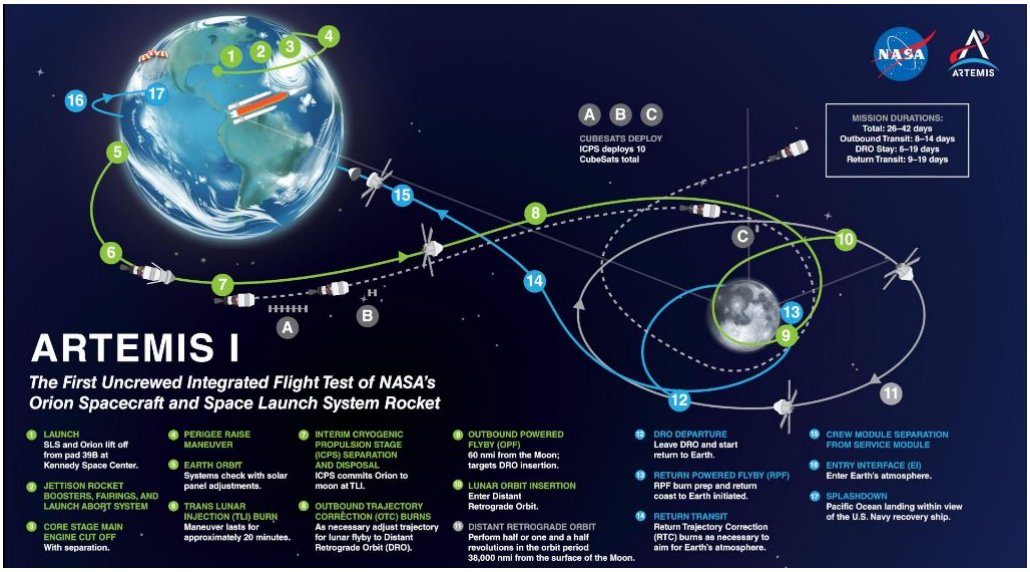


Figure 2.2.5: Artemis I trajectory [credit by NASA]

readily gained on the ground in preparation for critical rendezvous, proximity operations, docking, undocking operations on Artemis III. After this procedure, the crew will spend the remainder of the day-long orbit verifying system performance in the space environment. In HEO, they will assess the life support systems necessary to generate breathable air and remove metabolically produced water vapor and carbon dioxide. They will remove the Orion Crew Survival System suit they wear for launch and spend the rest of the in space mission in plain clothes until they put their suits again to prepare for entry into Earth's atmosphere and recovery from the ocean. After completing checkout procedures in HEO, Orion will perform the translunar injection (TLI) maneuver. With the ICPS having done most of the work to put Orion into HEO, Orion's service module will provide the last push to put the spacecraft on a path toward the Moon with a lunar free-return trajectory. The TLI will send the crew on an outbound trip of about four days and around the far side of the Moon. Orion will return on another four-day journey back home with a fuel- efficient trajectory that harnesses the relationship of the Earth-Moon gravity field. This ensures that after its trip around the far side of the Moon, Orion will be pulled back naturally by Earth's gravity, with no propulsive moves required. The day before the crew returns home, they will prepare for Earth entry, descent, and landing by pressurizing and testing the crew module propulsion systems and storing loose equipment before descent through Earth's atmosphere. On entry day, the crew will put on their pressurized spacesuits and strap into their seats before the crew module separates from the European service module. During reentry, the Orion spacecraft will be traveling at nearly 11 Km/s as it reenters Earth's atmosphere, which will slow it down. Parachutes will then slow it further for splashdown, ending a mission that will exceed over 1,000,000 km. Recovery forces, already positioned at the target landing zone, will be ready to recover the crew from the Pacific Ocean.

After more than two million miles traveled in space on NASA's deep space transportation systems during the previous missions, **Artemis III** is the upcoming mission that will mark the culmination of this rigorous testing. Orion spacecraft will once again

travel to the Moon, this time with the aim of making history by landing the first woman and next man on its surface. To enable a rapid return to the Moon, NASA is aiming to minimize the number of systems involved in landing humans on the surface. Therefore, the agency's procurement process for a commercially provided Human Landing System (HLS) has left the door open for proposals by private companies. Thus, early Artemis missions do not use the Gateway. The agency is planning to begin crewed exploration missions to the lunar surface in December 2025. These initial missions represent a human return to the Moon for the first time since 1972, but with several key differences, including the use of 21st century technologies and access to more parts of the Moon. The exact landing site for Artemis III astronauts will depend on various factors, including specific science objectives and launch dates. NASA's Lunar Reconnaissance Orbiter has provided high-resolution data and detailed mapping of the lunar surface, and through the CLPS initiative, the agency will gain even more surface data through science investigations that will help identify additional areas of interest for human exploration. To achive this, more supplies may be placed on the surface prior to crew arrival. Moreover, CLPS providers may be used to deliver pre-emplaced science instruments and equipment for use by the first human return crew while exploring the lunar surface. During the Artemis III surface expedition, the crew will characterize and document the regional geology, including small permanently shadowed regions, if available. They will collect a variety of samples to return to Earth for later research, such as rock samples to help date the sequence of impact events on the Moon, core tube samples to capture ancient solar wind trapped in regolith layers, and paired samples of material within and outside a permanently shadowed region to characterize the presence of volatiles and assess geotechnical differences between materials inside and outside of permanent shadows. While on the surface, the crew will live in the cabin of the ascent vehicle, which is the upper part of the landing system that they will use to get back to lunar orbit when the surface expedition concludes. NASA requires a minimum of two moonwalks during the Artemis III surface expedition, and is currently working to drive down the HLS vehicle mass to allocate more resources to spacesuit life support systems. The goal, if mass allows, is to conduct four planned extravehicular activities (EVAs) during the mission.

Following the Artemis III mission, NASA and its collaborators will undertake missions on and around the Moon, which will also serve as preparation for the mission durations and operations that will be necessary for human missions to Mars. This will involve expanding the Gateway's capabilities, achieving high confidence in commercial lunar landers departing from it, and establishing the Artemis Base Camp at the South Pole of the Moon. Artemis Base Camp represents our inaugural human settlement on the lunar frontier, and consists of three primary mission elements. The first element is the Lunar Terrain Vehicle (LTV), an unpressurized rover designed to transport suited astronauts around the lunar surface. The second element is the habitable mobility platform, a pressurized rover that enables long-duration trips away from Artemis Base Camp. This component is particularly important, as it is similar to the spacecraft that will be used in the first mission to Mars. Finally, the foundation surface habitat will serve as a home for four crew members on the lunar surface, anchoring Artemis Base Camp at the South Pole. Together, these elements, along with supporting infrastructure such as communications, power, radiation shielding, and waste disposal and storage planning, comprise a sustainable capability on the Moon that can be revisited and expanded upon over the coming decades. Furthermore,

Figure 2.2.6: Artistic representation of astronauts conducting EVAs

the testing of these systems will be crucial in preparing for human missions deeper into the Solar System.

Chapter 3

GN&C Elements

The Lunar landing mission requires a highly accurate Guidance Navigation and Control (GN&C) system to ensure a gentle and precise touchdown on the moon's surface, while meeting all necessary constraints during the descent phase. The system must be able to detect and avoid hazards, such as rocks and craters, through the Hazard Detection and Avoidance (HDA) system. It should reliably find an alternate landing site and command a re-targeting maneuver if the pre-defined landing site is unsafe. To achieve high precision position and attitude control, a vision-based relative navigation system is also needed, along with advances in sensor capability for surface feature detection. The navigation design must consider motion and attitude constraints, as well as time and altitude margin, to allow the sensors to see the targeted area. The trajectory guidance must ensure dynamic tridimensional retargeting to avoid potential uncertainties, and modifications capability of downrange and crossrange should be implemented. The major challenge comes from the need to comply with high precision position and attitude control requirements while using the available propulsion system, which is composed of a set of fixed-thrust main engines, a set of pulsed assist engines, and possibly the Reaction Control System.

In the first paragraph of the chapter, it will be presented the configuration of the Lunar Module for the Apollo missions and a brief description of the guidance law adopted for the descent and ascent phases. Following, the fundamental aspects of nowdays GN&C systems for landing missions on celestial bodies will be detailed.

3.1 Apollo Lunar Module GN&C

In this section, we will discuss the hardware and guidance laws utilized by the Apollo Lunar Module (LM) for its Guidance, Navigation, and Control system during the nominal descent and ascent trajectories. Figure 3.1.1 illustrates an overview of the Apollo LM GN&C system. It can seen that the onboard digital computer plays a central role during operations. The computer receives and transmits data and commands to and from other components and subsystems. The Inertial Measurement Unit (IMU) serves as the navigation sensor, comprising accelerometers and gyros to sense changes in velocity and attitude. The IMU information is sent to the computer. The LM optical system is a simple periscope that measures star direction for IMU alignment. The landing radar is mounted at the bottom rear of the LM, which measures local altitude, altitude rate, and components of horizontal velocity during descent. It has four radar beams, with one for providing

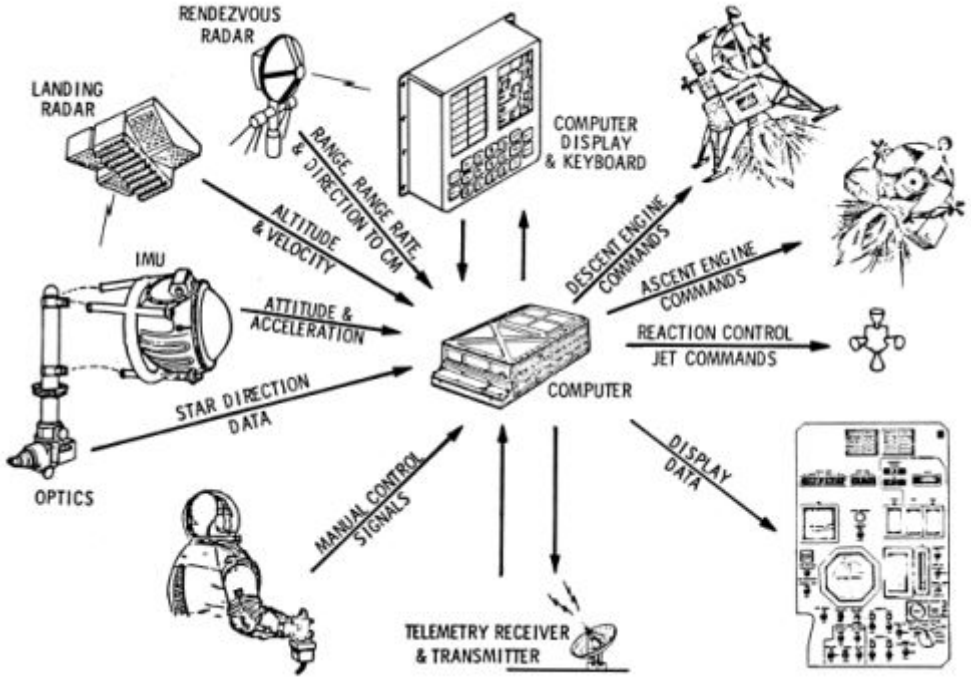


Figure 3.1.1: Apollo GN&C System Overview

ranging measurements and three for providing velocity measurements. The computer utilizes this input for the navigation of the lander and to calculate steering commands. The rendezvous radar tracks the Command and Service Module (CSM) in its lunar orbit, providing the computer with its direction, range, and range rate. The LM uses the rendezvous radar when it is on the lunar surface and during rendezvous in orbit. The LM has two system-driven displays, the computer display and keyboard, and the attitude error display. Additionally, the computer drives computed altitude rate and horizontal velocity component displays during the landing, and CM range and range rate displays during ascent and rendezvous. Finally, the computer sends steering commands to the descent and ascent engines and the Reaction Control System (RCS). The deep-throttling descent engine is capable of throttling between 10 and 60 percent for controlled operations or can be set to maximum, i.e., 45,000 Newton. The throttle can be controlled automatically or by manual commands from the crew. Additionally, the engine can move ±6 degrees from the center for slow attitude rate commands, which is automatically controlled by the digital autopilot. For high rate changes, the digital autopilot controls the RCS, consisting of four groups of four small 445 Newton control rockets. This RCS is also used during the ascent. The main ascent engine is a ±15,000 Newton fixed constant- thrust engine. Since only altitude, velocity, and orbit plane are required for targeting and not downrange position control, engine throttling is not required during ascent. Moreover,the LM has an abort guidance system, which serves as a redundant GN&C system in case of failure.

In order to successfully land on the lunar surface and complete the ascent and rendezvous maneuvers, the Apollo landers employed quadratic guidance laws to keep the spacecraft on the desired trajectory. These laws utilize acceleration commands that are a quadratic function of time. Figure 3.1.2 presents a simplified flow chart of the Apollo

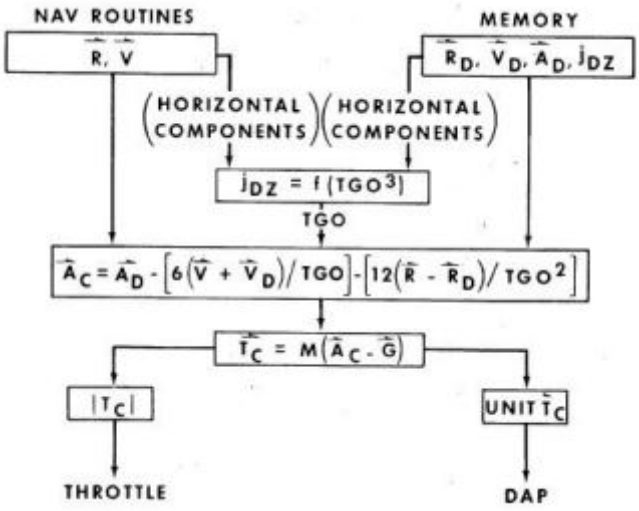


Figure 3.1.2: Apollo LM's descent guidance laws

guidance algorithm. During the descent phase, the navigation sensors monitor the current position and velocity vectors of the Lunar Module, denoted as r and v, respectively. The desired position, velocity, acceleration, and horizontal (down-range) component of the time derivative of acceleration vectors, represented as r_D, v_D, a_D, J_D respectively, are stored in the memory. The horizontal components of these current and desired state vectors are used in the Jerk equation to calculate the time to go (t_{GO}) from the current to desired conditions. Once the current state vector, desired state vector, and$t_{GO}$are known, the commanded guidance acceleration vector a_{cmd} can be calculated from the following quadratic guidance law:

$$a_{cmd} = a_D - (6(v + v_D))/t_{GO} - (12(r - r_D))/t_{GO}^2 \tag{3.1}$$

The above equation is a simplified version since the actual law contains multiplicative terms. At this point, The commanded thrust vector T_C is obtained by using the lander's mass m, the difference between the commanded acceleration and lunar gravity, $a_{cmd} - g$, and applying Newton's law. The magnitude of T_C is then utilized for the automatic throttling of the descent propulsion system. Several sequential programs were implemented to enable the Apollo spacecraft to achieve a successful powered descent. These programs, which are briefly explained below, can be observed in Figure 3.1.3.The first program, P-63 braking phase guidance, utilizes an ignition algorithm and the fundamental quadratic guidance law. The ignition algorithm determines the descent propulsion system's ignition time based on the surface range to the landing site. Once the propulsion system is ignited, the quadratic guidance law is used to guide the lander to the desired start conditions for the approach phase. When the time-to-go from the current state to the desired conditions reaches a preselected value, the guidance program automatically switches from P-63 to P-64. This is called the approach phase guidance, and also employs the quadratic guidance law. Trajectory shaping is provided throughout the approach phase by selecting desired targets, which provide suitable conditions for the beginning of the vertical descent, and by maintaining the trajectory angle at a value that allows the crew to use the window as a

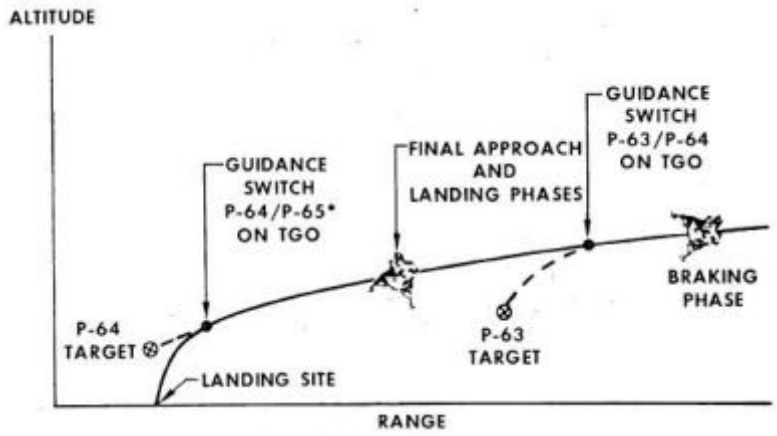


Figure 3.1.3: Overview of sequential programs (*Guidance P-65 is velocity nulling only)

guide to approach the landing site. If the crew decides to change the landing site, they can adjust the trajectory manually. Although the target point is not anticipated to be reached, it guides the lander to the desired point near or directly above the landing site. Again, once t_{GO} reaches a preselected value, the guidance program automatically switches to P-65, named velocity nulling guidance. It nullifies all velocity components to preselected values, allowing for an automatic vertical descent to the surface. No position control is utilized during this guidance mode. The vertical and horizontal velocity are controlled separately. The vertical guidance algorithm controls the altitude rate by adjusting the throttle, while the horizontal guidance algorithm nullifies the horizontal velocity by making small changes to the thrust angle. It assumes a constant vertical thrust component, with a thrust angle limit of 20 degrees. P-66 (rate of descent) and P-67 (manual guidance) are optional modes that the crew can use at any time during the different phases. During P-66, the crew controls the spacecraft's attitude while the computer maintains the throttle of the descent propulsion system at a desired value. This mode was typically utilized during late P-64 operation before switching to P-65 for manual control of the final touchdown position. P-67, instead, mode is only used when P-66 fails to function properly.

3.2 Trends in Autonomous Landing Technology

In the field of Guidance, Navigation, and Control systems for planetary landers, there are several noteworthy advancements being made, particularly in the areas of Hazard Detection and Avoidance systems and visual navigation. This section will provide a more detailed introduction to these technologies, enabling an analysis of the guidance requirements for future missions.

- **Vision-based Navigation for Precision Landing.** Vision-based technologies can improve the precision of lunar landings beyond what was previously achievable. Two types of visual navigation can be distinguished: relative vision-based navigation and absolute vision-based navigation. Absolute vision-based navigation involves improving the position estimate by matching features or landmarks extracted from an on-board camera image with features contained in a database stored on-board. Techniques differ in the type of features used, how these are matched,

and how the database is built. These techniques can use both imagery and topographic data. However, absolute navigation presents challenges related to illumination conditions, the generation of the feature database, and on-board implementation. Therefore, relative vision-based navigation is used in the final landing phase to precisely locate the vehicle relative to the features in the on-board camera images of the landing site. The high-precision achievable with visual navigation technologies allows for a reduction in landing ellipses from kilometers to hundreds of meters, and even to tens of meters or single-digit meter precision if previously landed surface assets are available for relative navigation. This represents a significant improvement over past autonomous lunar landings. Visual navigation, in combination with HDA, broadens the range of possible mission scenarios. However, it poses challenging new requirements for autonomous landers, such as the need to maintain a visual on the landing site throughout the landing phase and possible retargeting. These requirements are detailed in the next section.

- **Hazard Detection and Avoidance for Safe Landing.** The historical data indicated that the empirical safe landing probability was only about 80% percent at the beginning of the last decade. However, with sub-meter resolution imagery of the Moon's surface obtained by NASA's Lunar Reconnaissance Orbiter in recent years, significant progress has been made in automatic hazard detection. While hazards such as slopes, craters, boulders, or shadows can be identified to some extent from these orbital data, smaller hazardous features, like rocks, remain unresolved and can compromise touchdown safety. To improve the probability of safe landing, automatic Hazard Detection and Avoidance technologies are being developed. These systems are designed to characterize the sensed surface in terms of hazardous features within the designated landing area in real-time and determine whether the chosen landing site is safe. If not, the system will command a diversion to a safer landing site, reducing the probability of failed landing by at least a factor of 4 in any given terrain. From an engineering point of view, the successful implementation of HDA systems would have a positive impact, as previously deemed unsafe landing areas can now be targeted. Previous missions would only target landing sites considered to be 98% safe a priori, but with HDA, the safe landing probability for those areas previously only 50% safe can be increased to 98% as well. The major challenge for the HDA system is the illumination of the prospective landing sites, which can result in dark and patchy images. To overcome the limitations of camera-based hazard detection, an imaging LIDAR is used. However, imaging LIDARs often build an image of the terrain through a scanning mechanism, requiring compensation for the Lander's motion during image acquisition, stability of the Lander during image acquisition, accuracy of the velocity and attitude estimates used to correct the image, and the required processing power.

- **Requirements for the Guidance System.** The HDA and vision-based landing systems have several implications on the design of the guidance system for planetary landers. Pinpoint-landing ability is necessary to achieve precision landing, which requires the guidance mode to target a precise position on the surface. The on-board data processing required for HDA and visual navigation puts a heavy load on the CPU and memory, which limits guidance in terms of computational resources.

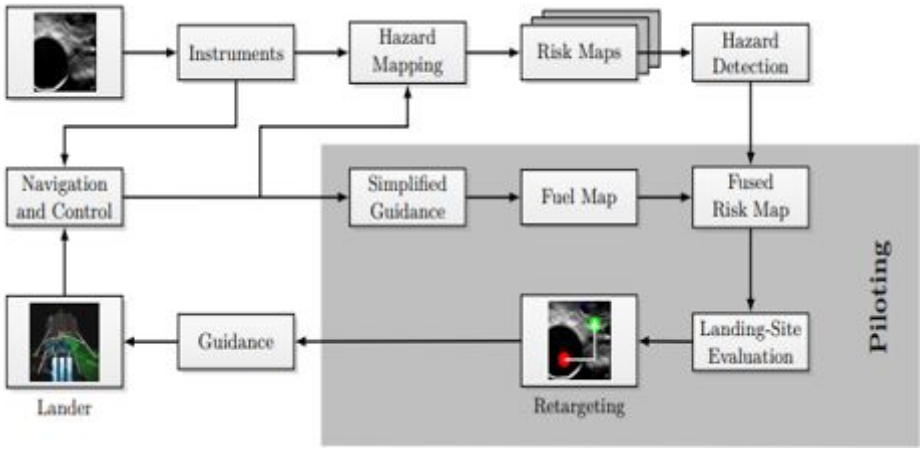


Figure 3.2.4: HDA Architecture

Thus, the guidance algorithm must work efficiently to optimize the use of resources. The guidance system must also be robust to maintain operational functionality and to performance against uncertainties within the HDA framework, which is crucial for safe landing probability. Adaptivity is another critical aspect since the guidance system must be capable of re- planning the trajectory to a new target in real-time. The HDA system may command diversions to alternative landing-site locations if the current site is deemed unsafe, and a safer one exists. To accommodate the landing sensors' requirements, a constrained trajectory must be generated, including path constraints to ensure that the landing site remains in the sensors' field of view, limits on the viewing angle, and thrust limitations that may affect engine throttling during retargeting. This is the direct interface to guidance, which receives the retargeting command and computes a new trajectory. The guidance law must also be fuel optimal to maximize feasible payload mass. This is of particular importance when used in combination with HDA, as diversion maneuvers may require extra propellant. These issues become clearer when considering guidance in the context of other GN&C system elements, such as instruments that map the body's surface and evaluate the safety of landing sites by creating risk maps. The guidance algorithm computes a reachability or fuel map to determine the range of feasible diversions, which is used by the piloting function to confirm the current landing site or determine a new one. Figure 3.2.4 represents the direct interface to guidance, which receives the retargeting command and computes a new trajectory.

Chapter 4

Optimal Control Theory

Optimal Control is a branch of applied mathematics that deals with finding the best control strategy for a dynamical system in order to optimize (i.e. minimize or maximize) a certain performance index, subject to constraints on the system's inputs and outputs. The applications of optimal control theory can be found in a wide range of fields. Quoting the famous Swiss mathematician Leonhard Euler:

> *"Since the building of the universe is perfect and is created by the wisdom creator, nothing arises in the universe in which one cannot see the sense of some maximum or minimum."*

For instance, Optimal control theory has been employed in medicine to control blood glucose levels in diabetic patients by designing personalized treatment plans that regulate glucose levels and minimize the risk of complications through the use of mathematical models and optimal control techniques. In neuroscience, optimal control methods can help understand the behavior of neurons and networks of neurons by modeling the dynamics of neural systems. In everyday life, the theory of optimal control plays a crucial role in regulating the global economy. Governments can use it to determine the optimal fiscal policies, considering the maximization of social welfare or the reduction of public debt. Moreover, the theory of optimal control is also used in natural resource management, such as fishing and agriculture, to maximize crop yield. Similarly, optimal control theory is used in the space field to design spacecraft trajectories that minimize travel time and/or fuel consumption, or other mission performance parameters. For example, in the case of the recent James Webb Space Telescope mission, optimal control theory was employed to develop the guidance and pointing controls, maximizing the precision of the telescope's positioning and minimizing the effects of vibrations and external disturbances, allowing the telescope to maintain a precise position and orientation in space for long observations of celestial objects. The main goal of this chapter is to introduce the basic concepts of optimal control theory, starting with the formulation of the optimal control problem. This section will also cover different methods for solving optimal control problems. Among all these categories, the indirect approach is widely used in space engineering due to its high level of precision and accuracy. For this reason, we shall employ an indirect method in this work, and apply variational techniques to the optimal control problem. In this way it will be possible to derive the necessary conditions for optimality, that will then be utilized to determine the optimal control law. Additionally, we shall introduce Pontryagin's

Maximum Principle (PMP) as a generalization of the fundamental theorem of the calculus of variations. Finally, the shooting method will be presented as a means to solve the Multi-Point Boundary Value Problem (MPBVP).

4.1 Optimal Control Theory

The primary objective of optimal control theory is to determine the control signals that will cause a process to adhere to the physical constraints and, at the same time, optimize a specific performance criterion. Later, we will provide a more detailed mathematical statement of the optimal control problem. However, before that, it is important to discuss the process of problem formulation. Defining an optimal control problem necessitates the creation of a model description of the process to be controlled, introducing the relevant system variables. It is also necessary to state the physical constraints to be considered, outlining the limitations for both the variables and the controls. Ultimately, it is essential to specify a performance measure that needs to be optimized, whether it is maximizing or minimizing it.

The Mathematical Model

Modeling the process is a crucial aspect of any control problem and involves obtaining a concise mathematical description that can accurately predict the response of the physical system to all anticipated inputs. This discussion will focus on systems described by ordinary differential equations in state variable form. In this regard, suppose that the *state variables* or *states* of the process at time t are denoted by $x_1(t), x_2(t), \ldots, x_n(t)$, while the *control inputs* to the process at time t are represented by $u_1(t), u_2(t), \ldots, u_m(t)$. Then, the system can be described using the following n first-order differential equations:

$$\begin{cases} \dot{x}_1 = a_1(x_1(t),\, x_2(t),\, \ldots,\, x_n(t),\, u_1(t),\, u_2(t),\, \ldots,\, u_m(t),\, t) \\ \dot{x}_2 = a_2(x_1(t),\, x_2(t),\, \ldots,\, x_n(t),\, u_1(t),\, u_2(t),\, \ldots,\, u_m(t),\, t) \\ \quad \cdot \\ \quad \cdot \\ \quad \cdot \\ \dot{x}_n = a_n(x_1(t),\, x_2(t),\, \ldots,\, x_n(t),\, u_1(t),\, u_2(t),\, \ldots,\, u_m(t),\, t) \end{cases}$$

Note that $\dot{x}(t)$ is in general a nonlinear time-varying function a of the states, the control inputs, and time. Thus, it is possible to define **x(t)** as the *state vector* of the system, and **u(t)** as the *control vector*. The state equations can then be written as:

$$\dot{x}(t) = \mathbf{a}(\mathbf{x}(\mathbf{t}),\, \mathbf{u}(\mathbf{t}),\, t) \tag{4.1}$$

This differential equations based on the state and control variables describe the time evolution of x(t) and u(t) between the initial and final states (*outer boundaries*). On the other side, the physical quantities that can be measured are called the *outputs* and are denoted by the vector $y(t)) = (y_1(t),\, y_2(t),\, \ldots, y_q(t))$.

Let the system be described by this equation for $t \in [t_0, t_f]$. The following definitions can be found:

- A history of control input values during the interval $[t_0, t_f]$ is denoted by $\mathbf{u}$ and is called a *control history*, or simply a *control*.

- A history of state values in the interval $[t_0, t_f]$ is called a *state trajectory* and is denoted by $\mathbf{x}$.

In control problems, it is often necessary to divide the trajectory of the process into a specific number of sub-intervals, referred to as *arcs*. This technique is used to ensure that the variables remain continuous within each arc. Specifically, the $j-th$ sub-interval starts at $t_{(j-1)+}$ and terminates at t_{j-}. Additionally, $x_{(j-1)+}$ and xj- represent the values assumed by the state variables during these instants. It is essential to note that the signs $'-'$ and $'+'$ denote the values of the state variables immediately before and after a specific point. This framework allows the consideration of possible discontinuities of the state variables, such as impulsive maneuvers, as well as internal constraints that may occur among the junction points between each contiguous arc (*inner boundaries*). By using arcs, a complex control problem can be simplified into a series of smaller, more manageable sub-problems.

Physical Constraints

Once a mathematical model has been chosen for a control problem, the subsequent step is to define the physical constraints imposed on both the state and control variables. Firstly, let us consider state constraints. Assuming that t0 is the time of leaving the starting point while tf is the time of arrival at the final position, $x(t_0)$ and $x(t_f)$ respectively represent the boundary conditions for the state variables within the interval $[t_0, t_f]$. Furthermore, additional constraints must be placed on the control inputs, which are bounded by some upper and lower limits that depend on the capability of the engine. In addition, if the system starts with a given amount of fuel or energy, denoted as G, and it cannot spend more than this value, then another possible constraint may be imposed as follow:

$$\int_{t_0}^{t_f} k_1 \, u(t) dt \leq \; G \tag{4.2}$$

which assumes that the rate of consumption is proportional to controls with constants of proportionality k_1.

Having outlined the common constraints encountered in control problems, it is necessary to define these concepts more precisely.

- A control history which satisfies the control constraints during the entire time interval $[t_0, t_f]$ is called an *admissible control*.

- A state trajectory which satisfies the state variable constraints during the entire time interval $[t_0, t_f]$ is called an *admissible trajectory*.

We shall denote the set of admissible controls by U, and the notation $\mathbf{u} \in U$ means that the control history $\mathbf{u}$ is admissible. On the other hand, the set of admissible state trajectories will be denoted by X, and $\mathbf{x} \in X$ means that the trajectory $\mathbf{x}$ is admissible. Admissibility is a crucial concept in control theory, as it restricts the range of values that can be taken on by both state and control variables. By imposing admissibility constraints, we limit our consideration to only those trajectories and control inputs that satisfy the physical

constraints and are thus feasible. Instead of evaluating all possible control histories and trajectories to determine the best one according to some performance criterion, we focus solely on admissible trajectories and control inputs. This approach not only simplifies the problem but also provides a more realistic and practical solution.

The Performance Measure

In order to evaluate the performance of a system quantitatively, the designer have to select a performance measure. An *optimal control* is defined as one that *minimizes* or *maximizes* the performance measure. In all that follows it will be assumed that the performance of a system is evaluated by a measure of the form

$$J = h\left(x\left(t_f\right),\ t_f\right) + \int_{t_0}^{t_f} g(x(t), u(t), t)dt \tag{4.3}$$

where h and g are scalar functions and the final time t_f may be specified or "$free$", depending on the problem statement. Starting from the initial state $x(t_0) = x_0$ and applying a control signal u(t), for $t \in [t_0, t_f]$, causes a system to follow some state trajectory. The performance measure represents the *cost functional* of the control problem, and it assigns a unique real number to each trajectory of the system. With the foundation we have established, we can now provide a formal statement of the optimal control problem as follows:

"Find an admissible control u^* which causes the system

$$\dot{x}(t) \ = \ a(x(t),\ u(t),\ t)$$

to follow an admissible trajectory x^* that minimizes or maximizes the performance measure J."

Thus, u^* is called an *optimal control* and x^* an *optimal trajectory*.

There are a few important points to be made here. Firstly, it is possible that we do not know in advance whether an optimal control exists. It may be impossible to find a control that is admissible and leads the system to follow an admissible trajectory. Therefore, in most cases, we attempt to find an optimal control rather than trying to prove that one exists, as existence theorems are rare. Secondly, even if an optimal control exists, it may not be unique. Non-unique optimal controls may make computational procedures more complicated, but they also provide the opportunity to choose from multiple controller configurations. This is advantageous to the designer, as they can consider additional factors such as cost, size, and reliability that may not have been incorporated in the performance measure. Finally, when we say that u^* causes the performance measure to be minimized or maximized, we mean that

$$J = h\left(x^*\left(t_f\right),\ t_f\right) + \int_{t_0}^{t_f} g\left(x^*(t), u^*(t), t\right) dt \ \leq \ h\left(x\left(t_f\right),\ t_f\right) + \int_{t_0}^{t_f} g\left(x(t), u(t), t\right) dt \tag{4.4}$$

for all $u \in U$, which make $x \in X$. The above inequality states that an optimal control and its trajectory cause the performance measure to have a value smaller than (or perhaps equal to) the performance measure for any other admissible control and trajectory. Thus, we are seeking the absolute or global minimum of J, not merely local minima. It is worth noting that determining the global minimum can be achieved by finding all local

minima and selecting the one that yields the smallest value for the performance measure. Additionally, it should be emphasized that the theory developed here can be applied to problems where the objective is to maximize some measure of system performance, as this is equivalent to minimizing the negative of the performance measure. Therefore, we will use the terms "minimizing" and "maximizing" interchangeably when referring to the performance measure. With this in mind, we can now introduce the following definition:

- If a functional relationship of the form

$$u^*(t) = f(x(t), t) \tag{4.5}$$

  can be found for the optimal control at time t, then the function f is called the *optimal control law*, or the *optimal policy*.

Notice that the optimal control law specifies how to generate the control value at time t from the state value at time t. The presence of t as an argument of f indicates that the optimal control law may be time-varying.

In coclusion, let's introduce the notions of *controllability* and *osservability*.

- If there is a finite time $t_1 > t_0$ and a control u(t), $t \in [t_0, t_1]$, which transfers the state x_0 to the origin at time t_1, the state x_0 is said to be controllable at time t_0. If all values of x_0 are controllable for all t_0, the system is *completely controllable*, or simply *controllable*.

The concept of controllability is crucial since we will be analyzing problems that involve moving a system from any initial state to a specific target state while minimizing a performance measure. Therefore, the ability to control the system is a necessary prerequisite for the existence of a solution.

- If by observing the output y(t) during the finite time interval $[t_0, t_1]$ the state $x(t_0) = x_0$ can be determined, the state x_0 is said to be observable at time t_0. If all states x_0 are observable for t_0, the system is called completely *completely observable*, or simply *observable*.

4.2 Survey of Methods

The evolution of numerical methods for trajectory optimization have closely paralleled the development of digital computers and space exploration. The latter presented scientists and engineers with complex technical challenges, while the former provided the means to tackle such problems. The purpose of this section is to provide an overview of the current state of trajectory optimization. Specifically, we introduce direct and indirect methods as a means of classification. However, it should be noted that some techniques do not fit neatly into either category. We will then discuss the advantages and drawbacks of each method and and then suggest that the techniques may ultimately merge hybrid methods. Finally, we will focus on a specific type of method, namely evolutionary algorithms.

Indirect Methods

Indirect methods (IM) utilize the first-order necessary conditions of optimality derived from the application of the variational approach on optimal control problem (OCP). These methods aim to identify strong optimal control and state variables. The application of these conditions to any given OCP typically results in a Two-Point Boundary Value Problem or, in cases where multiple intervals of integration are present, a Multi-Point Boundary Value Problem. These problems can be solved using an appropriate boundary value solver. Indirect methods are known to provide highly accurate solutions, even for problems with a large number of continuous-valued controls and states, provided that the boundary value solver converges. The most basic IM is the *shooting method*. In a typical shooting approach, an initial guess is made for the unknown boundary conditions at one end of the interval. Upon reaching tf, the terminal conditions obtained from numerical integration are compared to the known terminal conditions. If the integrated terminal conditions differ from the known terminal conditions by more than a specified tolerance, the unknown initial conditions are adjusted, and the process is repeated until the difference between the integrated terminal conditions and the required terminal conditions is less than some specified threshold. However, the shooting method poses poor characteristics when the optimal control problem is hypersensitive. To overcome the numerical difficulties of the simple shooting method, a modified method, called the *multiple-shooting method*, has been developed. In this case, the time interval $[t_0, t_f]$ is divided into several sub-intervals, and the shooting method is applied over each subinterval $[t_j, t_{j+1}]$. The multiple-shooting method requires extra variables to be introduced into the problem, such as the values of the state and adjoint at the interface points. Despite the increased size of the problem due to these extra variables, this method is an improvement over the standard shooting method, as the sensitivity to errors in the unknown initial conditions is reduced because integration is performed over significantly smaller time intervals. On the other side, one major drawback of indirect methods is the requirement to derive the first-order necessary conditions analytically, which can be bukly for high-dimensional systems and requires at least some knowledge of optimal control theory to deduce the first-order necessary conditions properly. Indirect methods also suffer from numerical difficulties, making them unfeasible in many practical scenarios. For instance, a good initial guess for the approximated costates is required to achieve convergence. The construction of a good initial guess is complicated and requires an estimate of the switching structure of the linear controls. If the guess is not accurate enough, the solution will not converge. Moreover, problems with state constraints might be intractable for indirect methods, leading to a large class of problems that cannot be solved using these methods. In such cases, alternatives such as direct methods for optimal control are required.

Direct Methods

Direct methods are commonly used in trajectory optimization, as they transform the optimal control problem into a *nonlinear programming* (NLP) problem. This involves discretizing the trajectory and approximating state and control variables using a piecewise constant parameterization for each trajectory arc. The resulting parameter optimization problem can be approached through various discretization techniques, depending on the nature of the problem. While direct methods can handle any problem formulation (dynamics and constraints) with minimal programming effort, they can be computationally

expensive and may produce inaccurate results due to the large number of parameters required for an accurate trajectory description. However, one major advantage of the direct approach is that the user need not be concerned with adjoint variables or switching structures. Instead, differential equations are transformed into a finite set of equality constraints, and the optimal control problem is solved strictly through the numerical integration of these constraints. Nevertheless, direct methods are known to produce less accurate solutions than indirect methods. For instance, in our numerical experimentation with aeronautics problems, we found that direct methods tend to yield minimum functional values with relatively low accuracy (i.e., errors of about one percent). While seemingly small, such errors could be significant in the context of space flight missions. Furthermore another disadvantage of direct methods is that the discretized optimal control problems may sometimes have multiple minima, leading the direct methods to converge to one of these "pseudominima".

Hybrid Methods

In order to address the limitations of the direct method for solving optimal control problems, a hybrid approach has been proposed which combines the advantages of the direct method with the accuracy and reliability of the multiple shooting method. However, this approach requires careful consideration of several factors, including the adjoint differential and the optimal control laws. Additionally, the proper selection of the multiple shooting nodes, initial values of the adjoint variables, and switching structure must be determined in advance. The grid points used in the direct method can serve as a good starting point for the multiple shooting nodes. Moreover, reliable estimates for the adjoint variables can be obtained from the parameters and the Lagrangian function of the nonlinear program, which do not explicitly appear in the direct formulation. This hybrid approach has been successfully applied to several test examples as well as new real-life problems with unknown solutions, such as the maximum range trajectory of a hang glider.

Evolutionary Algorithms

Evolutionary Algorithms (EA) belong to the heuristic optimization methods. They are inspired by Darwin's theory evolutionary genetic processes, such as reproduction, mutation, recombination, and selection. Among the various techniques available, the most widely used EA is the Genetic Algorithm (GA). Initially, the code randomly generates a set of potential solutions to the problem at hand. The quality of these solutions is determined by the *fitness function*, which plays the role of assessing their fitness. The *candidate solutions* to the optimization problem represent individuals in a population. Analogous to biological processes, it is as if each chromosome has a particular gene that confers a specific characteristic. In nature, genes can give rise to mutations, modifying both the genotype and the phenotypic manifestation. In other words, the distinctive trait associated with that gene can change and evolve. Evolutionary algorithms recombine solutions using a crossover mechanism, generating additional populations. This process is iterated to form generations of populations of solutions where the identification, through the fitness function, of the characteristic that survived the original populations allows the selection of the best solution for the optimal control problem.

4.3 Calculus of Variations

In Paragraph 4.1 the ultimate objective of optimal control problem was identified as obtaining a controller that achieves desirable system performance. It is common for other factors, such as weight, volume, cost, and reliability, to influence controller design, leading to compromises between performance requirements and implementation considerations. Once the mathematical model of the system is established, and a performance measure is chosen, the next step is to find a control function that minimizes this criterion. Two main methods are used for this maximization: the maximum principle of Pontryagin and the method of dynamic programming developed by R. E. Bellman. The method of dynamic programming involves a functional equation that can be solved using a digital computer, but this approach will not be considered in this study. Instead, the variational approach of Pontryagin leads to a nonlinear Two-Point or Multi- Point Boundary Value Problem that must be solved to obtain an optimal control.

Before delving further, an introduction to the calculus of variations is necessary. It is a mathematical branch that proves to be extremely useful in solving optimization problems. Queen Dido of Carthage was apparently the first person to attack a problem that can readily be solved by using variational calculus. But even if the history of the calculus of variations can be traced back to ancient Greeks, substantial progress was made in the seventeenth century in Western Europe. Sir Isaac Newton used variational principles to determine the shape of a body moving in air that experiences the least resistance, while Johann Bernoulli posed the brachistochrone problem, which involves finding the shape of the wire that causes a bead to move from point A to point B under the influence of gravity in minimum time. Thus, the calculus of variations is concerned with finding the optimal curve that satisfies some criterion. For example, in the case of Queen Dido, having been promised all of the land she could enclose with a bull's hide, cleverly cut the hide into many lengths and tied the ends together. Having done this, her problem was to find the closed curve with a fixed perimeter that encloses the maximum area. The calculus of variations enables us to prove that a circle is the optimal solution to this problem. In optimal control problems, we seek a control function that minimizes a performance measure. The connection between the calculus of variations and optimal control should now be apparent.

In optimal control problems the objective is to determine a function that minimizes a specified cost functional, i.e. the performance measure. The analogous problem in calculus is to determine a point that yields the minimum value of a function. In the present section, we will present new concepts related to functionals by leveraging some well-known findings from function theory. To begin, let us review the definition of a function.

- A *function* f is a rule of correspondence that assigns to each element q in a certain set D a unique element in a set R. D is called the *domain* of f and R is the *range*.

The definition of a functional parallels that of a function.

- A *functional* J is a rule of correspondence that assigns to each function x in a certain class Ω a unique real number. Ω is called the *domain* of the functional, and the set of real numbers associated with the functions in Ω is called the *range* of the functional.

Notice that the domain of a functional is a class of functions; intuitively, we might say that a functional is a "function of a function".

Next, let us define the norm of a function.

- The *norm* of a function is a rule of correspondence that assigns to each function $x \in \Omega$, defined for $t \in [t_0, t_f]$, a real number. The norm of x, denoted by ||x||, satisfies the following properties:

  1. $||x|| > 0$ *or* $||x|| = 0$ if and only if $x(t) = 0$ for all $t \in [t_0, t_f]$.

  2. $||\alpha x|| = ||\alpha||\ ||x||$ for all real numbers α.

  3. $||x_1 + x_2|| \ \leq \ ||x_1|| + ||x_2||$.

The norm is useful to compare the closeness of two functions y and z that are defined for $t \in [t_0, t_f]$, as $x(t) = y(t) - z(t)$. Intuitively speaking, the norm of the difference of two functions should be zero if the functions are identical, small if the functions are "close," and large if the functions are "far apart".

In order to consider extreme values of a function, we now define the concept of an *increment*.

- If q and $q + \Delta q$ are elements for which the function f is defined, then the increment of f, denoted by Δf, is

$$\Delta f = f(q + \Delta q) - f(q) \tag{4.6}$$

Notice that Δf depends on both q and Δq, in general, so to be more explicit we would write $\Delta(q, \Delta q)$.

In an analogous manner, we next define the increment of a functional.

- If x and $x + \delta x$ are functions for which the functional J is defined, then the increment of J, denoted by ΔJ, is

$$\Delta J = J(x + \delta x) - J(x) \tag{4.7}$$

Again, to be more explicit, we would write $\Delta J(x, \delta x)$ to emphasize that the increment depends on the functions x and δx. δx is called the *variation* of the function x.

Let us proceed with a review of the definition of an extreme value of a function, which will then enable us to define the maxima and minima of functionals.

- A function f with domain D has a *relative extremum* at the point q^* if there is an $\varepsilon > 0$ such that for all points q in D satisfy $||q - q^*|| < \varepsilon$ the increment of f has the same sign. If

$$\Delta f = f(q) - f(q^*) \ \geq \ 0, \tag{4.8}$$

  $f(q^*)$ is a *relative minimum*; if

$$\Delta f = f(q) - f(q^*) \ \leq \ 0, \tag{4.9}$$

  $f(q^*)$ is a *relative maximum*.

If these equations are satisfied for arbitrarily large ε, then $f(q^*)$ is a *global*, or *absolute, minimum/maximum*. Typically, the strategy for finding extrema of a function is to locate points where the differential of the function is equal to zero, which is a necessary condition for an extremum at an interior point of the domain. Assuming that such points exist and can be computed, one can then analyze the behavior of the function in the vicinity of these points.

Next, consider a functional J which is defined for all functions x in a class Ω.

- A functional J with domain Ω has a *relative extremum* at the point x^* if there is an $\varepsilon > 0$ such that for all functions x in Ω satisfy $||x - x^*|| < \varepsilon$ the increment of J has the same sign. If

$$\Delta J = J(x) - J(x^*) \geq 0, \tag{4.10}$$

  $J(x^*)$ is a *relative minimum*; if

$$\Delta J = J(x) - J(x^*) \leq 0, \tag{4.11}$$

  $J(x^*)$ is a *relative maximum*.

If these equations are satisfied for arbitrarily large ε, then $J(x^*)$ is a *global*, or *absolute, minimum/maximum*. x^* is called an *extremal*, and $J(x^*)$ is referred to as an *extremum*.

At this point, the *Fundamental Theorem of the Calculus of Variations* can be introduced. The fundamental theorem used in finding extreme values of functions is the necessary condition that the differential vanish at an extreme point. In variational problems, the analogous theorem is that the variation must be zero on an extremal curve, if there are no bounds imposed on the curves. Let x be a vector function in the class Ω, and $J(x)$ be a differentiable functional of x. Assume that the functions in Ω are not constrained by any boundaries. The fundamental theorem of the calculus of variations states:

- If x^* is an extremal, the variation of J must vanish on x^*; that is:

$$\delta J(x^*, \delta x) = 0, \tag{4.12}$$

  for all admissible δx.

The fundamental theorem of the calculus of variations yields a set of differential equations, known as the *Eulerequations*, that are applicable to a specified functional regardless of the boundary conditions. However, unlike in the case of the calculus of variations, in control problems the trajectory of the system is determined by the applied control, and the optimization is constrained by the dynamics of the process. As a result, the differential equations obtained in control problems are generally nonlinear and have boundary values that are subject to constraints imposed by the system's dynamics.

4.4 Pontryagin's Maximum Principle

We will now apply the methods introduced in the previous section to establish necessary conditions for optimal control. The problem at hand involves $n+m$ functions that must satisfy n differential equations. Initially, we assume that the state and control regions are unbounded and that the initial conditions $x(t_0) = x_0$ and the initial time t_0 are given.

The m control inputs represent the independent functions. However, by assuming that h is a differentiable function, we can express the optimization problem as follows:

$$h\left(x\left(t_f\right),t_f\right) = \int_{t_0}^{t_f} \frac{d}{dt}\left[h\left(x\left(t\right),t\right)\right]\, dt + \, h\left(x\left(t_0\right),t_0\right) \tag{4.13}$$

so that the performance measure can be expressed as

$$J\left(u\right) = \int_{t_0}^{t_f} \left\{ g\left(x\left(t\right),\, u\left(t\right),\, t\right) + \, \frac{d}{dt}\left[h\left(x\left(t\right),t\right)\right] \right\} dt + \, h\left(x\left(t_0\right),t_0\right) \tag{4.14}$$

Since $x(t_0)$ and t_0 are fixed, the minimization does not affect the last term of the above equation, so we need consider only the functional

$$J\left(u\right) = \int_{t_0}^{t_f} \left\{ g\left(x\left(t\right),\, u\left(t\right),\, t\right) + \, \frac{d}{dt}\left[h\left(x\left(t\right),t\right)\right] \right\} dt \tag{4.15}$$

Using the chain rule of differentiation, we find that this becomes:

$$J\left(u\right) = \int_{t_0}^{t_f} \left\{ g\left(x\left(t\right),\, u\left(t\right),\, t\right) + \, \left[\frac{\partial h}{\partial x}\left(x\left(t\right),t\right)\right]^T \dot{x}(t) + \, \frac{\partial h}{\partial t}\left(x\left(t\right),t\right) \right\} dt \tag{4.16}$$

At this point, to include the differential equation constraints, we form the *augmented cost functional*:

$$\begin{aligned} J\left(u\right) = \int_{t_0}^{t_f} \{\, g(x\left(t\right),\, u\left(t\right),\, t\right) + \, \left[\frac{\partial h}{\partial x}\left(x\left(t\right),t\right)\right]^T \dot{x}(t) \\ + \frac{\partial h}{\partial t}\left(x\left(t\right),t\right) + p^T(t)[\, a\left(x\left(t\right),\, u\left(t\right),\, t\right) - \dot{x}(t)]\}dt \end{aligned} \tag{4.17}$$

by introducing the *Lagrange multipliers* $p_1(t), p_2(t), ..., p_n(t)$. Let's define

$$\begin{aligned} g_a\left(x\left(t\right),\, \dot{x}\left(t\right),\, u\left(t\right),\, p\left(t\right),t\right) \, &= g\left(x\left(t\right),\, u\left(t\right),\, t\right) \\ + \left[\frac{\partial h}{\partial x}\left(x\left(t\right),t\right)\right]^T \dot{x}\left(t\right) &+ \frac{\partial h}{\partial t}\left(x\left(t\right),t\right) + p^T(t)[a\left(x\left(t\right),\, u\left(t\right),\, t\right) - \dot{x}\left(t\right)] \end{aligned} \tag{4.18}$$

so that

$$J_a\left(u\right) = \int_{t_0}^{t_f} g_a\left(x\left(t\right),\, \dot{x}\left(t\right),\, u\left(t\right),\, p\left(t\right),t\right)\, dt \tag{4.19}$$

We shall assume that the end points at $t = t_f$ can be specified or free. To determine the variation of J_a, we introduce the variations $\delta x,\ \delta \dot{x},\ \delta u,\ \delta p,$ *and* δt_f. Therefore, as

discussed in Section 4.3, on an extremal we have

$$\begin{aligned}\delta J_a(u^*) = 0 = &\left[\frac{\partial g_a}{\partial \dot{x}}\left(x^*(t_f),\ ,\ u^*(t_f),\ p^*(t_f),t_f\right)\right]^T \delta x_f(t) \\ &+[g_a\left(x^*(t_f),\ \dot{x^*}(t_f),\ u^*(t_f),\ p^*(t_f),t_f\right) \\ &-\left[\frac{\partial g_a}{\partial \dot{x}}\left(x^*(t_f),\ \dot{x^*}(t_f),\ u^*(t_f),\ p^*(t_f),t_f\right)\right]^T \dot{x^*}(t_f)]\delta t_f \\ &+[\int_{t_0}^{t_f}\left[\frac{\partial g_a}{\partial \dot{x}}\left(x^*(t),\ \dot{x^*}(t),\ u^*(t),\ p^*(t),t\right)\right]^T dt \\ &-\frac{d}{dt}\left[\frac{\partial g_a}{\partial \dot{x}}\left(x^*(t),\ \dot{x^*}(t),\ u^*(t),\ p^*(t),t\right)\right]^T]\delta x(t) \\ &+\left[\frac{\partial g_a}{\partial u}\left(x^*(t),\ \dot{x^*}(t),\ u^*(t),\ p^*(t),t\right)\right]^T \delta u(t) \\ &+\left[\frac{\partial g_a}{\partial p}\left(x^*(t),\ \dot{x^*}(t),\ u^*(t),\ p^*(t),t\right)\right]^T \delta p(t) \end{aligned} \tag{4.20}$$

We first observe that

$$\dot{x^*}(t) = a(x^*(t),\ u^*(t),t) \tag{4.21}$$

must be satisfied by an extremal so that the coefficient of $\delta p(t)$ is zero. The Lagrange multipliers are arbitrary, so let us select them to make the coefficient of $\delta x(t)$ equal to zero, that is,

$$\dot{p}^*(t) = -\left[\frac{\partial a}{\partial x}\left(x^*(t),\ u^*(t),t\right)\right]^T p^*(t) - \frac{\partial g}{\partial x}(x^*(t),\ u^*(t),t) \tag{4.22}$$

From now on, we will refer to this set of equations as the *costate equations* and p(t) the *costate* or *adjoint variables*.

The remaining variation $\delta u(t)$ is independent, so its coefficient must be zero:

$$0 = \frac{\partial g}{\partial u}(x^*(t),\ u^*(t),t) + \left[\frac{\partial a}{\partial u}\left(x^*(t),\ u^*(t),t\right)\right]^T p^*(t) \tag{4.23}$$

It's noted that even when the admissible controls are bounded, only this last equation is modified. However, there are still the terms outside the integral to deal with. Therefore, as the variation needs to be null, we have:

$$\begin{aligned}&\left[\frac{\partial h}{\partial x}\left(x^*(t_f),t_f\right) - p^*(t_f)\right]^T \delta \mathbf{x}_f \\ &+\left[g\left(x^*(t_f),\ u^*(t_f),t_f\right) + \frac{\partial h}{\partial t}\left(x^*(t_f),t_f\right) + p^{*T}(t_f)[a\left(x^*(t_f),u^*(t_f),t_f\right)]\right]\delta \mathbf{t}_f = 0\end{aligned} \tag{4.24}$$

The first three equations are the *necessary conditions* we set out to determine. Notice that these necessary conditions consist of a set of $2n$ first-order differential equations and a set of m algebraic relations which must be satisfied throughout the interval $[t_0, t_f]$. The

solution of the state and costate equations will contain $2n$ constants of integration. To evaluate these constants, we use then equations $x*(t_0) = x_0$ and an additional set of n or $n+1$ (depending on whether or not tf is specified) relationships. Thus, we are faced with a two-point boundary-value problem, as expected.

In the following, we will use the function H, called the *Hamiltonian*, for convenience. It is defined as:

$$H(x(t),\ u(t),\ p(t),t)\ =\ g(x(t),\ u(t),\ t)+\ p^T(t)\left[a(x(t),\ u(t),\ t)\right] \tag{4.25}$$

Using this notation, we can write the necessary conditions for optimality, named *Euler-Lagrange equations*, as follows:

$$\dot{x}^*(t) = \frac{\delta H}{\delta p}(x^*(t),\ u^*(t), p^*(t), t) \tag{4.26}$$

$$\dot{p}^*(t) = -\frac{\delta H}{\delta x}(x^*(t),\ u^*(t), p^*(t), t) \tag{4.27}$$

$$0 = \frac{\delta H}{\delta u}(x^*(t),\ u^*(t), p^*(t), t) \tag{4.28}$$

for all $t \in [t_0, t_f]$. And

$$\left[\frac{\partial h}{\partial x}\left(x^*\left(t_f\right), t_f\right) - p^*\left(t_f\right)\right]^T\ \delta\mathbf{x}_f + \left[H\left(x^*\left(t_f\right),\ u^*\left(t_f\right), p^*\left(t_f\right) t_f\right) + \frac{\partial h}{\partial t}\left(x^*\left(t_f\right), t_f\right)\right]\ \delta\mathbf{t}_f = 0 \tag{4.29}$$

is the *boundary condition equation*.

Up to this point, the admissible controls have been assumed to have no boundary restrictions. However, in practical systems, such limitations frequently arise. Physically realizable controls often have constraints on their magnitude. For instance, the thrust of a rocket engine cannot surpass a certain value, or attitude control mass expulsion systems can provide only limited torque. Constraints may also arise due to safety or structural limitations. Thus, it is necessary to investigate the impact of control constraints on the fundamental theorem presented in the previous section, and subsequently, to illustrate the modified necessary conditions. This generalization of the fundamental theorem leads to the *Pontryagin's Maximum Principle*.

By definition, the control u^* causes the functional J to have a relative maximum if

$$\Delta J = J(u) - J(u^*) \leq\ 0 \tag{4.30}$$

for all admissible controls sufficiently close to u^*.

If we let $u = u^* + \delta u$, the increment in J can be expressed as:

$$\Delta J(u^*, \delta u) = \delta J(u^*, \delta u)\ +\ \textit{higher order terms} \tag{4.31}$$

Considering all admissible variations with $||\delta u||$ small enough so that the sign of ΔJ is determined by δJ, it can be seen that for the control problem the necessary conditions for u^* to maximize J is

$$\delta J(u^*, \delta u) \leq\ 0 \tag{4.32}$$

if u^* lies on the boundary during any portion of the time interval $[t_0, t_f]$, and

$$\delta J(u^*, \delta u) = 0 \tag{4.33}$$

if u^* lies within the boundary during the entire time interval $[t_0, t_f]$.

Now, let us examine how this modification affects the necessary conditions derived using the assumption of unconstrained admissible control values. If we consider these equations and the definition of Hamiltonian, if the state as well as the boundary condition equations are satisfied, and if $p^*(t)$ is selected so that the coefficient of $\delta x(t)$ in the integral is identically zero, the increment of J is:

$$\Delta J(u^*, \delta u) = \int_{t_0}^{t_f} \left[\frac{\delta H}{\delta u}(x^*(t),\ u^*(t), p^*(t), t) \right]^T \delta u(t)\, dt + \textit{ higher order terms} \tag{4.34}$$

The integrand is the first-order approximation to the change in H caused by a change in u alone. Therefore:

$$\Delta J(u^*, \delta u) = \int_{t_0}^{t_f} \left[H(x^*(t),\ u^*(t) + \delta u(t), p^*(t), t) - H(x^*(t),\ u^*(t), p^*(t), t) \right] dt + \textit{ higher order terms} \tag{4.35}$$

If $u^* + \delta u$ is in a sufficiently small neighborhood of u^* ($||\delta u|| < \beta$) then the higher order terms are small, and the integral dominates the expression for ΔJ. Thus, for u^* to be a maximizing control it is necessary that:

$$\int_{t_0}^{t_f} \left[H(x^*(t),\ u^*(t) + \delta u(t), p^*(t), t) - H(x^*(t),\ u^*(t), p^*(t), t) \right] dt \leq 0 \tag{4.36}$$

for all *admissible* δu, such that $||\delta u|| < \beta$.

We assert that to satisfy the above equation for all admissible δu in the specified neighborhood, it is necessary that:

$$H(x^*(t),\ u^*(t) + \delta u(t), p^*(t), t) \leq H(x^*(t),\ u^*(t), p^*(t), t) \tag{4.37}$$

for all *admissible* δu(t) and for all $t \in [t_0, t_f]$.

In conclusion, a necessary condition for u^* to minimize the functional J is

$$H(x^*(t),\ u^*(t), p^*(t), t) \geq H(x^*(t),\ u(t), p^*(t), t) \tag{4.38}$$

for all $t \in [t_0, t_f]$ and for all admissible controls.

This equation, which indicates that *an optimal control must maximize the Hamiltonian*, is called *Pontryagin's Maximum Principle*. By substituting it into the third equation of the first set of equations, the necessary conditions for an optimal control are found. The principle, although derived for controls with values in a closed and bounded region, can also be applied to problems in which the admissible controls are unbounded, as seen in the previous case. It should be noted that we have only established a necessary condition for optimality, which is not sufficient in general. An optimal control must satisfy the Pontryagin's Maximum Principle. However, there may be controls that satisfy the maximum

principle that are not optimal. If the Hamiltonian can be expressed in the following matrix form:

$$\frac{\partial^2 H}{\partial u^2}(x^*(t),\ u^*(t), p^*(t), t) = R(t) \tag{4.39}$$

If R(t) is *positive definite*, then the control

$$u*(t) = -R^{-1}(t)c(x*(t), p*(t), t) \tag{4.40}$$

maximizes (globally) the Hamiltonian. c is an $mx1$ array that does not have any terms containing $u(t)$.

Thus, now we have necessary and sufficient conditions for $\mathbf{H}(x^*(t),\ u^*(t), p^*(t), t)$ to be an *absolute maximum* and $u^*(t)$ the *optimal control law*. Then, the maximum principle provides a useful starting point for obtaining numerical solutions to optimal control problems. It allows us to determine the form of the optimal control (if it exists) and provides a statement of the two-point boundary-value problem, which, when solved, yields an explicit relationship for the optimal control.

However, if there exists a time interval $[t_1, t_2]$ during which the necessary condition of the maximum principle provides no information about the relationship between $u^*(t)$, $x^*(t)$, *and* $p^*(t)$, then the problem is said to be *singular*. This interval is called an interval of singularity or *singular interval*, and the corresponding trajectory is called a *singular arc*. To identify singular intervals, we examine the Hamiltonian and determine whether there are situations in which the minimum principle does not yield sufficient information to determine the relationship between $u^*(t)$, $x^*(t)$, *and* $p^*(t)$). If such a situation occurs, we can use the fact that the Hamiltonian must be zero (and that its derivatives H', H'', ... are equal to zero) to determine other necessary conditions for the existence of singular intervals and, consequently, the optimal control law.

4.5 Boundary Value Problem

The PMP applied to a system typically results in a Two-Point Boundary Value Problem (TPBVP) or a Multi- Point Boundary Value Problem (MPBVP) in the case of multiple integration intervals. To solve this problem, initial values of the unknown variables are sought to satisfy the boundary conditions by integrating the differential equations. The integration intervals are generally divided into several sub- intervals, and and each of them can be characterized by different differential equations. The duration of various sub-intervals is often unknown and boundary conditions may be nonlinear, involving both outer and inner boundaries. At the inner boundaries, variables may be discontinuous and their values unknown. The main challenge of indirect methods is thus finding a solution to the MPBVP. This is where the discretization process begins. To deal with the unknown duration of the sub-intervals of integration, a change of independent variable is applied. For each sub-interval j, time is replaced by:

$$\varepsilon = j - 1 + \frac{t - t_{j-1}}{t_j - t_{j-1}} \tag{4.41}$$

where $t_j - t_{j-1}$ is the duration of the sub-interval. By differentiating this expression with respect to time, we obtain:

$$\frac{d\varepsilon}{dt} = \frac{1}{t_j - t_{j-1}} \tag{4.42}$$

which will be useful later on.

By using this technique, the limits of integration for each sub-interval are established and are represented by consecutive integer values of the new independent variable ε at the boundaries. The process for solving the MPBVP using the shooting method is described below for a generic system:

$$\frac{dy}{dt} = f(y(t), t) \tag{4.43}$$

where the state variables and the adjoint variables are now directly grouped in the vector $y(t) = (x(t), p(t))$. Given that the problem is also characterized by constant parameters, such as the duration of the sub-intervals, it is expedient to create a new vector $z(t) = (y(t), c)$ that aggregates both state and adjoint variables and the constant parameters, with c representing the vector of constant parameters, which includes all unknown quantities that are not time-varying. Introducing the change of independent variable ε, the following definition of the vector z is obtained:

$$z(\varepsilon) \;=\; (y(\varepsilon),\, c) \tag{4.44}$$

Thus, the state equation becomes:

$$\frac{dz}{d\varepsilon} \;=\; f(z(\varepsilon),\, \varepsilon) \tag{4.45}$$

where the second term of the equation is obtained by taking the derivative with respect to ε:

$$\frac{dz}{d\varepsilon} \;=\; (\frac{dy}{d\varepsilon}, \frac{dc}{d\varepsilon}) \tag{4.46}$$

Regarding this, it should be noted that the vector of constant parameters does not depend on time, nor on ε, so:

$$\frac{dc}{d\varepsilon} \;=\; 0 \tag{4.47}$$

Recalling the chain rule of differentiation, we can obtain the derivative with respect to ε of the state variable y as:

$$\frac{dy}{d\varepsilon} \;= \frac{dt}{d\varepsilon}\,\frac{dy}{dt} \;= (t_j - t_{j-1})\,\frac{dy}{dt} \tag{4.48}$$

The boundary conditions do not distinguish between assigned and optimal boundary conditions, and are usually expressed as:

$$\Psi(s) \;=\; 0 \tag{4.49}$$

The vector s appearing in the boundary equations contains the value of the y variables at every boundary, both external and internal, as well as the vector of constant parameters:

$$s \;=\; (y_0,\, y_1,\, \ldots\, y_\varepsilon,\, \ldots,\, y_n,\, c) \tag{4.50}$$

In general, the initial values of the variables, $z(t_0)$, are unknown, so the purpose of the solution algorithm is to search for such values, so that they can satisfy the boundary conditions. In the following discussion, the symbol r is used to indicate the generic integration step of the solution algorithm, while q^r indicates the vector of initial values at

the $r-th$ iteration. The operation of the software is based on some key steps. At the first iteration $(r = 1)$, it is necessary to provide appropriate tentative values for the variables of the q1 vector, so that:

$$z(t_0) = q^1 \tag{4.51}$$

So, the equations are integrated over the entire trajectory, taking into account the discontinuities at internal boundaries. In this way, at the end of the process, it will be possible to correct the vector of initial values q^1, in order to try to nullify the error ψ^1 on the boundary conditions q^2. Then, the second iteration step $(r = 2)$ is performed. This process is repeated m times for all integration steps, correcting the initial values vector q^r to q^{r+1} every time, which will become the new initial values vector for the next step. The process terminates when a certain level of precision on the error ψ^m committed in the boundary conditions is satisfied. In the generic r-th iteration, following the integration of the trajectory equations, the error on the boundary conditions, that is ψ^r, is found. The correction of the initial values vector q^r by an amount Δq determines:

$$q^{r+1} \;=\; q^r + \Delta \mathrm{q} \tag{4.52}$$

It can be observed how the variation on the initial values vector Δq leads to a variation of the error on the boundary conditions $\Delta\Psi$, which, considering only first-order terms, is equal to:

$$\Delta\Psi \;=\; \frac{\partial \Psi}{\partial q}\,\Delta \mathrm{q} \tag{4.53}$$

Since the aim is to reduce the errors on the boundary conditions to zero, the goal for each subsequent iteration is to achieve $\Delta\Psi = -\Psi^r$. Based on this observation, the initial values are corrected progressively by an amount equal to:

$$\Delta \mathrm{q} = -\left[\frac{\partial \Psi}{\partial q}\right]^{-1} \Psi^r \tag{4.54}$$

where iteration $r+1$ is started by integrating the differential equations with initial values. The iterations continue until the boundary conditions are satisfied with the desired level of precision.

The matrix of this equation can be computed as the product of two other matrices:

$$\left[\frac{\partial \Psi}{\partial\, q}\right] = \left[\frac{\partial \Psi}{\partial s}\right]\left[\frac{\partial s}{\partial q}\right] \tag{4.55}$$

One can easily obtain the gradient of boundary conditions with respect to the vector s by analytical derivation. However, the derivative of the vector s with respect to the vector of initial values q is equivalent to the matrix that contains the derivative of the vector z calculated at the inner and outer boundaries $\varepsilon = (0, 1, \ldots, n)$:

$$[g(\varepsilon)] = \left[\frac{\partial z}{\partial q}\right] \tag{4.56}$$

By deriving the system with respect to the vector of initial values q, we obtain the following equation:

$$[\dot{g}(\varepsilon)] = \frac{d}{d\varepsilon}\left[\frac{\partial z}{\partial q}\right] = \left[\frac{\partial}{\partial q}\left(\frac{dz}{d\varepsilon}\right)\right] = \left[\frac{\partial f}{\partial q}\right] \tag{4.57}$$

By defining the Jacobian matrix as $\left[\frac{\partial f}{\partial z}\right]$, we can further develop the equation above:

$$[\dot{g}(\varepsilon)] = \left[\frac{\partial f}{\partial z}\right]\left[\frac{\partial z}{\partial q}\right] = \left[\frac{\partial f}{\partial z}\right][g(\varepsilon)] \tag{4.58}$$

The identity matrix can be easily obtained by setting the initial values for the homogeneous system as follow:

$$[g(t_0)] = \left[\frac{\partial z(t_0)}{\partial q}\right] = [I] \tag{4.59}$$

The method used to obtain $[g(\varepsilon)]$ allows for the treatment of cases where there are discontinuities in the variables. Assuming that there is a discontinuity in the $i-th$ boundary, the method is still valid. It is sufficient to update both the vector of variables z and the expression of the matrix g, through a relation l that connects the values of the variables immediately before and after the discontinuity, mathematically given by:

$$z_{i_+} = l(z_{i_-}) \tag{4.60}$$

$$[g_{i_+}] = \left[\frac{\partial l}{\partial z}\right][g_{i_+}] \tag{4.61}$$

This explains why there is no distinction made between y_{i_+} and y_{i_-} in the vector s, as it is possible to relate the two values through the relationship l, as expressed in the above equation. If some of the initial values of the variables are known, the problem is simplified, and the vector q only contains the estimation of unknown initial values of the variables $z(t_0)$, while vector Ψ only contains the boundary conditions that are not explicit at the initial time (hence the term implicit conditions). Since the analytical effort required for the determination of the error gradient matrix $\left[\frac{\partial \Psi}{\partial q}\right]$ is significant and the computational time may be quite long, another alternative method can be adopted. It involves numerically evaluating the error gradient matrix: row i of the matrix is obtained by perturbing the $i-th$ component of q by a small variation Δq and subsequently integrating. By doing so for each component of q, it is possible to evaluate the variation of the errors $\Delta\Psi(\Delta q)$ and the corresponding $i-th$ row as $\Delta\Psi^T/\Delta q$ through a linearization procedure. Empirical values for (Δq) are typically in the order between 10^{-6} and 10^{-7}. Although this numerical method is faster than the analytical one, it may not ensure numerical convergence. Its implementation depends on the complexity and sensitivity of the problem. This numerical procedure can also be used to calculate the Jacobian matrix $\left[\frac{\partial f}{\partial z}\right]$ and the matrix $\left[\frac{\partial \Psi}{\partial s}\right]$. However, the analytical technique is preferred in these cases, and a comparison between analytical and numerical results for both matrices is performed to verify accuracy and precision.

Chapter 5

Model Definition

In Chapter 4 was introduced the theory of optimal control, beginning with a historical background of the problem and then providing a general mathematical formulation, considering the definitions and theorems involved. and giving an overview of the numerical method used to write the guidance algorithm. At this point, it is necessary to define the model examined in this thesis work. Specifically, this chapter presents the mission architecture, showing the main phases that the lander will have to follow. Then, a brief description of the Moon will be provided, starting with the hypotetis adopted in the system's dynamics. The reference frames used to derive the motion equations will be explained, also providing an outline of the boundary conditions adopted for the landing maneuver. Additionally, the main propulsion system's for translational dynamics and the reaction control system for attitude control will be specified in detail, highlighting how the vehicle's orientation influences the path to follow and explaining their limitations. The assumptions made to simplify the model will be delineated step by step. Once the foundation for understanding the adopted model is laid, in the next chapter it will be possible to address fuel-optimal trajectory for powered descent and soft-landing.

5.1 Mission Scenario

To begin with, it is important to provide an overview of the mission architecture that the lander will have to follow in order to achieve a soft-landing on the lunar surface. Referring to the Apollo Lunar Module concept of operations, Figure 5.1.1 presents the major phases of the descent and landing profile adopted for this work. The mission commences with a low lunar orbit (LLO), which is a circular parking orbit at an altitude of approximately 100 km above the lunar surface. During this phase, the trajectory for the descent and landing phase is established based on the landing site's position. Once the exact ignition time is calculated, the descent begins with the descent orbit insertion (DOI), where a de-orbit burn places the lander into a Hohmann-transfer orbit with a lowered periapsis between 10 and 15 km altitude. The Powered Descent Phase then begins. It is divided into three sub-phases: Braking, Approach, and Terminal. Figure 5.1.2 provides a scale drawing of this mission stage, representing the altitude, velocity, and timeframe of the different sub-phases. The main braking phase, like the Apollo Lunar Module, begins approximately 500 km downrange from the landing area with the powered descent initiation (PDI) maneuver. The propulsion system is ignited to provide maximum thrust and effi-

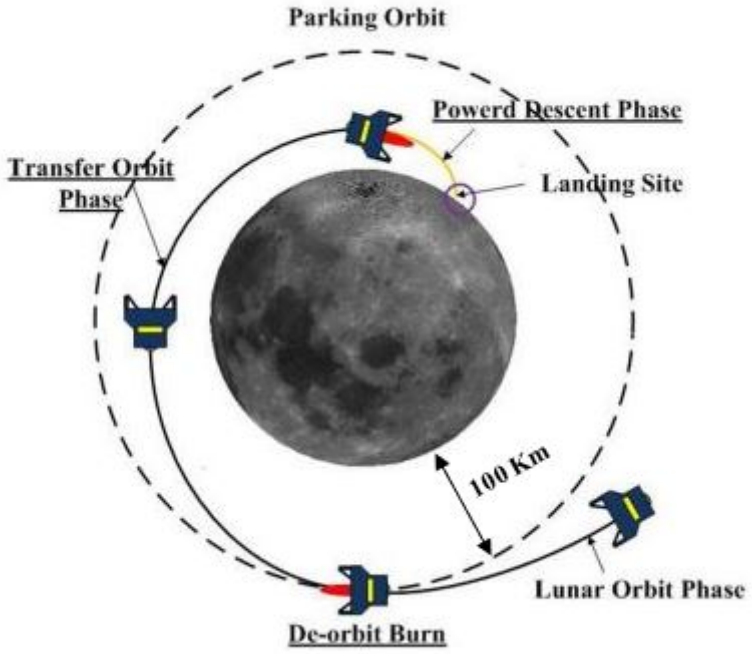


Figure 5.1.1: Phases for Lunar Descent Trajectory

ciently remove as much of the orbital velocity as possible. During this phase, most of the horizontal velocity is eliminated, reducing it from 1630 m/s to around 70 m/s. However, the correct thrust direction must be maintained to avoid excessive vertical velocity and an accidental hard-landing. At the High Gate (HG) point, the pre-selected landing site (LS) enters the Field of View (FoV) of lander sensors. A few seconds later, at Approach Gate (AG), the thrust level in the horizontal direction is reduced to gain manoeuvrability and achieve a trajectory compatible with navigation and Hazard Detection and Avoidance constraints, such as the LS's visibility and time for image processing. This is the Approach Phase, where the vehicle begins actively targeting the landing site and may include retargetings to divert to a safe area if the original one is found to be unsafe. This phase is critical in terms of safety, landing precision, and propellant usage. When visual contact with the landing site is lost, the low gate (LG) is reached. At approximately 15 m above the surface, all horizontal velocity must be nullified, and the approach phase ends at Terminal Gate (TG). The Terminal Phase starts from here, and the lander descends vertically with constant deceleration to the surface until touchdown (TD). Although this thesis focuses on a lunar landing, it is crucial to consider that this mission scenario can also be adjusted for entry trajectories on asteroids or Mars.

5.2 Moon

The Moon, the brightest celestial body in our night sky, plays a crucial role in maintaining Earth's habitable environment by moderating the planet's axial wobble, which results in a relatively stable climate. Additionally, its gravitational pull creates tides, pro-

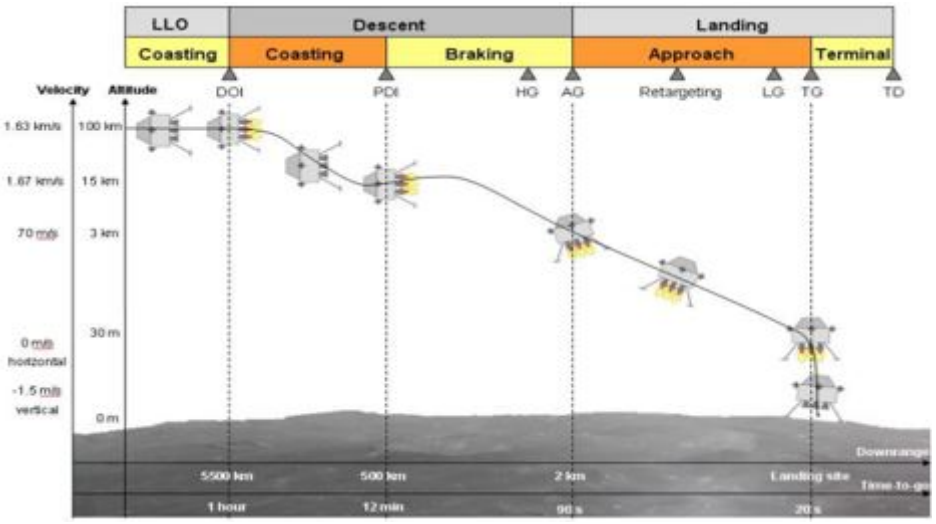


Figure 5.1.2: Lunar Landing Phases

Surface Area $[10^7 Km^2]$	**3.793**
Mass $[10^2 2Kg]$	**7.346**
Volume $[10^1 0Km^3]$	**2.1968**
Mean Radius [Km]	**1737.4**
Mean Density $[Kg/m^3]$	**3344**
Surface Gravity $[m/s^2]$	**1.622**
GM $[km^3/s^2]$	**4900**

Table 5.2.1: Moon Physical Characteristics

viding a rhythmic pattern that has guided humanity and other life forms for millennia. Luna, the Latin name of the Moon, is the root adjective for all lunar entities, derived from the Indo-European term luc, meaning "to shine". By geophysical definition, the Moon is a planetary mass object with a rocky body, qualifying it as a satellite planet. Its diameter is comparable to that of Australia, making it the largest and most massive satellite relative to its parent planet in the Solar System, surpassing also all known dwarf planets in size. Table 5.2.1 provides an overview of the key physical and gravitational characteristics of the Moon.

Isotopic dating of lunar samples indicates that the Moon formed about 4.51 billion years ago, shortly after the formation of the Earth, although its origin remains unknown. Over the years, several theories have been proposed to explain its formation, but none satisfactorily explain the high angular momentum of the Earth-Moon system. The capture of a pre-formed Moon through gravitational forces depends on an implausibly extended Earth's atmosphere to dissipate the energy of the passing Moon. The co-formation of the Earth and the Moon in the primordial accretion disk does not explain the depletion of metals in the Moon. The most widely accepted theory postulates that the Moon formed from debris resulting from a giant impact between Earth and a hypothetical Mars-sized body called Theia. During the first chaotic phase of the Solar System, the impact ejected material into space, which accumulated to form the Moon in an orbit much closer to Earth

Semimajor axis [10^5 Km]	*3.8443*
Perigee [10^6 Km]	*0.3633*
Apogee [10^6 Km]]	*0.4055*
Revolution period [days]	*27.3217*
Synodic period [days]	*29.53*
Mean orbital velocity [Km/s]	*1.022*
Inclination to ecliptic [deg]	*5.145*
Orbit eccentricity	*0.0549*
Sidereal rotation period [hrs]	*655.720*
Obliquity to orbit [deg]	*6.68*
Recession rate [$cm/year$]	*3.8*

Table 5.2.2: Moon Orbital Parameters

than its current location. In fact, due to tidal forces, the Moon exerts a friction force on the Earth, causing its rotation to slow down and acquiring kinetic energy. As a result, the Moon is slowly moving away from the Earth at a rate of about 3.8 centimeters per year. Since its formation, the Moon's orbit has become significantly larger, placing it at approximately 384,000 kilometers from us. At this distance it completes an orbit relative to the fixed stars in a sidereal period of about 27 days, and since it rotates on its own axis at the same speed, it is in synchronous rotation with the Earth. This means that the Moon always faces the Earth with the same hemisphere, known as the near side. However, since the Earth is also in motion, from our point of view the Moon appears to orbit us in a synodic period of 29 days, 12 hours, 44 minutes and 2.8 seconds (lunar month), which is equal to the length of the solar day on it. More information on the Moon's orbit is provided in Table 5.2.2. Since orbital and rotational motions have a much longer duration than powered descent and landing maneuvers, as a first approximation, they will not be needed to derive the lander's equations of motion. Therefore, we define the following simplification for the model:

Assumption 1.

"The orbital and rotational motion of the Moon will be neglected for the entry trajectory."

The post-formation Moon was very different from today. Long ago, between 4.2 and 1.2 billion years ago, the Moon had active volcanoes that produced the dark areas visible from Earth with the naked eye, called maria (Latin for seas). Almost all maria are on the near side and fill the spaces between the bright areas of the Moon, representing the ancient crustal highlands. These light and dark areas represent rocks of different compositions and ages, providing evidence for how the early crust may have crystallized from a magma ocean. The intense terrestrial activity in the early years of the Moon's life was primarily due to the release of large amounts of heat from the core to the mantle. For this reason, the early Moon may have developed an internal dynamo, which is the mechanism for generating global magnetic fields for terrestrial planets. However, today the Moon only has a very weak magnetic field, less than 0.2 nanoteslas. Therefore, the presence of a

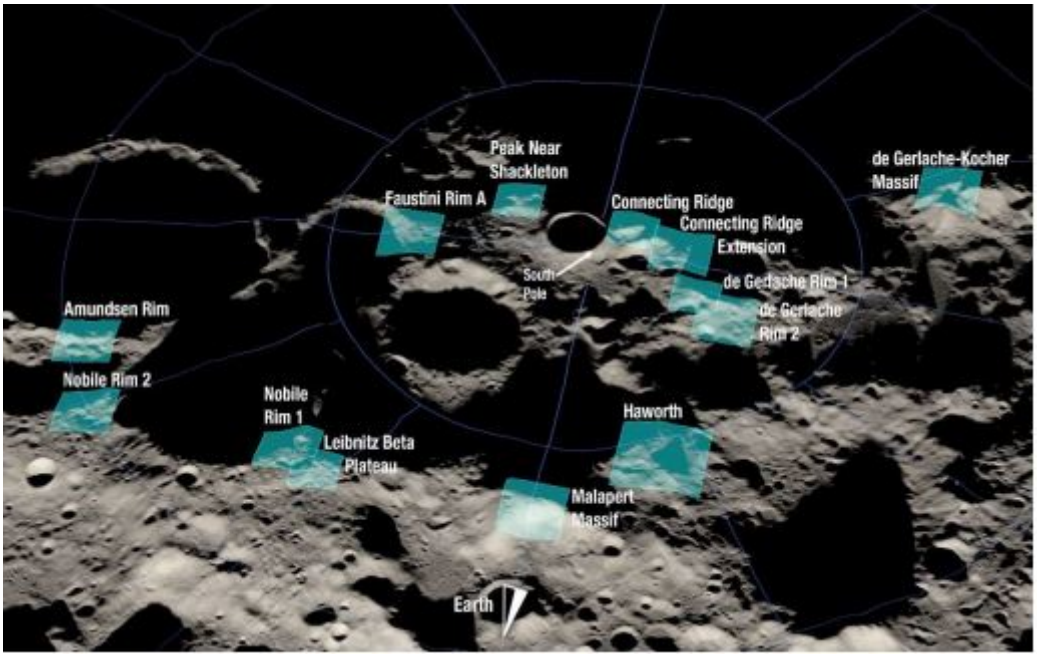


Figure 5.2.3: Candidate landing regions for Artemis III

global magnetic field for the Moon was not considered in this study. Let's define the following hypothesis:

Assumption 2.

"The attitude dynamics are not perturbed by electromagnetic forces."

As the dynamo mechanism stops in the core, the Moon's surface starts to cool, leaving space for the impact of a steady rain of asteroids, meteoroids, and comets. The craters have been preserved for billions of years, providing an impact history for the Moon and other bodies in the inner Solar System. The South Pole–Aitken basin is the largest crater known. It exhibits fascinating scientific features due to its permanently shadowed regions that are abundant in resources and high concentrations of water ice. As a result, NASA has planned the landing of Artemis III in this region. The agency has already identified 13 potential landing sites shown in Figure 5.2.3. Each of these areas measures approximately 15 by 15 kilometers.

The primary reason for the Moon's topography is due to the very thin and weak atmosphere, known as the exosphere. This atmosphere does not provide any protection against meteoroid impacts. Studies of magma samples retrieved from the Moon by the Apollo missions reveal that it once had a relatively thick atmosphere for a period of 70 million years between 3 and 4 billion years ago. This atmosphere, which originated from gases ejected during lunar volcanic eruptions, was twice as thick as the present day atmosphere of Mars. However, the ancient lunar atmosphere was eventually stripped away by solar winds and dissipated into space. Nowadays, the atmospheric pressure at the surface is a few picopascals. As a result, the interaction between particles in the lunar environment and spacecraft is very weak.

Therefore, the following assumption is made for modeling the trajectory:

Assumption 3.

"Aerodynamic forces acting on the lander do not affect its dynamics."

5.3 Reference Frames

To define the system dynamics and hence the state variables for the optimization problem, it is necessary to first specify the Cartesian coordinate systems to which reference should be made. For orbital dynamics, the inertial reference frame $F = (X, Y, Z)$ has been adopted. It is defined as the Moon Inertial (MI) frame, which has its origin at the center of mass (CoM) of the Moon and is assumed to be fixed with respect to the fixed stars, meaning that the axes are non-rotating. The MI frame is based on the Moon's principal equatorial plane, corresponding to the X-Y plane, with the axes oriented parallel to the Earth-Centered Inertial (ECI) frame. Therefore, the following definitions can be derived for the axes:

- The X-axis points in the vernal equinox
- The Z-axis lies at 90 degrees to the equatorial plane and extends through the North Pole
- The Y-axis axis completes the right-handed system The X-axis points in the vernal equinox

According to Assumption 1 of the previous paragraph, the orbital motion of the Moon around the Earth will not be considered in deriving the motion equations. Therefore, assuming that the vehicle is not disturbed by external forces along its orbit, the restricted two-body problem can be used as the mass of the lander is much smaller than that of the Moon. In this way, the MI frame can be used to determine the orbital motion of the lander as well as for estimates of the Hohmann transfer that will lead to the initial phase of the descent maneuver.

On the other hand, the reference system $F_S = (X_S, Y_S, Z_S)$ centered at the coordinates of the landing site is used to study the powered descent phase.Starting from the F frame, the locus of points given by the intersection between the position vector connecting the center of mass of the Moon to that of the lander and the lunar surface, i.e., the ground track, can be plotted. In this way, the range on the surface that the vehicle must traverse in order to reach the landing site is obtained. This distance can be decomposed into a component along the ground track, defined as downrange, while the distance traveled in a direction orthogonal to the ground track is called crossrange. Therefore, first, the equations are derived with respect to the xyz frame. Since the initial and final positions for the descent will be given with respect to locations on the target, the equations of motion are converted to downrange (d), altitude (h), and crossrange (c) coordinates. The representation of the frame is shown in Figure 5.3.4. The transformation assumes a reference sphere of radius r. The position can be described in spherical coordinates, with the equatorial plane of the spherical coordinates equal to the $d-h$ plane. The angle in the plane is denoted as δ, and the angle perpendicular to the plane is γ. Defining crossrange and downrange as the projection on the reference sphere, the angles can be expressed as:

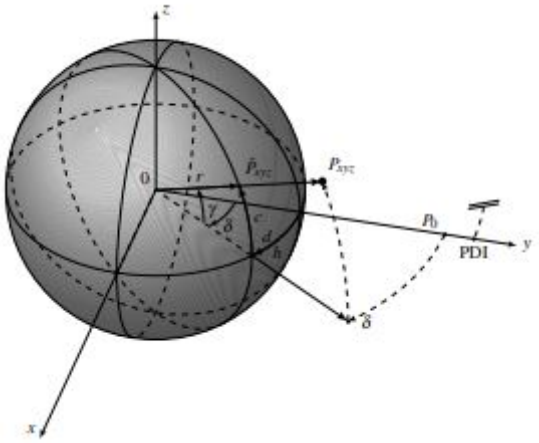


Figure 5.3.4: Downrange, Crossrange, Altitude visualization

$$\delta = \frac{d}{r} \;\; ; \;\; \gamma = \frac{c}{r} \tag{5.1}$$

In accordance with Assumption 1, since the timescale of the soft-landing trajectory is significantly shorter than that of the Moon's rotational motion, the effects of its rotation can be disregarded during the maneuver. Consequently, the F_S system is inertial with respect to the motion of the lander. Furthermore, the following simplification can be adopted:

Assumption 4.

"The surface of the Moon is assumed to be flat."

Due to the fact that the lander's descent trajectory begins only a few tens of kilometers above the surface, the curvature of the surface can be neglected.

This causes the ground-track to degenerate into a straight line. Therefore, the following axis definitions can be given for the F_S reference frame:

- X_S-axis parallel to the downrange direction
- Y_S-axis parallel to the crossrange direction
- Z_S-axis perpendicular to the $c-d$ plane and pointing outward from the lunar surface

At this point, in order to determine the attitude dynamics of the lander, it is necessary to define the body reference frame $F_B = (X_B, Y_B, Z_B)$. The origin of the body reference frame is located at the center of mass of the vehicle. The XB-axis is chosen to be positive in the direction of the top of the lander. The body reference frame of the lunar lander is based on the Apollo Lunar Module, which is illustrated in Figure 5.3.5. As shown in the figure, the Z_B-axis is aligned with the line that passes between two of the LM's legs. The Y_B-axis completes the right-handed system and is aligned with the line that connects the other two legs. The definition of the Y_B-axis and Z_B-axis for this project, will be given in the following. First, let's introduce another hypothesis:

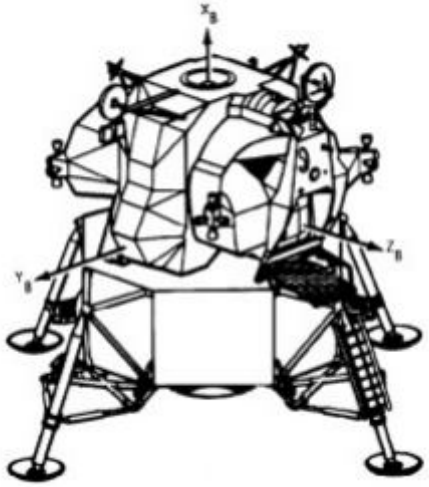


Figure 5.3.5: Body frame of the Apollo Lunar Module

Assumption 5.

"The lander is assumed to be a body with cylindrical symmetry."

Compared to the Lunar Module, in this case, the landing vehicle is assumed to have two planes of symmetry and the X_B-axis coincides with the axis of symmetry of the cylindrical body. This allows neglecting the effects of the roll angle, so any presence of a spin motion of the spacecraft will be neglected in the study of rotational dynamics. Furthermore, the presence of two planes of symmetry allows for the cancellation of centrifugal moments in the inertia tensor, resulting in a diagonal matrix that simplifies the equations of kinematics.

By comparing the body frame F_B with the inertial system F_S, it is possible to determine the orientation of the vehicle and its angular velocity with respect to the landing site. In particular:

- The rotation of F_B with respect to the Z_S-axis is the yaw angle ψ
- The rotation of F_B with respect to the Y_S-axis is the pitch angle θ
- The rotation of F_B with respect to the X_S-axis is the roll angle ϕ

The triad of angles $[\psi, \theta, \phi]$ are defined in this order as Euler angles. The rotations, as written in sequence, lead to aligning the body reference frame with the inertial reference frame F_S. Further assumptions for attitude dynamics are:

Assumption 6.

"The yaw angle e the yaw rate are set to be null."

This ensures that there is no thrust component, and consequently no relative motion, in the crossrange direction. For this reason, the trajectory will be studied on the $X_S - Z_S$

plane. Thus, the initial orbit of the lander must necessarily pass over the landing site, if it is known beforehand. Alternatively, once the site has been identified, a correction maneuver must be evaluated before the Breaking Phase. From this assumption it is also possible to deduce that Y_B-axis remains parallel to Y_S-axis along the entire trajectory. Consequently, Z_B-axis is also determined completing the Cartesian frame F_B.

Assumption 7.

"The main engine is fixed to the lander."

Since the descent engine is fixed to the lander, the thrust is always directed along the positive X_B-axis. Therefore, possible losses due to misalignment are neglected. Moreover, not having any other thrust components, variations in the angle θ are the only way to change the direction of thrust. In this way, the attitude dynamics of the lander are coupled with translational dynamics. Specifically, this will be discussed in the Paragraph 5.5 concerning the equations of motion.

5.4 Controls

As discussed in the previous chapter, another fundamental aspect for defining the dynamics of the system is the control system. Moreover, the only degree of freedom with which the thrust direction can be varied is the pitch angle, as seen in the last paragraph. Therefore, it is necessary to find the right thrust profile at each point of the descent trajectory in order to guarantee a soft-landing and optimize propellant consumption, while respecting the imposed constraints. This section introduces the control systems used by the lander to manage translational and rotational dynamics.

To perform the parking orbit insertion phase, as well as the Hohmann maneuver, and the subsequent Breaking and Vertical Descent phases, a main propulsion system is used. As previously introduced, the main thruster is supposed to be tightly docked to the spacecraft lower part and fixed along the X_B-axis of symmetry. It operates on an on/off basis. In *on-mode*, it is able to provide maximum thrust, while in *off-mode* it does not produce any force on the vehicle, which will be affected only by the lunar gravity force. Therefore, unlike Apollo missions in which the magnitude of the thrust is variable, in this thesis work it is supposed to be constant along the entire descent trajectory. As will be seen later, the thrust level will be gradually increased from a reference value to study its effects, and to investigate possible improvements on the landing maneuver.

On the other side, for attitude control was assumed a Reaction Control System (RCS), taking inspiration from the one used by the Lunar Module (LM) of the Apollo 11 mission. A cluster of LM's RCS is reported in Figure 5.4.6. In that case, thrusters were arranged in clusters of four, mounted on four outriggers equally spaced around the ascent stage. Therefore, a 16-thruster configuration is assumed. In each cluster, two thrusters were mounted facing in opposite directions, parallel to the LM $X-$axis; the other two were spaced 90 degrees apart, in a plane normal to the $X-$axis and parallel to the $Y-$ and $Z-$axis. This study adopt the same configuration, but the engines will have a lower thrust capacity. In fact, they are capable of generating an integer value of torque equal to 133 Newtonmeters. They lie in the plane passing through the lander's center of mass and are symmetrically arranged with respect to the cylinder's axis. Although the presented RCS

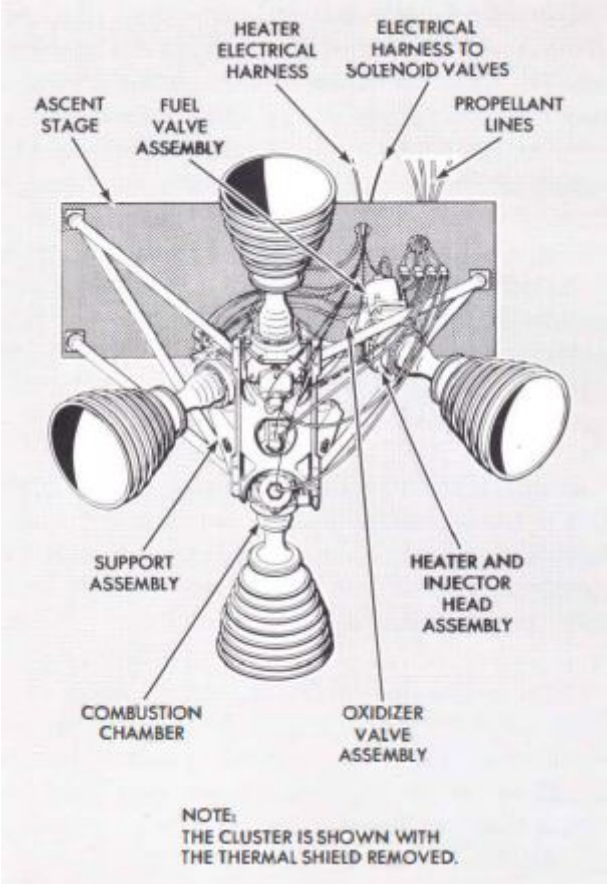


Figure 5.4.6: Cluster for Apollo 11 Lunar Module RCS

system is able to control every rotational degree of freedom of the lander, in accordance with previous assumptions, only the pitch motion will be evaluated.

Assumption 8.

"Perfectly expanded rocket engines are assumed for both main engine and RCS."

With this hypothesis, the vehicle is able to generate a proper thrust and torque. In general, spacecrafts produce a force in a certain direction by expelling propellant through the usage of a chemical rocket engine. Thus, the thrust force (T) is derived from the third law of Newton and can be mathematically described as follows:

$$T = \dot{m}_p V_e + (P_e - P_0)A_e \tag{5.2}$$

Here, $\dot{m}_p$ is the propellant mass flow rate, V_e is the mean exit velocity of propellant particles from the rocket engine at the nozzle exit area A_e. The ambient pressure P_0 and the pressure at the nozzle exit P_e contribute to the thrust force as well. Assuming that the rocket nozzle is perfectly expanded (i.e. $P_e = P_0$), the thrust force and the mass flow rate are related by:

$$T = \dot{m}_p V_e \; = \; \dot{m}_p c \tag{5.3}$$

where c is the effective exhaust velocity of the propellant. Therefore, the thrust force is simply the product of the exhaust velocity and the mass flow rate of the propellant expelled. By introducing the specific impulse (I_{sp}) of the rocket engine, the equation can be rewritten as:

$$T = \dot{m}_p g_0 I_{sp} \tag{5.4}$$

where g_0 is the acceleration due to gravity at the Earth's surface. The specific impulse values for the main engine and the RCS thrusters take inspiration from the systems used in the Lunar Module. They are 290 seconds and 280 seconds, respectively.

5.5 Motion Equations

Taking up the pseudo-inertial reference system F_S introduced in Paragraph 5.3, let's define the unit vectors associated with the axes:

- $\hat{u}_x$ is the unit vector of the X-axis $(downrange)$ of the reference system F_S
- $\hat{u}_y$ is the unit vector of the Y-axis $(crossrange)$ of the reference system F_S
- $\hat{u}_z$ is the unit vector of the Z-axis $(altitude)$ of the reference system F_S

In this way, since the goal is to study the trajectory in this coordinate frame, we can express the lander's position $\bar{R}$, velocity $\bar{V}$, and thrust $\bar{T}$ vectors as follows:

$$\bar{R} = R_x\hat{u}_x + R_y\hat{u}_y + R_z\hat{u}_z \tag{5.5}$$

$$\bar{V} = V_x\hat{u}_x + V_y\hat{u}_y + V_z\hat{u}_z \tag{5.6}$$

$$\bar{T} = T\cos(\theta)\hat{u}_x + T\sin(\theta)\hat{u}_z \tag{5.7}$$

Assumption 6 shows that the Y_B-axis remains parallel to the Y_S-axis throughout the maneuver. For this reason, the expression of thrust shows the lack of a component in the crossrange direction.

On the other hand, according to Assumption 7, the torque $\bar{M}$ acting on the lander and its angular velocity $\bar{\omega}$ can be defined in this way:

$$\bar{\omega} = \omega\hat{u}_y \tag{5.8}$$

$$\bar{M} = M\hat{u}_y \tag{5.9}$$

since only the pitch motion is considered.

$\bar{F}$ and the gravitational force are the only forces that affect the dynamics of the lander. Therefore, referring to the second law of Newton, we obtain:

$$\bar{F} = m\,\bar{g}_m + \bar{T} = m\,\dot{\bar{V}} = m\,\ddot{\bar{R}} \tag{5.10}$$

This is the equation for *translational motion* along the downrange, crossrange, and altitude coordinates in terms of velocity and position. In the above equation, $\bar{g}_m = -g\hat{u}_z$ is the gravitazional acceleration of Moon at the surface. Since we want to simulate the trajectory starting from an altitude of 15 km, we can neglect variations in lunar gravity. Consequently, we can carry out the following simplification of the model:

Assumption 9.

"The gravitational acceleration acting on the lander during the powered descent maneuver is constant."

Similarly, the equations for *rotational dynamics* around the pitch axis can be derived by taking into account the following expression:

$$\bar{M} = m\,\dot{\bar{\omega}} = m\,\ddot{\theta} \tag{5.11}$$

However, the expression for thrust obtained above is valid for a system with constant mass. As seen in Section 5.4, thrust is obtained by expelling a certain mass with a specific velocity. Consequently, it is possible to use a constant mass formulation at the cost of adding an equation for the change in spacecraft mass induced by propellant consumption. In the considered system, the mass lost over time can be expressed as follows:

$$\dot{m} = \frac{dm}{dt} = -\dot{m}_p \tag{5.12}$$

since it is equal in magnitude to the propellant flow rate. Therefore, by substituting the relationship obtained in the previous paragraph for thrust in this equation, we obtain:

$$\dot{m} = \frac{dm}{dt} = -\frac{T}{c} \tag{5.13}$$

The above equation allows for the evaluation of the lander's mass variation induced by propellant consumption. Specifically, considering the configuration described in Section 5.4, propellant consumption is due to both the main engine activation and the torque-generating thrusters of the RCS system. However, if the thrust value and effective exhaust velocity of the two systems are substituted into the equation, it can be seen that the main engine activation is the predominant term in propellant consumption, being three orders of magnitude greater than the consumption generated by the RCS system. As a result, this statement would result in a small error, that can be considered negligible in first approximation.

Assumption 10.

"The propellant consumption due to the RCS system is neglected."

Furthermore, it is easy to see that if the initial spacecraft mass is m_0 and the mass of propellant used to perform a certain maneuver is m_p, then the final mass mf will be equal to the difference between the two, as shown below:

$$m_f = m_0 - m_p \tag{5.14}$$

In this regard, it should be noted that the main objective of this work is to implement a guidance algorithm that is capable of finding the optimal landing trajectory. In Chapter 4, the theory of optimal control was introduced, highlighting that optimal control is the one that, among all the admissible controls, can satisfy a given problem with its boundary conditions and constraints, maximizing a certain system performance. In this case, the variable to be maximized is the final mass of the lander. From the above equation, it is understood why this is equivalent to minimizing the propellant consumption used for the maneuver.

5.6 Boundary Conditions and Constraints

At this stage, it is necessary to impose specific initial and final conditions, as well as physical and engineering constraints, to bound the problem. The boundary conditions will only refer to the external limits of the trajectory: the starting point of the trajectory, i.e., $t = t_0$, coincides with the endpoint of the DOI and at this time, the attitude conditions of the lunar lander are considered unassigned. In this way, the algorithm can search for the optimal orientation of the thrust with which the lander must start the PDI without particular constraints. Let us then define the *initial conditions* as follows:

$$\begin{cases} \bar{R}_0 = \bar{R}(t_0) \\ \bar{V}_0 = \bar{V}(t_0) \\ \bar{m}_0 = \bar{m}(t_0) \end{cases}$$

At the end of the landing trajectory, i.e., for $t = t_f$, it is required that the lander has zero velocity at the landing site and that its symmetry axis is perpendicular to the lunar surface, ensuring a vertical descent. To complete a safe soft-landing and allow for the correct deployment and use of the landing gear, the additional condition of zero angular velocity at the endpoint is imposed. Therefore, the *final conditions* are as follows:

$$\begin{cases} \bar{R}_f = \bar{R}(t_f) = 0Km \\ \bar{V}_f = \bar{V}(t_f) = 0Km/s \\ \theta_f = \theta(t_f) = 90 \\ \omega_f = \omega(t_f) = 0rad/s \end{cases}$$

It may be necessary to introduce certain internal conditions at specific points. For example, it will be necessary to indicate the height at which the vertical descent should begin, at which point the horizontal velocity should already be zero. Other requirements may also concern attitude feedback. Another crucial variable to define is time. In this case, the initial time t_0 is fixed, while the final time t_f is set free. The constraint also involves the capability of the thruster, which includes the thrust magnitude of both the main engine and RCS thrusters. Therefore, let's assume these other limitations.

Assumption 11.

"Thrust and Torque magnitudes are bounded."

The maximum and minimum rates at which propellant mass can be expelled from the chemical rocket engine correspond to the upper and lower limits of thrust, respectively. Therefore, limits on propellant mass rate are directly converted into limits on thrust and torque magnitude, represented by:

$$T_{min} \leq T \leq T_{max} \tag{5.15}$$

$$M_{min} \leq M \leq M_{max} \tag{5.16}$$

At this point, having fully defined the model for the landing mission, all the necessary ingredients are available to apply to the optimization problem. In the next chapter, we will see how these conditions will be applied to the optimal control problem. Additionally, the trajectory structure will be presented, explaining why it will be necessary to divide the problem into a series of sub-intervals of integration.

Chapter 6

Optimal Control for Lunar Landing

Chapter 5 introduces the dynamic model of the lander for the powered descent maneuver to be optimized. The choice of a fuel-minimizing strategy is driven by future prospects, which will see a higher frequency of missions to the lunar surface. This is a fundamental aspect of the lunar landing design, as it has important implications for the vehicle's mass budget. Therefore, this chapter presents the formulation of the Boundary Value Problem that allows us to maximize the mass of the lander at the end of the trajectory, obtained by applying the Pontryagin's Maximum Principle to the optimal control problem, as seen in Chapter 4. The dynamics describing the system, the assumptions made, and the constraints outlined for the mission will be taken into account. Then, we will proceed with the presentation of the bang-bang control and the constraints imposed on the maneuver to ensure its efficiency and safety. To this end, the subdivision into subintervals introduced in the trajectory will be explained in detail. Finally, the algorithm implemented to solve this optimization problem will be discussed.

6.1 Optimal Lunar Landing Problem

Let's recall the definition of the optimal control problem provided in Chapter 4:

> *"Find an admissible control u* which causes the system*
>
> $$\dot{x}(t) = a(x(t), u(t), t) \tag{6.1}$$
>
> *to follow an admissible trajectory x^* that minimizes or maximizes the performance measure*
>
> $$J = h(x^*(t_f), t_f) + \int_{t_0}^{t_f} g(x^*(t), u^*(t), t)dt \leq h(x(t_f), t_f) + \int_{t_0}^{t_f} g(x(t), u(t), t)dt\text{"} \tag{6.2}$$

Considering the control variables of the lander, its dynamics in the F_S frame can be described by the following set of equations:

$$\begin{cases} \dot{R}_x = \dot{V}_x \\ \dot{R}_y = \dot{V}_y \\ \dot{R}_z = \dot{V}_z \\ \dot{V}_x = T/m * cos(\theta) \\ \dot{V}_y = 0 \\ \dot{V}_z = T/m * sin(\theta) - g \\ \dot{\theta} = \omega_y \\ \dot{\omega}_y = M/I_y \\ \dot{m} = -T/c = -T/(g_0 I_{sp} \end{cases}$$

On the other hand, the performance to be maximized is the final mass of the vehicle. To achieve this, the form chosen to define J corresponds to the so-called *Mayer form*

$$J = h(x(t_f\ ), u(t_f\ ), t_f) \tag{6.3}$$

which represents the *terminal cost*. This form allows for the optimization of the final condition performance measure J, that can be seen as the profit to be maximized by the desired set of optimal controls.

The next step involves the calculation of the Lagrange multipliers λ, specifically defined as costate (or adjoint) variables. They are represented by the following vector:

$$\lambda \ = \ (\lambda_{R_x}), \lambda_{R_y}), \lambda_{R_z}), \lambda_{V_x}, \lambda_{V_y}, \lambda_{V_z}), \lambda_\theta, \lambda_\omega, \lambda_m) \tag{6.4}$$

which are associated with the state variables of the problem. It is important to note that these variables are introduced to define the Hamiltonian and simplify the calculation of the gradient of the cost function with respect to the controls, which is necessary to update the control variables iteratively until convergence to the optimal solution is achieved. Consequently, we proceed with the definition of the Hamiltonian as follows:

$$\begin{aligned} H = \lambda_{R_x} V_x + \lambda_{R_y} V_y + \lambda_{R_z} V_z + \lambda_{V_x}(T)/m * cos(\vartheta) \\ + \lambda_{V_z} T/m * sin(\vartheta) + \lambda_\theta\ \omega + \lambda_\omega M I_y - \lambda_m T/c \end{aligned} \tag{6.5}$$

By differentiating the Hamiltonian function with respect to the state variables, the costate equations can be obtained as follows:

$$\frac{d\lambda}{dt} = -\frac{\delta H}{\delta x} \tag{6.6}$$

From these equations, it is possible to define one of the three necessary conditions for optimality of the solution. In particular, for the problem under consideration, we have:

$$\begin{cases} \dot{\lambda_x} = 0 \\ \dot{\lambda_y} = 0 \\ \dot{\lambda_z} = 0 \\ \dot{\lambda_{V_x}} = -\lambda_{R_x} \\ \dot{\lambda_{V_y}} = -\lambda_{R_y} \\ \dot{\lambda_{V_z}} = -\lambda_{R_z} \\ \dot{\lambda_\theta} = \frac{T}{m}\left(\lambda_{V_x}\sin(\theta) - \lambda_{V_z}\cos(\theta)\right) \\ \dot{\lambda_{\omega_y}} = -\lambda_\theta \\ \dot{\lambda_m} = \frac{T}{m^2}\left(\lambda_{V_x}\cos(\theta) + \lambda_{V_z}\sin(\theta)\right) \end{cases}$$

Firstly, it should be noted that the adjoint differential equations are independent of the cost functional J. Furthermore, combining the state variable vector with the adjoint variable vector, we obtain a total number of NY=18 variables for the lunar landing problem. Therefore, boundary conditions must be added for the adjoint equations. The lander's downrange and initial attitude are considered free variables of the problem, while its mass will be an imposed parameter. Hence, the constraints for the costate variables are:

$$\begin{cases} \lambda_{R_{x_0}} = \lambda_{R_x}(t_0) = 0 \\ \lambda_{\theta_0} = \lambda_\theta(t_0) = 0 \\ \lambda_{\omega_0} = \lambda_\omega(t_0) = 0 \\ \lambda_{m_0} = \lambda_m(t_0) = 1 \end{cases}$$

The free attitude conditions of the lander at the initial point of the trajectory determine a *singular arc* condition, since the Bang-Bang control for torque will not be defined. Next, Section 6.4 will provide a detailed explanation of the singular arc condition.

6.2 Primer Vector

But before proceeding, it is necessary to reorganize the equations by introducing the primer vector **p**(t) defined by D.F. Lawden in 1963 on *"Optimal Trajectories for Space* Navigation". In Optimal Control Theory, Pontryagin's Maximum Principle states that the performance measure is optimized if and only if the optimal control value maximizes (or minimizes) the Hamiltonian. However, in addition to the thrust module (fixed), the Hamiltonian is also influenced by its direction, i.e., the pitch angle, considering a main engine with a fixed mounting on the spacecraft. According to the definition given by D.F. Lawden, the *optimal thrust vector* is in the direction of the *primer vector*. In the case of the landing problem, neglecting the effects of crossrange and yaw, the primer vector is parallel to the adjoint variables vector λ_V.

Thus, consider the unit vector $\hat{u}$ that indicates the direction of the thrust in the inertial reference frame centered at the landing site F_S. In the lunar landing problem with the assumption of bidimensional motion, its components are:

$$\hat{u} = [cos(\theta)\hat{u}_x, sin(\theta)\hat{u}_z] \tag{6.7}$$

In order to maximize the Hamiltonian on the choice of thrust direction, primer vector has to be parallel to the adjoint variable. This is definable as:

$$\mathbf{p}(t) = \lambda_V(t) = [\lambda_{Vx}(t), \lambda_{Vz}(t)] \tag{6.8}$$

Therefore, if the optimal thrust direction coincides with the components of the primer vector, these are:

$$\begin{cases} p_x(t) = \lambda_{V_x}(t) = cos(\theta)\hat{u}_x \\ p_z(t) = \lambda_{V_z}(t) = sin(\theta)\hat{u}_z \end{cases}$$

From this, the optimal thrust direction θ_{opt} can be derived for each instant in time:

$$\theta_{opt} = tan^{-1}[\frac{\lambda_{V_z}(t)}{\lambda_{V_x}(t)}] \tag{6.9}$$

At this point, taking the equation for $\frac{\partial H}{\partial \theta}$ mentioned in Chapter 4 and differentiating it, we obtain:

$$\frac{\partial^2 H}{\partial \theta^2} = -\frac{T}{m}(\lambda_{V_x} cos(\theta) + \lambda_{V_z} sin(\theta)) \tag{6.10}$$

as expressed earlier, the term in parentheses becomes unity and thus disappears from the relation, yielding the necessary and sufficient condition for control optimality:

$$\frac{\partial^2 H}{\partial \theta^2} = -\frac{T}{m} < 0 \tag{6.11}$$

From this, the maximum point of the Hamiltonian is derived, and it guarantees that the found solution is the optimal one.

6.3 Bang-Bang Control

This Paragraph will address the applied control framework for the problem. Applying the Lagrange multipliers, a new form of functional J can be obtained, which is a rewrite of the functional J that takes into account the boundary conditions of the problem through the constants and the evolution laws through the adjoint variables. Imposing the first derivative of this equal to zero allows the maximization of the functional J. This condition is satisfied if and only if the Euler-Lagrange differential equations are satisfied. The first necessary condition for optimality was presented in the Section 6.1. Now, in order to satisfy the stationarity of this new functional, the algebraic equations for the controls must also be defined, namely:

$$\frac{\delta H}{\delta u} = 0 \tag{6.12}$$

where u is the control vector, defined as:

$$u = (T,\ M) \tag{6.13}$$

consisting of the thrust T generated by the main propeller and the torque M generated by the thrusters of the Reaction Control System. Differentiating the expression gives:

$$\frac{\delta H}{\delta T} = \frac{\lambda_{V_x} cos(\theta) + \lambda_{V_z} sin(\theta)}{m} - \frac{\lambda_m}{c} \tag{6.14}$$

$$\frac{\delta H}{\delta M} = \frac{\lambda_\omega}{I_y} \tag{6.15}$$

The coefficient of the control variable that appears in the Hamiltonian expression is defined as the *Switching Function* and can be expressed as:

$$S_T = \frac{\lambda_{V_x}\cos(\theta) + \lambda_{V_z}\sin(\theta)}{m} - \frac{lambda_m}{c} \tag{6.16}$$

$$S_M = \frac{\lambda_\omega}{I_y} \tag{6.17}$$

In case the Hamiltonian is linear with respect to the control variables, it is evident that the control cannot be determined because it disappears from the resulting equation after the differentiation of the necessary condition. However, if the Switching Function associated with the control variables is not null, then the controls fall under the category of what are defined as *Bang-Bang controls*. They can be defined as follow:

$$\begin{cases} T = T_{max} \;\; if \; S_T > 0 \\ T = T_{min} \;\; if \; S_T < 0 \end{cases}$$

where T_{min} in this work is equal to zero, and

$$\begin{cases} M = Mmax = +M \;\; if \; S_M > 0 \\ M = Mmin = -M \;\; if \; S_M < 0 \end{cases}$$

Based on PMP, it can be claimed that in order to maximize the Hamiltonian, the Bang-Bang control should be used. The concept is based on the ignition and shutdown times of the engines, i.e. the minimum amount of thrust required to complete the trajectory, and therefore on the amount of mass that it is advantageous to consume or save along the trajectory while respecting the necessary conditions of optimality.

Algebraic control equations have been computed for the 2-D case for convenience, but formulas maintain the same expression and meaning even for the 3-D case. It should be noted that in a 3-D framework, in addition to control variables and θ, the other control variables are ψ and ϕ, since crossrange is now an active direction of motion. In this study, the complete analytical solution for a six-degree of freedom system has been omitted, due to its high complexity. In this way, except for the constant value of the moment of inertia evaluated with respect to the spacecraft's axis of symmetry I_y, the sign of the Switching Function for torque depends only on the value of the adjoint variable λ_ω.

6.4 Singular Arc

The formulation of Bang-Bang Control is valid only if the Switching Function associated with the control is different from zero. Otherwise, a singular arc turns out. In this case, the search for the optimal control occurs by imposing all the successive derivatives of the Switching Function (made with respect to time) equal to zero, until the control appears explicitly. Therefore, the optimal control value is obtained by imposing the k-th

derivative equal to zero. If the adjoint variable $\lambda_\omega = 0$, as at the beginning of trajectory, then S_M is null. This means that a singular arc condition for torque exists. For the *first-order derivative* we have:

$$\dot{S}_M = \frac{\dot{\lambda}_\omega}{I_y} = -\lambda_\theta = 0 \tag{6.18}$$

Since the control is not explicit yet, we proceed with the *second-order derivative*:

$$\ddot{S}_M = -\dot{\lambda}_\theta = \frac{T}{m}(\lambda_{V_x} sin(\theta) - \lambda_{V_z} cos(\theta)) = 0 \tag{6.19}$$

At this point, the parameter Δ is defined as:

$$\Delta = \sqrt{\lambda_{V_x}^2 + \lambda_{V_z}^2} \tag{6.20}$$

As a result, the terms $cos(\theta)$ and $sin(\theta)$ can be expressed in terms of Δ and in a way that satisfies the relationship for $\ddot{S}_M$ as follow:

$$cos(\theta) = \lambda_{V_x}/\Delta \tag{6.21}$$

$$sin(\theta) = \lambda_{V_z}/\Delta \tag{6.22}$$

Using the product rule and the chain rule of differentiation, we can express the *third-order derivative* of the Switching Function:

$$\dddot{S}_M = \frac{T}{m}(\lambda_{R_x} \sin(\theta) - \lambda_{R_z} \cos(\theta)) - \frac{T}{m}(\lambda_{V_x} \cos(\theta) + \lambda_{V_z} \sin(\theta))\omega = 0 \tag{6.23}$$

At this point, from the following relationship and considering the relationships derived earlier for $cos(\theta)$ e $sin(\theta)$, we obtain the expression for angular velocity:

$$\omega = \frac{\lambda_{R_x} sin(\theta) - \lambda_{R_z} cos(\theta)}{\Delta} \tag{6.24}$$

Making use of this relationship and the adjoint equations derived in Paragraph 6.1, we obtain the *fourth-order derivative*e of the Switching Function for torque:

$$S_M^{IV} = \frac{T}{m}[2\omega(\lambda_{R_x} cos(\theta) + \lambda_{R_z} sin(\theta)) - \frac{M}{I_y}(\lambda_{V_x} cos(\theta) + \lambda_{V_z} sin(\theta))] = 0 \tag{6.25}$$

Finally, we obtain the value of the optimal control **M*** by solving it:

$$M^* = 2\omega I_y \frac{\lambda_{R_x} cos(\theta) + \lambda_{R_z} sin(\theta)}{\Lambda} \tag{6.26}$$

The presence of a singular arc for the torque determines new boundary conditions to be satisfied, coming by $\dot{S}_M = \ddot{S}_M = \dddot{S}_M = 0$. However, the condition $\dot{S}_M = 0$ is equivalent to the necessary optimal condition. For this reason, the additional boundary conditions for the singular arc are:

$$\lambda_{V_{x_0}} sin(\theta_0) - \lambda_{V_{z_0}} cos(\theta_0) = 0 \tag{6.27}$$

$$\omega_0 = \frac{\lambda_{R_{x_0}} sin(\theta_0) - \lambda_{R_{z_0}} cos(\theta_0)}{\Delta} \tag{6.28}$$

Moreover, the Hamiltonian does not explicitly depend on time, hence the final time does not appear neither in the expression of the Hamiltonian function nor in the performance measure. This leads to an additional boundary condition that must be satisfied in the end point of the singular arc (t_S):

$$H(t_S) = H_S = 0 \tag{6.29}$$

6.5 BVP Parameters

When a singular interval occurs, it persists throughout the considered interval. In order to follow the mission steps described in Section 5.1 and perform a soft-landing, the trajectory has been divided into the following sub-intervals:

- *Singular Arc* with duration t_S
- *Second Arc* with duration t_2
- *Third Arc* with duration t_3
- *Vertical Descent* with duration t_4

At the beginning of the powered-descent trajectory, the lander's attitude is left free. As explained in the previous section, this leads to a singular arc. This results in the RCS's thrusters being turned off throughout the arc, and the lander being subject only to inertial moment. At the end of the singular arc, the vehicle's attitude dynamics are still free, resulting in $\theta(t_S) \neq 0$. This can be considered as the value of a state variable in an internal boundary, corresponding to an adjoint variable $\lambda_\theta(t_S)$. In order to satisfy the final boundary conditions imposed on the maneuver, which require a vertical landing ($\theta = 90$) with zero angular velocity ($\omega = 0$ rad/s), a discontinuity in the adjoint variable λ_θ can be exploited at the end of the singular arc, so that $\lambda_\theta(t_{S-}) \neq \lambda_\theta(t_{S+})$. Therefore, unlike before, the value of $\lambda_\theta(t_{S+})$ for the variable representing the rotational dynamics of the lander is imposed, obtaining it from the optimization process. This is an additional parameter for the Boundary Value Problem. Consequently, a second arc is introduced with the imposition of a switch in the torque so that $M = M_{min}$. The negative torque is necessary in order to reduce the pitch angle before the vertical descent phase, as it was high in order to break the horizontal velocity. On the other hand, the third arc is introduced to stop the induced rotational motion in the second interval. Thus, by applying a positive torque, $M = M_{max}$, for an unknown duration t_3, the lander's angular velocity is stopped and the pitch angle converges to 90°. Despite the explicit imposition in the two arcs, the rotation maneuver is optimized while respecting the formulation of Bang-Bang control for the torque. At the end of this interval, the desired conditions for pitch angle and angular velocity must be satisfied in order to start the vertical descent. Furthermore, the horizontal velocity at the end of t3 must be zero ($Vz(t_3) = 0$). At this point, in order to obtain a well-posed problem, an additional boundary condition must be introduced, that is the altitude at which to start the fourth phase, $Rz(t_3)$. Therefore, the vertical descent phase does not involve applying torques to the lander, but only thrust in the vertical direction in order to achieve a soft-landing.

It is noticed that, by imposing the torque in these final phases, the optimal control law for thrust cannot be respected. In fact, the switching function S_T, starting from the second arc, begins to decrease until it becomes negative, which would correspond to the shutdown of the propulsion system according to the Bang-Bang control. However, by doing so, the algorithm calculates a trajectory that passes through negative distances, i.e., it would arrive at the landing site from below. This obviously is not reasonable and in reality would lead to a hard-landing or the crash of the lander on the lunar surface. Therefore, since these phases have a short duration, $T = T_{max}$ is imposed, even though the PMP is not respected and the optimal solution is not obtained.

Explicit B.C	Implicit B.C	Unknown Initial Values
R_{y_0}, R_{z_0}	$R_{x_f}, R_{y_f}, R_{z_f}$	t_S, t_2, t_3, t_4
$V_{x_0}, V_{y_0}, V_{z_0}$	$V_{x_f}, V_{y_f}, V_{z_f}$	$\lambda_\theta(t_{S+})$
m_0, λ_{m_0}	θ_f, ω_f	$R_{x_0}, \theta_0, \omega_0$
λ_{Rx_0}	$\ddot{S}_M$, $\dddot{S}_M, H_S$	$\lambda_{Ry_0}, \lambda_{Rz_0}$
$\lambda_{\theta_0}, \lambda_{\omega_0}$	$R_z(t_3), V_z(t_3)$	$\lambda_{Vx_0}, \lambda_{Vy_0}, \lambda_{Vz_0}$

Table 6.5.1: Boundary Conditions and Unknown Initial Values

In this way, the Boundary Value Problem (BVP) for the considered study consists of five parameters, KP, which are given by the duration of the arcs considered and the free discontinuity introduced in the previous paragraph, and eighteen NY variables, as mentioned in Section 6.1. Therefore, the total size of the BVP is equal to $N = KP + NY = 23$. To ensure that the problem is well-posed and to guarantee the convergence of the code, the same number of Boundary Conditions (B.C) are needed. Once the B.C. and differential equations are defined, one of the major challenges in solving a BVP is finding the most suitable initial tentative solution to start the integration, due to the high sensitivity of the shooting methods. The unknown initial values refer to those for which a tentative guess is required to run the code. Since ten variables have been introduced in the explicit conditions, the remaining eight values, the duration of the four phases and the free discontinuity introduced at the end of the singular arc, constitute the thirteen unknown initial values of the BVP. These together with the explicit-initial and implicit-final ones are presented in Table 6.5.1.

6.6 Algorithm Overview

The resolution algorithm implemented for this thesis work in order to solve the Boundary Value Problem is based on a Fortran code derived from Gasparo, Macconi, and Pasquali of the University of Florence. The BVP reduces to a sequence of initial value problems, as extensively discussed in Section 4.5. The integrator for the system's differential equations is based on a variable step and order integrator derived from Adams method. The modifications to the code introduced by Lorenzo Casalino, professor at Politecnico di Torino, and Guido Colasurdo, professor at Sapienza di Roma, consider the presence of several arches, discontinuities in internal boundaries, as well as mixed boundary conditions. This algorithm can also be adapted for other applications in aerospace trajectories, such as orbital or interplanetary transfers. It provides a precision on the boundary conditions, i.e., the maximum error, of:

$$E_{max} = (max)_i(\Psi_i) < 10^{-7} \tag{6.30}$$

As discussed in Section 4.5, the initial trial solution is corrected by a linearized corrective factor Δq, for which the approximation on the initial values can compromise the convergence of the algorithm, increasing the error on the boundary conditions instead of decreasing it with iterations. To improve the convergence a variable correction factor k_1 has been introduced, so as not to stray too far from the solution of the previous step. Thus:

$$q^{r+1} = q^r + k_1 \Delta q \tag{6.31}$$

Reference Radius	R_{ref}	15 Km
Reference Velocity	V_{ref}	1 Km/s
Reference Acceleration	$a_{ref} = V_{ref}^2/R_{ref}$	0.0667 Km/s^2
Reference Time	$t_{ref} = R_{ref}/V_{ref}$	15 s
Reference Mass	m_{ref}	1600 Kg
Reference Thrust	$T_{ref} = m_{ref}\ a_{ref}$	106.67 kN
Reference Torque	$M_{ref} = T_{ref}\ R_{ref}$	1600 kN/Km
Reference Inertia Moment	$I_{ref} = m_{ref}\ R_{ref}^2$	3.6*105 Kg/Km^2

Table 6.6.2: Dimensionless Factors

The value of k_1 ranges from 0.1 to 1, depending on the proximity of the initial trial solution to the desired one. Once the new vector of estimated initial values q^{r+1} is determined, the equations of motion are integrated, and the maximum error made at the new iteration step E_{max}^{r+1} is calculated. This is then compared with the error made in the previous step:

$$E_{max}^{r+1} \ < \ k_2 E^r{}_{max} \tag{6.32}$$

A factor k_2 between 2 and 3 ensures good converge to the desired solution, considering that the error made on the boundary conditions may increase in the early iterations. If this inequality is satisfied, it is evaluated whether the precision requirement on the boundary conditions is met. If it is not, it means that the error at step r is increasing too much compared to that at step r, so a bisection method is applied to the correction factor Δq determined at step r:

$$q^{r+1} \ = \ q^r \ + \ k_1 \frac{\Delta q}{2} \tag{6.33}$$

Therefore, the iteration process continues. The maximum number of bisections is set to 5, after which the process stops, meaning that convergence cannot be achieved for the selected initial tentative values.

Moreover,in order to enhance the robustness and efficiency of the equation integration, the mathematical problem has been normalized by introducing dimensionless factors. Consequently, the results provided by the developed Fortran code must be rescaled using reference quantities before generating output files. Therefore, Table 6.6.2 presents the scaling factors employed to appropriately configure variables for the integration process.

Chapter 7

Results

Once the model and the algorithm adopted for optimization have been defined, the downrange analysis was carried out for the lunar soft-landing trajectory. Starting from a reference case, the primary objective was to look for better solutions in terms of propellant saving. To this end, the effect of increasing the aviable thrust was evaluated in the first part. Additionally, it will be shown that further improvements can be obtained with an initial shutdown of the propulsion system. On the other hand, in order to ensure a good safety margin for the mission, a real-time change of the landing site was evaluated. Therefore, the algorithm must make a correction to the trajectory, lengthening or shortening it. The second part will present the results for this scenario.

7.1 Reference Case

To begin with, let's assume that the trajectory is unrestricted. The objective is to accomplish a precise and efficient touchdown on a specified point, in case the landing site has been identified Therefore, it is preferable to fix the final position R_{x_f}. Specifically, it must be null since the reference frame is centered on the landing site. On the other hand, the algorithm has to search for the value of the starting downrange distance R_{x_0} that optimizes the amount of propellant used for the maneuver. This way, the co-state variable $\lambda_{R_{x_0}}$ becomes an explicit condition of the Boundary Value Problem and set to zero so that R_{x_0} becomes a free parameter for which an initial guess solution needs to be hypothesized. Moreover, it is necessary to calculate the error on the new implicit boundary condition $R_{x_f} = 0$ to ensure a successful convergence of the code to the unknown initial position. The mission design is directly inspired by the Apollo Lunar Module in terms of altitude, mass, and thrust. The LM had a weight of approximately 15,000 kilograms and was capable of providing 45,000 Newtons of thrust, initiating the Powered Descent from an altitude of 15 kilometers above the lunar surface. In this case it has been considered to reduce the weight and the aviable thrust by a factor of approximately 10, assuming that there is no crew transport on board or subsequent ascent phase. Thus, while maintaining the same starting altitude as that of LM's maneuver, the model adopts a weight of 1,600 kilograms at the beginning of the PDI and a thrust output of 4,000 Newtons, which will be the reference case for subsequent analyses of the landing trajectories. Below the results and trends of the main variables are presented for analyzing the optimal soft-landing trajectory.

Thrust *[Newton]*	**Landing Site** *[Km]*	**Altitude PDI** *[Km]*	**Downrange Velocity** *[Km/s]*	**Mass** *[Kg]*
4000	0.0	15.0	1.6920	1600

Table 7.1.1: Reference Case Data

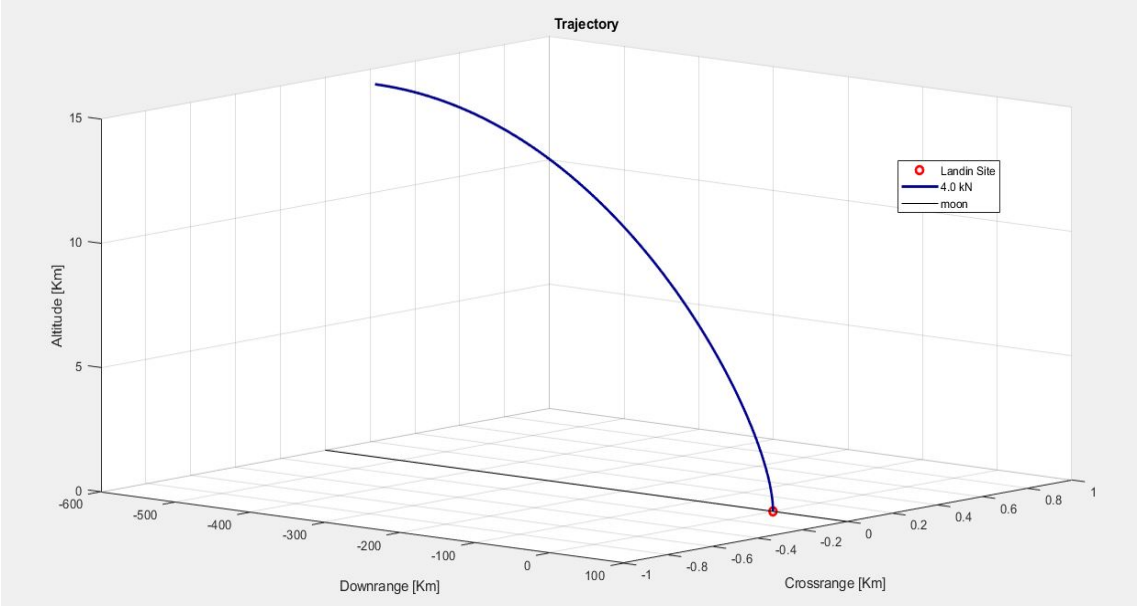


Figure 7.1.1: Trajectory for reference case

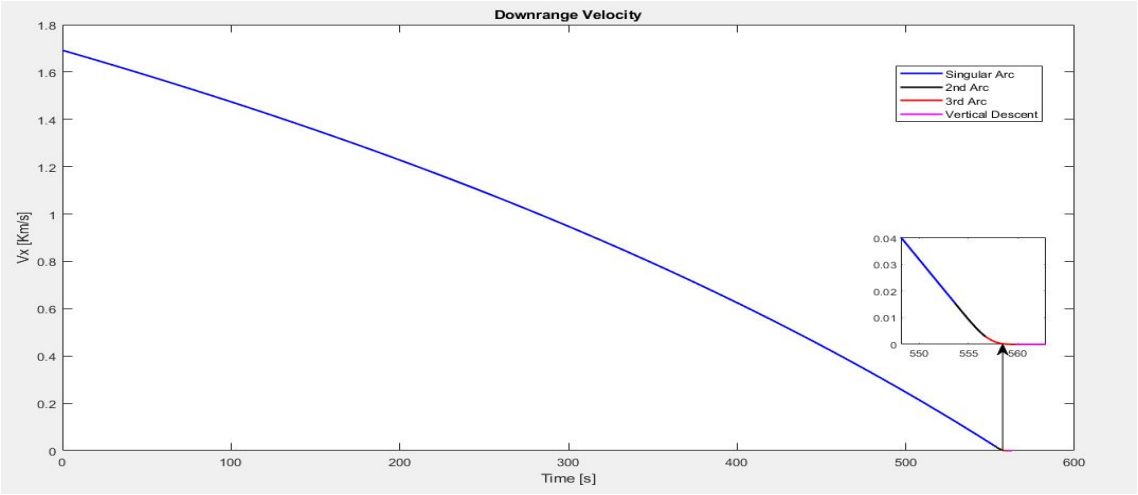


Figure 7.1.2: Dowrange Velocity in the reference case

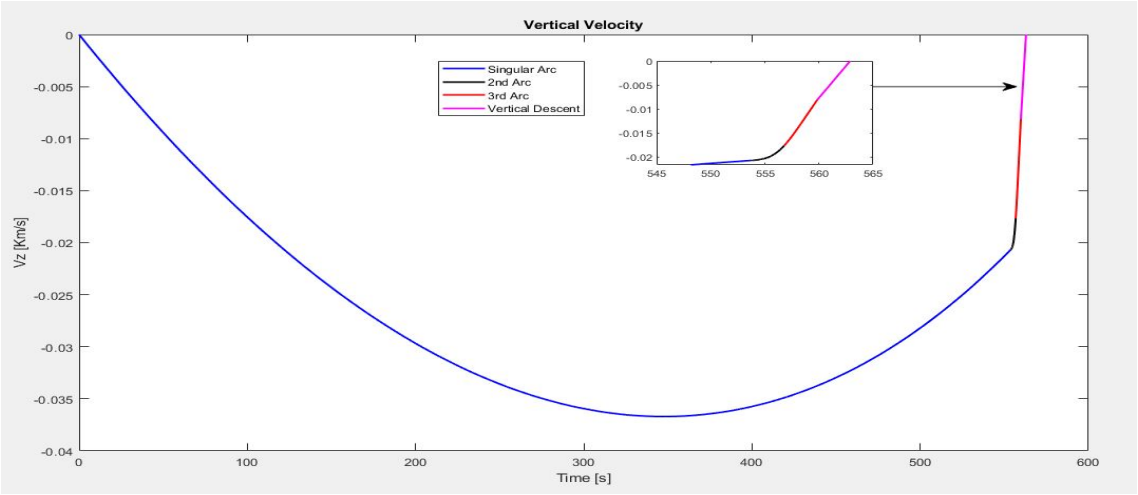


Figure 7.1.3: Vertical Velocity in the reference case

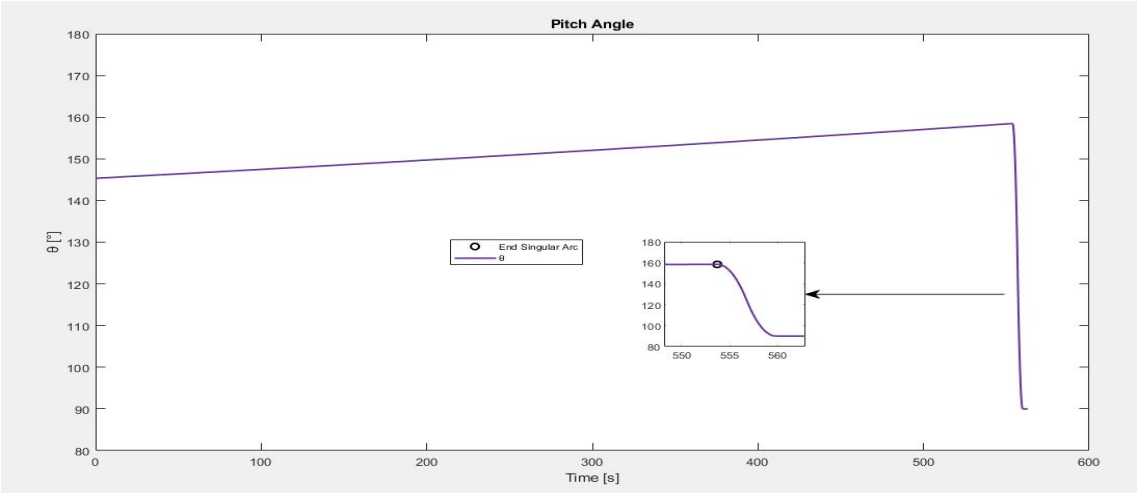


Figure 7.1.4: Pitch Angle in the reference case

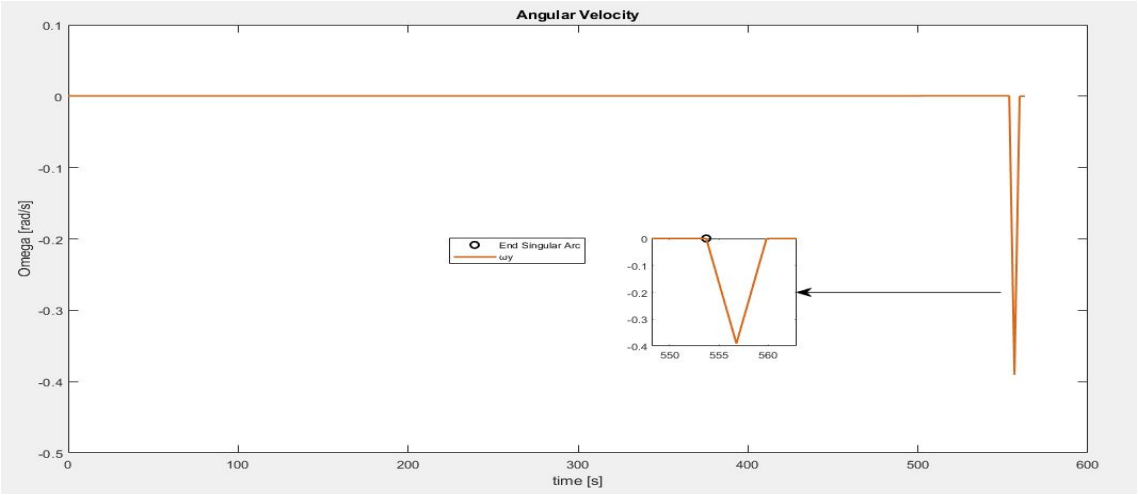


Figure 7.1.5: Angular Velocity in the reference case

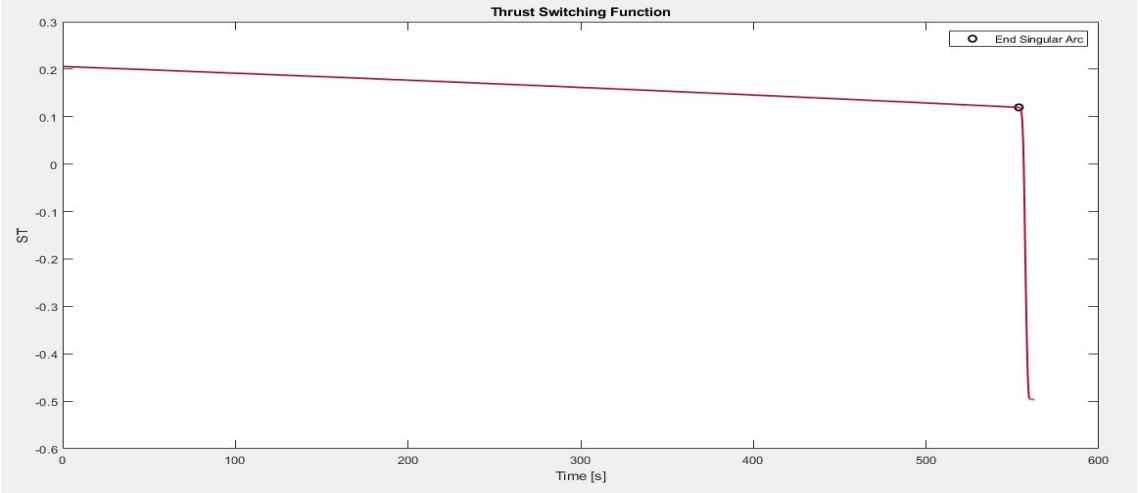


Figure 7.1.6: Thrust Switching Function in the reference case

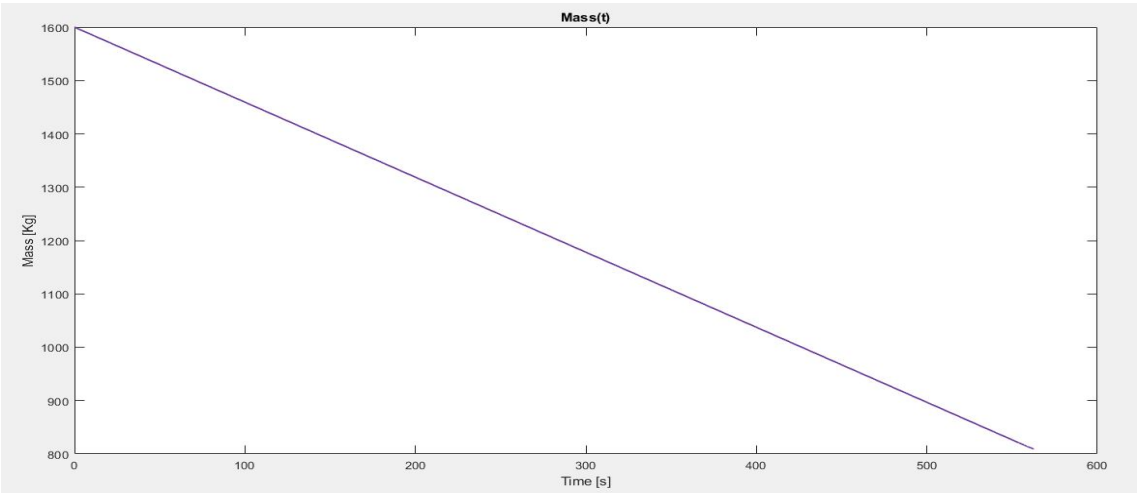


Figure 7.1.7: Mass in function of time for the reference case

7.1.1 Simulation Results

As previously mentioned, the trajectory evolves on the frontal plane ($X - Z$) of the presented reference frame. The crossrange motion is negated due to the absence of a third degree of freedom in the motion equations and the assumption of no yaw effects, as the thrust is oriented along the pitch angle. From this analysis, it can be deduced that the time evolution of velocities will vary depending on how θ evolves. The closer the Euler angle is to 180 degrees, the greater the braking along X, while the closer θ is to 90 degrees, the more rapidly the vertical velocity will increase, and so on. In this reference case, the lander begins its descent with an angle of about 145 degrees, primarily to reduce the significant horizontal velocity. On the other hand, it also has good thrust capability in the vertical direction, which allows it to sustain against lunar gravity and slightly decrease altitude over time, as shown by the reduced velocity values. The initial angular velocity is very low but positive. Moreover, since there is initially a singular arc for the torque, it will increase only slightly due to the moment of inertia. Therefore,

during this initial phase, the pitch angle progressively increases up to 158.47 degrees. This results in greater reduction of the downrange velocity over time. It is possible to observe from the corresponding graph in Figure 7.1.2 how the slope (i.e., the respective acceleration) of the V_x(t) curve arches. On the other hand, the vertical velocity shows an almost parabolic trend. In fact, despite the lander's increasingly horizontal orientation, the vertical thrust component applied over time is sufficient to reverse the trend of its velocity. If it initially decreases, after reaching a minimum point where the acceleration towards the ground is zero, it begins to increase. Therefore, the spacecraft continues to fall gradually with a lower value of velocity until the end of the singular arc. Subsequently, in order to prepare for the vertical descent phase, there are the two trajectory arcs for the rotation of the yaw axis. As mentioned, first a negative torque is applied to the vehicle, in order to suddenly reduce the pitch angle, and then a positive one to stop the resulting rotational motion. Therefore, it can be seen how this causes the progressive decrease of the angular velocity in the second phase of the mission. The following arc, instead, will have the duration necessary to cancel out the ω and the downrange velocity, as well as to converge θ to 90 degrees. It is also noted how the vertical velocity after the singular arc abruptly changes its slope following the decrease of the pitch angle. The Figure 7.1.7 and the Figure 7.1.6 show the time evolution of the mass and the switching function for thrust (S_T), respectively. They are useful for understanding and comparing with the next downrange analyses. . It can be observed that the mass decreases linearly according to equation 5.12, which considers only the contribution of thrust and not the one of the thrusters for attitude control. On the other hand, S_T is the variable responsible for the bang-bang control of thrust, with which the lander's performance can be optimized. In particular, it remains positive throughout the singular arc, indicating that the thrust must be equal to the maximum available thrust, as in this case. In the subsequent phases, there is a drop in its value, and it becomes negative. Therefore, the engine should be turned off, but in this case, constant control of the spacecraft's velocity and position is necessary to ensure a safe and soft landing. However, the cost function is maximized in the rotation phase, while it is not during the vertical descent since the Hamiltonian cannot be null. This is deliberately done to push vertically and then to arrest the fall accurately.

7.2 Augmented Thrust Cases

The initial value of the thrust switching function in the just presented case, being relatively high, suggests that an increase in thrust can be evaluated to search for solutions that further optimize the powered descent trajectory. According to Pontryagin's maximum principle, increasing available thrust leads to a reduction in the initial value of S_T for the problem considered. Consequently, it is supposed that for values of it close to zero, a more efficient maneuver should be obtained. Thus, the aim of this study is to search for the thrust level that results in the least amount of propellant consumed. Therefore, with the same philosophy as the previous case where the initial position was left free, trajectories that take into account a progressive increase in thrust capacity of 12.5% at a time will be analyzed. Some of these cases are reported below for a better understanding of what has been done.

As expected, having a higher thrust capability enables the lander to decelerate its horizontal velocity and descend faster. As shown in the Figure 7.2.9, the $V_x(t)$ curves become

T $[kN]$	R_{x_0} $[Km]$	t_S $[s]$	t_2 $[s]$	t_3 $[s]$	t_4 $[s]$	θ_i $[deg]$	ω_i $[rad/s]$	θ_f $[deg]$	ω_f $[rad/s]$	M_f $[Kg]$
4	-533.30	553.71	3.063	3.059	2.457	145.34	0.00036	158.47	0.00046	809.14
5	-405.15	427.04	3.112	3.110	2.142	155.26	0.00022	160.77	0.00023	834.51
7.5	-258.07	274.38	3.175	3.177	1.675	169.40	-0.00036	163.85	-0.00034	855.24
10	-191.09	202.89	3.188	3.197	1.415	177.88	-0.00116	164.75	-0.00108	859.16
12.5	-152.37	161.14	3.179	3.195	1.245	184.18	-0.00214	164.66	-0.002	858.25
15	-127.09	133.73	3.158	3.182	1.123	189.37	-0.00328	164.08	-0.0031	855.31

Table 7.2.2: Results for increasing thrust

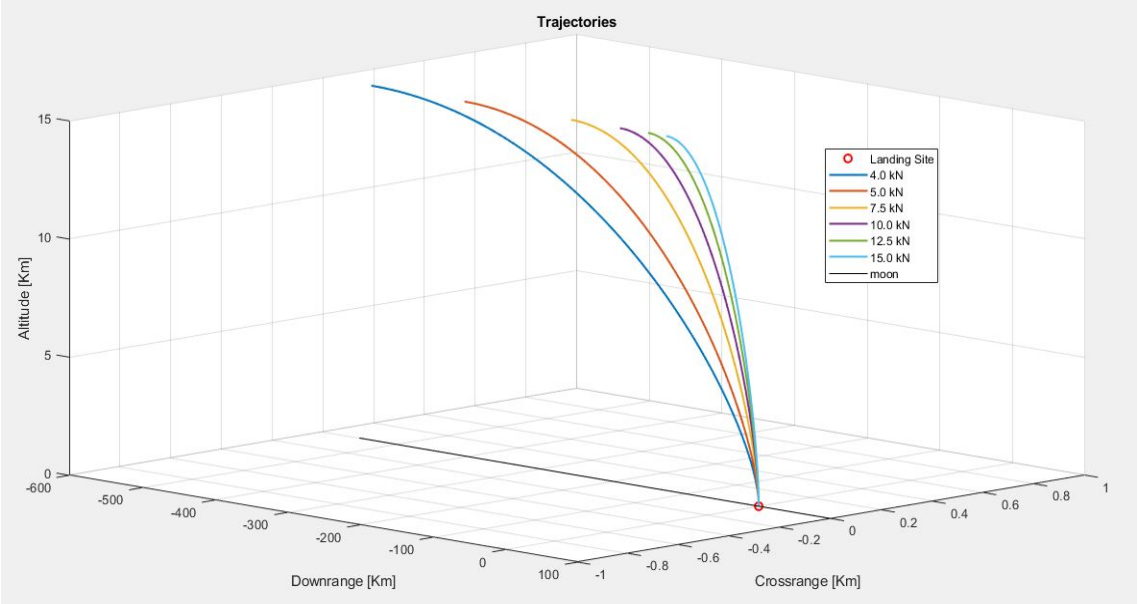


Figure 7.2.8: Trajectories for increasing thrust

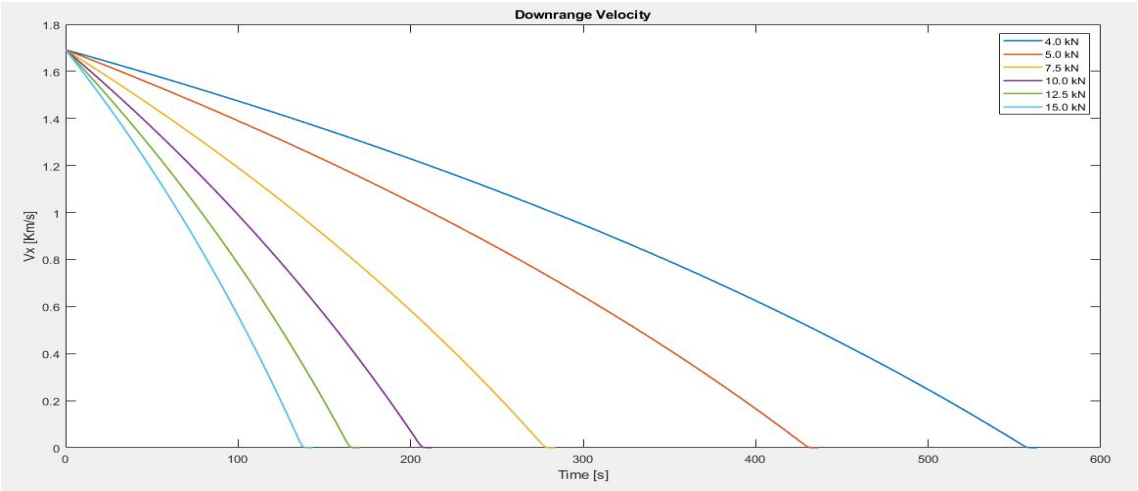


Figure 7.2.9: Dowrange Velocities for increasing thrust

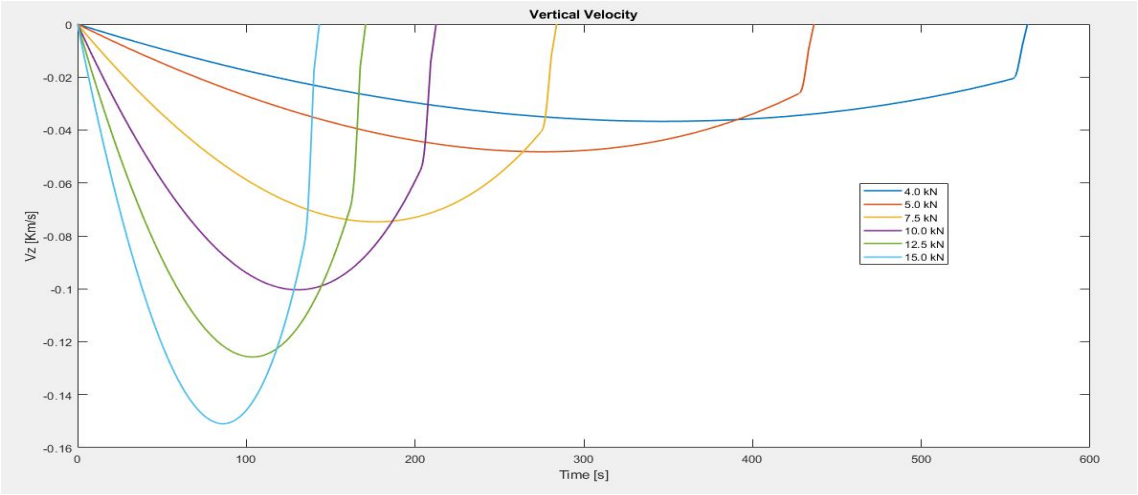


Figure 7.2.10: Vertical Velocities for increasing thrust

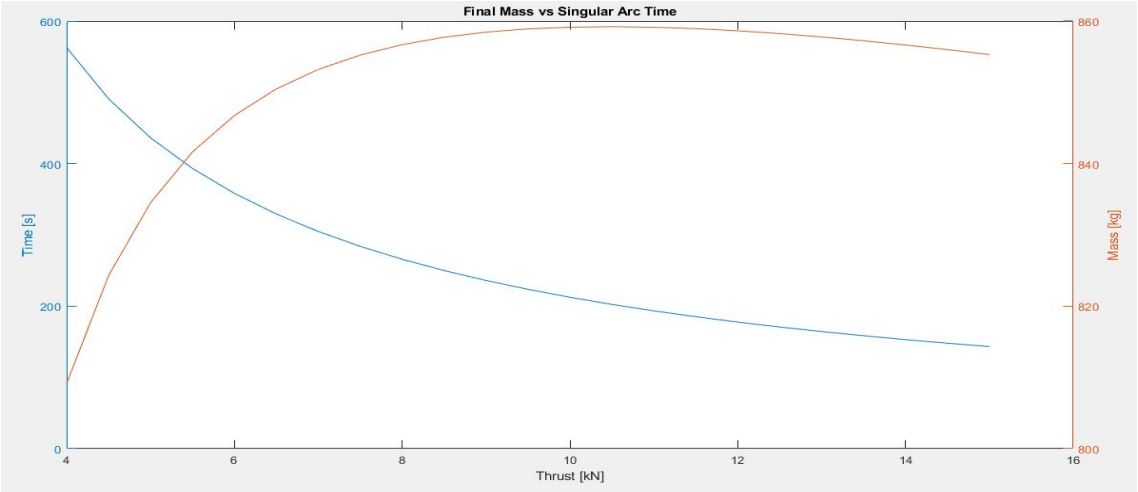


Figure 7.2.11: Final Mass vs Singular Arc time in function of thrust

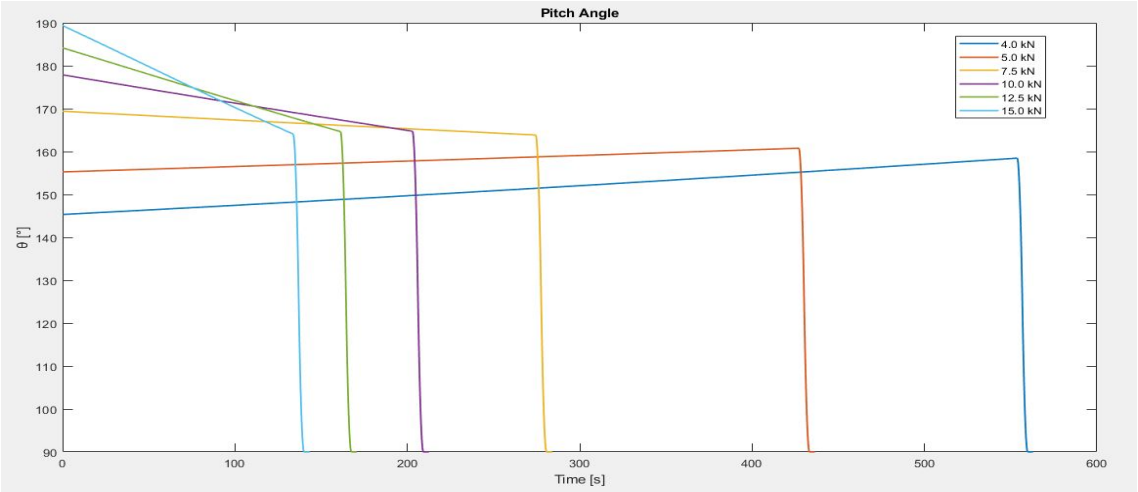


Figure 7.2.12: Pitch Angles for increasing thrust

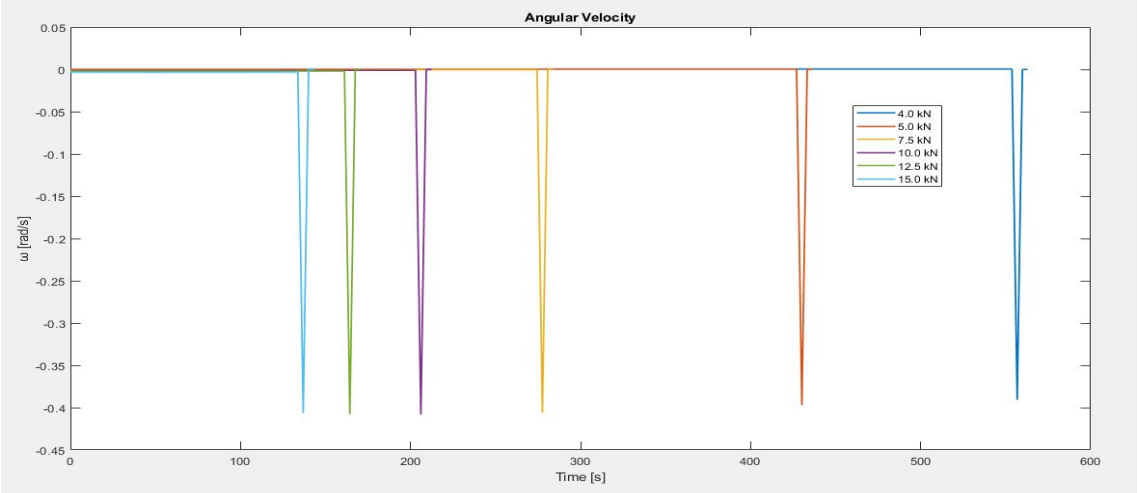


Figure 7.2.13: Angular Velocities for increasing thrust

steeper, indicating a more vigorous braking. Consequently, it can complete the descent maneuver in less time. Despite the greater thrust, this results in a lower total impulse and benefits the mass of propellant used, since we are assuming the same specific impulse of the thrusters. This translates into a shorter downrange distance, as clearly seen in the Figure 7.2.8. Therefore, since the vehicle must reach the same landing site, the initial starting point of the trajectory can progressively approach it. The modification of the initial downrange is possible by increasing the waiting time in the Parking Lunar Orbit, in order to anticipate the DOI phase. This operation does not influence the propellant consumption, as the spacecraft is assumed to orbit ideally around a circular orbit at an altitude of 100 km around the Moon, not affected by the perturbative effects generated by a non-uniform gravitational field. To achieve this reduction in downrange and the corresponding velocity, the lander at launch is oriented at an increasingly greater pitch angle, tending towards 180 degrees as shown in Figure 7.2.12. As mentioned previously, this indicates a sudden decrease in altitude as the orbital velocity is reduced more intensely. From Figure 7.2.10, it can be noted how the vertical velocity with which the spacecraft falls increases in magnitude as the thrust increases. In the case with a thrust of 15,000 Newtons, it even starts the descent with an angle close to 190 degrees, meaning it utilizes a small portion of the thrust to push downwards. By increasing the thrust, the lander is capable of accelerating its descent, consequently conserving propellant.

Additionally, it can be observed how there is a reversal in the trend of θ. Previously, during the silgular arc, the angle increased. However, as the thrust increases, the lander starts with a gradually smaller angular velocity, eventually becoming negative. It will continue to grow due to the moment of inertia, but it remains negative for the entire duration of this first phase if it was initially negative. Therefore, in the case of high thrust levels, the pitch angle decreases, although it always remains higher than the reference case at the beginning of the rotation phase. For this reason, the durations of the second and third arcs are longer. In fact, it is necessary to apply the torques, whose value does not change, for a longer time to make θ converge to 90 degrees. This also results in smaller minimum point in the trend of ω. In particular, it can be observed how for large thrusts, the third phase lasts longer than the second since in these cases it is necessary to obtain

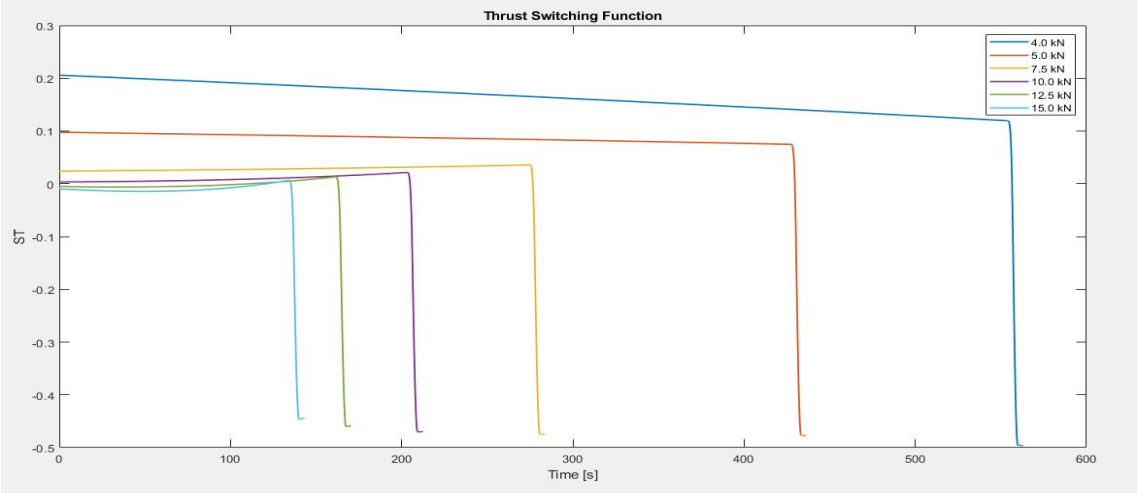


Figure 7.2.14: Thrust Switching Function for increasing thrust

a net positive moment in order to stop a negative rotational motion before the vertical descent.

Therefore, as introduced at the beginning of the paragraph, increased thrust results in a shorter maneuver and thus consumes less propellant. This can also be seen in Figure 7.2.14 as the initial value of S_T gradually approaches zero. From Figure 7.2.11, it can be observed that the optimal trajectory would be achieved with a thrust of approximately 10800 *Newtons*. For higher values of thrust, there is a slight decrease in the final mass compared to previous cases, which may seem to contradict the Pyontrigan principle. However, this is due to the fact that the thrust level is too high for the rotation arcs and vertical descent, where, due to the constraint to land with the engine on, PMP is violated. As a result, it is not possible to nullify the Hamiltonian, which becomes even larger in magnitude, and therefore maximize the cost function. Consequently, during these phases, a reduction in thrust can be considered in order to further increase the lander's performance, although it would not be advisable to turn off part of the propulsion system so close to the ground, considering that the loss in final mass is negligible from a mission safety standpoint.

7.2.1 Comparison with the shutdown case

In the previous paragraph, it was shown how for high values of thrust, the associated switching function decreased until it became negative. As already extensively discussed earlier, the guidance algorithm suggests the application of a bang-bang control of the $T_{min} - T_{max}$ type. Consequently, it is necessary to add an initial coasting arc for thrust in order to obtain the optimal control law that maximizes the cost function. This would require changing the code to introduce a new unpowered phase at the beginning of the maneuver. However, this modification would be complicated and delicate, considering that additional constraints would be needed to avoid affecting the control laws applied in the subsequent phases. Alternatively, the same result can be achieved by turning off the engines where the switching function is negative during the singular arc and continuing to leave the initial position of the trajectory free. This way, a coasting phase is obtained,

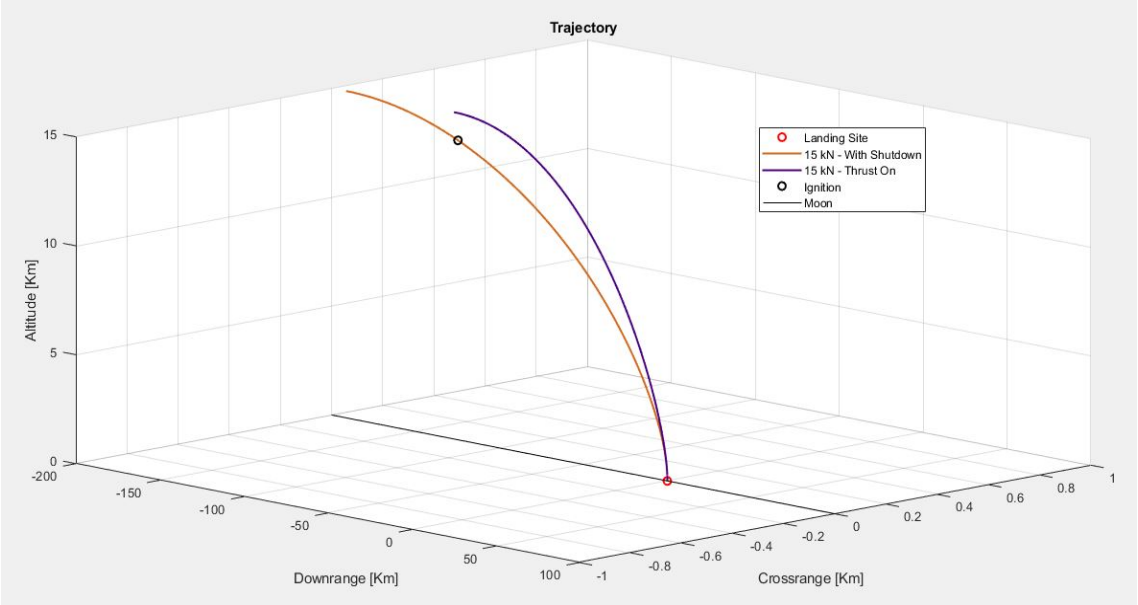


Figure 7.2.15: Trajectory with shutdown

and optimal thrust control is achieved. An improvement in performance is expected since part of the trajectory will be covered in free fall. The obtained results are presented below, assuming an available thrust of 15,000 Newton.

By comparing this case with the one where the thrusters remain on, the most noticeable aspect observed from the graphs is the increase in the time taken to complete the maneuver. In fact, in Figure 7.2.15, it can be seen that the lander starts from a position further away from the landing site, resulting in an increase in downrange due to the shutdown of the engines. During the first part of the trajectory, the horizontal velocity remains constant as revealed in Figure 7.2.16. The same is true for the time evolution of the mass, as displayed in Figure 7.2.20. On the other hand, since the spacecraft is subject only to the gravitational acceleration, the reduction of its vertical velocity component is less pronounced than in the previous case, as shown in Figure 7.2.17. Thus, since the lander is in free fall, it can start its descent by orienting itself parallel to the lunar surface, i.e., with a pitch angle of exactly 180 degrees. Also, in this case, the angular velocity remains negative throughout the first phase, but its magnitude is smaller than in the case with continuous thrust. As a result, the angular displacement decreases more slowly over time, reaching a value of approximately 170 degrees at the end of the singular arc, which is higher than in the previous case. For this reason, the rotation phases will have a longer duration to achieve a greater angular velocity change in the vehicle. Finally, as expected, by introducing the coasting phase, it is possible to obtain a thrust control law that further optimizes the landing maneuver. This is evident in Figure 7.2.21, where the final mass improves, giving the propulsion system the opportunity to be shut down.

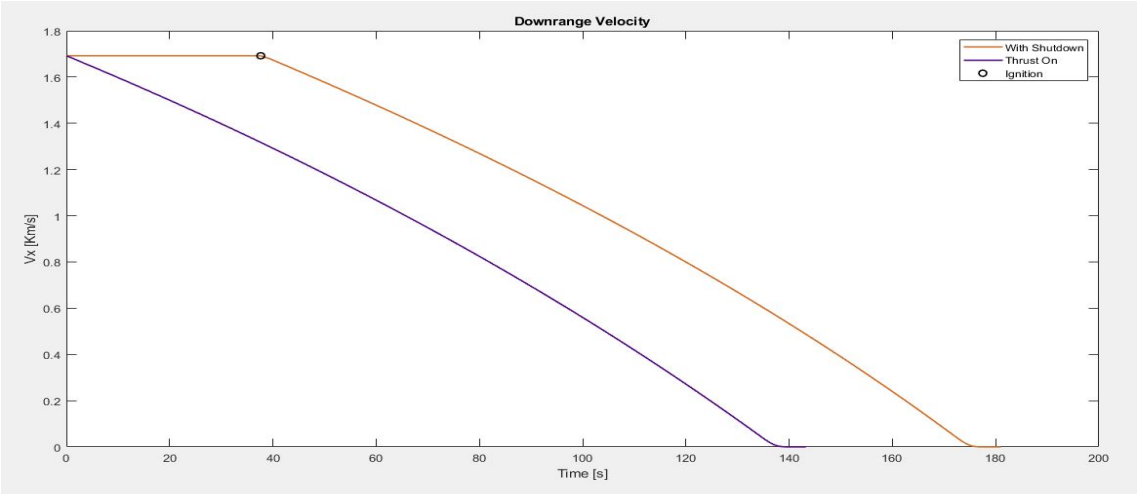


Figure 7.2.16: Dowrange Velocity with shutdown

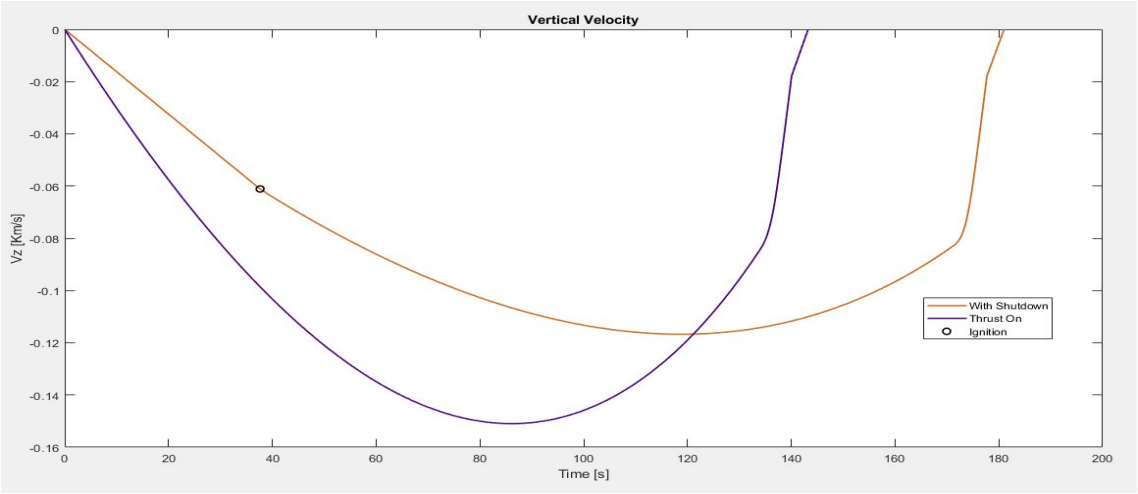


Figure 7.2.17: Vertical Velocity with shutdown

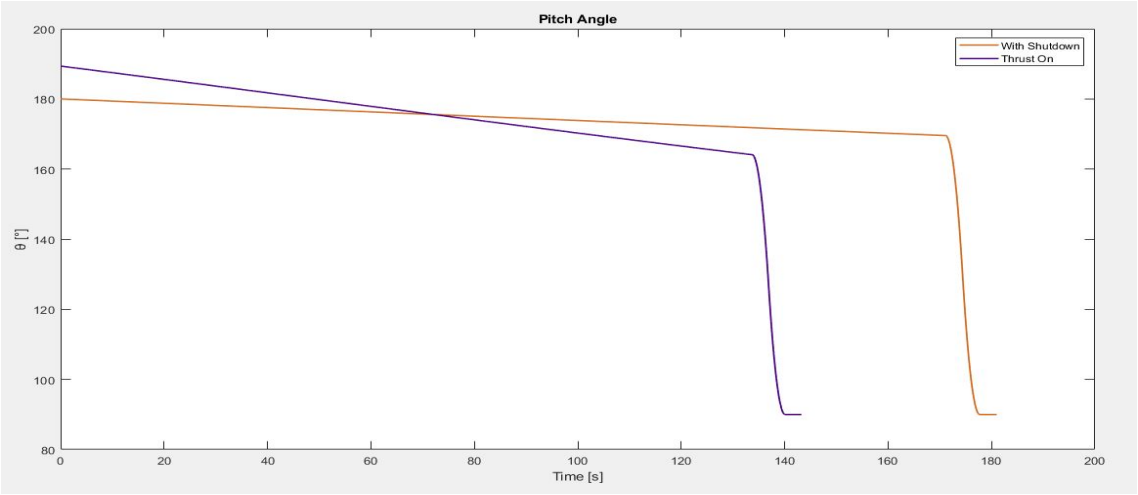


Figure 7.2.18: Pitch Angle with shutdown

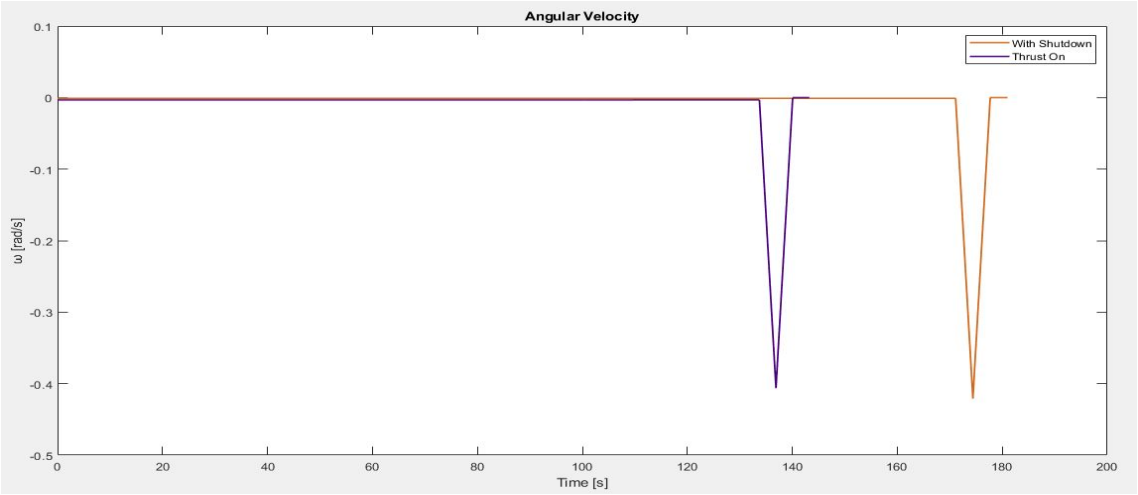


Figure 7.2.19: Angular Velocity with shutdown

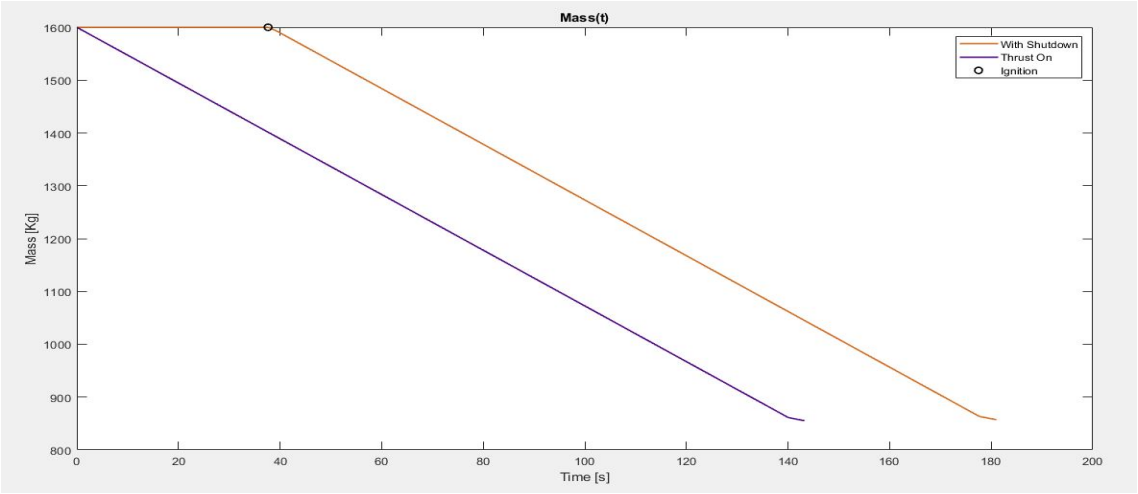


Figure 7.2.20: Mass in function of time with shutdown

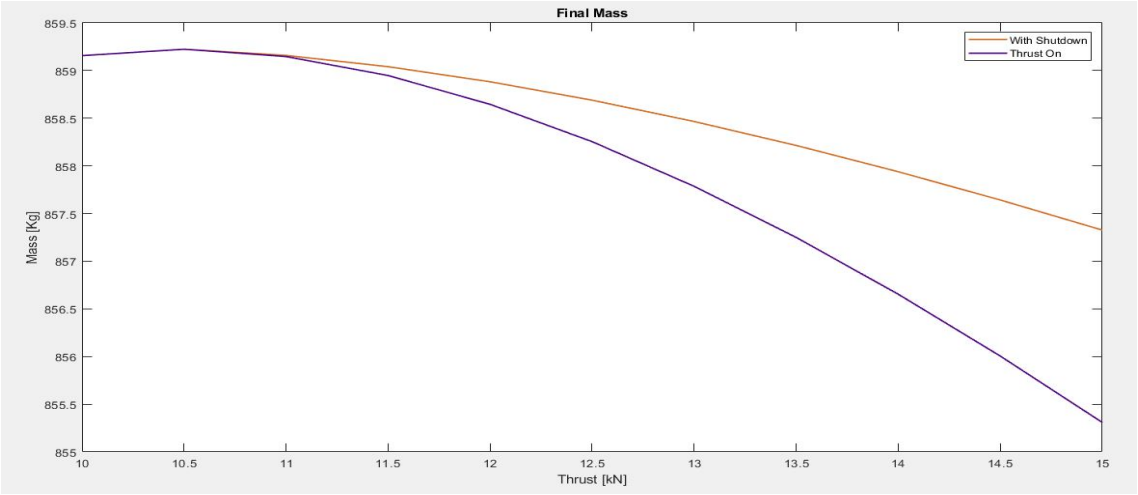


Figure 7.2.21: Final Mass in function of thrust with shutdown

7.3 Cases with fixed R_{x_0} *and* R_{x_f}

In the previous chapter, it was shown how by exploiting the Pontryagin principle it is possible to complete the landing maneuver with a lower expenditure of propellant. In doing so, it is possible to transport a higher payload mass to the lunar surface as well as to design a more compact lander by reducing the tank structure. Following this, it was also observed how for high levels of thrust, a coasting arc must be introduced in the first phase in order to improve losses due to vertical descent. In particular, this latter result will prove useful for the purposes of the following analysis. This section focuses on the search for possible trajectory correction maneuvers following a real-time variation of the touchdown point. This may be due to unforeseen circumstances or simply to uncertainty about the site, which is clarified as the vehicle approaches the ground. Therefore, we assume the scenario where the lander, once it reaches the initial position of the PDI, receives or identifies the coordinates of the new landing site. The objective of this study is thus to evaluate the effectiveness of the guidance algorithm in programming the adjustment of the powered descent maneuver. They will be considered both the case of lengthening and shortening of the downrange to be covered, in order to ensure a certain margin in which the lander is able to operate precisely and still complete the mission safely. To this end, the previous code has been modified by fixing the values of both the initial and final positions. Therefore, the adjoint variable λ_{x_0} in this case is no longer zero but will be an unknown parameter for which an initial value must be hypothesized. In the simulation, reference is made again to a thrust of 4000 Newtons. Consequently, the starting point is set at a distance of -533.30 kilometers from the center of the reference system, while the altitude remains unchanged. One would expect the lander to have the ability to extend its reach by decelerating less horizontally, reducing its speed less drastically. Conversely, it could shorten the maneuver by starting with an orientation that allows for greater downrange speed braking. Additionally, it is assumed that, in the first case, the trajectory time and propellant consumption associated with it would increase. In contrast, in the second scenario, the lander should complete the mission with a higher final mass, as it would cover a shorter distance. However, it should be noted that in both cases, the code will search for optimal solutions to adjust the trajectory.

7.3.1 Analysis

Initially, the need for an increase in downrange upon arrival at the PDI starting point was evaluated. To achieve this, the touchdown position was progressively increased. The results obtained for some cases are shown in the Table 7.3.3 and graphs below. It can be inferred that the lander immediately corrects its attitude in the first phase of the trajectory to reach the new landing site. It is particularly noticeable that it begins the maneuver with a gradually decreasing pitch angle as the downrange to be covered increases. The thrusters initially push more in the vertical direction compared to the reference case. By doing so, the vehicle can sustain itself against lunar gravity for a longer period, also due to the fact that horizontal braking is reduced. Consequently, it can actually travel a greater distance since the downrange decreases less rapidly. However, the case with high values of the required final position is of greater interest, as θ decreases so much that it reaches values less than 90 degrees. As shown in the Figure 7.3.23, the downrange velocity in-

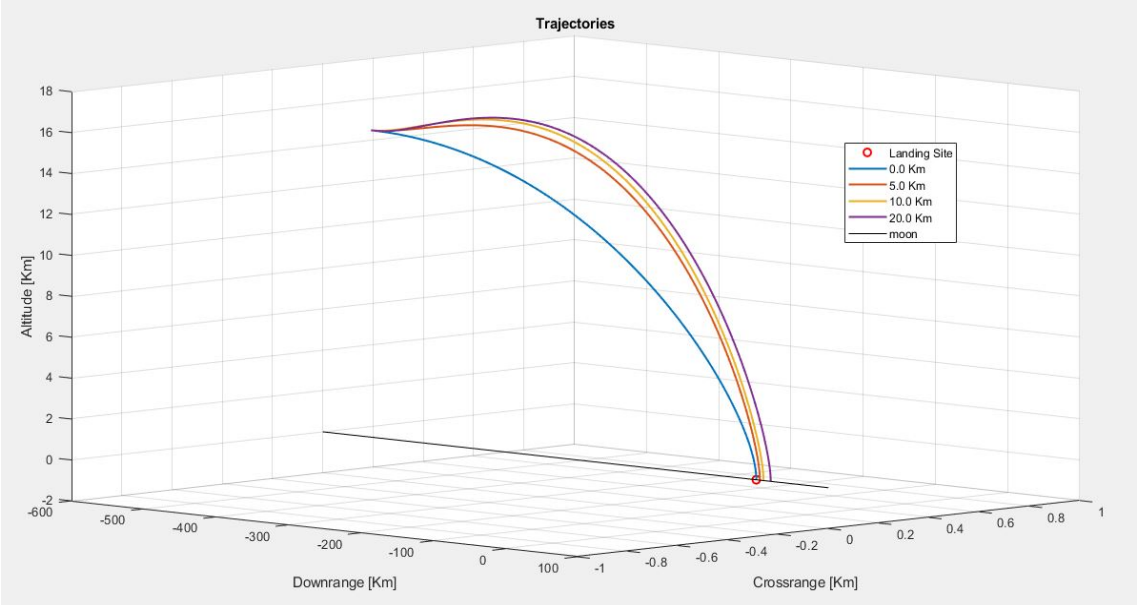


Figure 7.3.22: Trajectories for increasing landing site positions

R_{x_f} $[Km]$	t_S $[s]$	t_2 $[s]$	t_3 $[s]$	t_4 $[s]$	θ_i $[deg]$	ω_i $[rad/s]$	θ_f $[deg]$	ω_f $[rad/s]$	M_f $[Kg]$
0.0	553.71	3.063	3.059	2.457	145.34	0.00036	158.47	0.00046	809.14
1.0	554.04	3.012	3.011	2.456	137.99	0.00273	156.32	0.00012	808.81
2.5	554.60	2.998	2.997	2.454	125.78	0.01276	155.71	6.0e-05	808.07
5.0	555.54	2.991	2.990	2.451	105.50	0.04333	155.44	4.0e-05	806.76
7.5	556.50	2.990	2.989	2.448	87.72	0.07361	155.37	4.0e-05	805.43
10.0	557.45	2.990	2.989	2.445	73.21	0.09113	155.35	3.0e-05	804.10
15.0	559.35	2.990	2.989	2.439	52.59	0.09376	155.39	3.0e-05	801.43
20.0	561.25	2.992	2.991	2.433	39.52	0.08026	155.46	3.0e-05	798.77

Table 7.3.3: Results for increasing landing site positions

creases at the beginning of the maneuver, indicating that the spacecraft needs to accelerate horizontally to delay the start of the descent and to reach higher altitudes. In fact, from the Figure 7.3.24, it can be seen that the vertical velocity increases in every case, reaching positive values since it was initially zero. This is due to a greater vertical thrust component compared to the reference case and results in an adjustment of the altitude required to optimize the descent trajectory. This can be clearly observed from the plot of the different trajectories in the Figure 7.3.22. Afterward, the lander can start falling to the ground, and the two velocity components will assume their usual behavior. Therefore, after the initial instant, the pitch angle must increase progressively to reduce the downrange velocity.

For this reason, an increasing initial angular velocity is observed with increasing final landing site position. As shown in the Figure 7.3.27, a negative torque is applied at the beginning of the trajectory, which gradually disappears once the necessary position and attitude to start the descent have been achieved. Consequently, ω decreases over time and approaches zero at the end of the maneuver correction. In this way, the pitch angle increases rapidly and asymptotically reaches a value that is maintained until the end of the singular arc. This value is lower than that of the reference case, so the next two rotation phases will have a shorter duration. In this case, the duration of the second and third

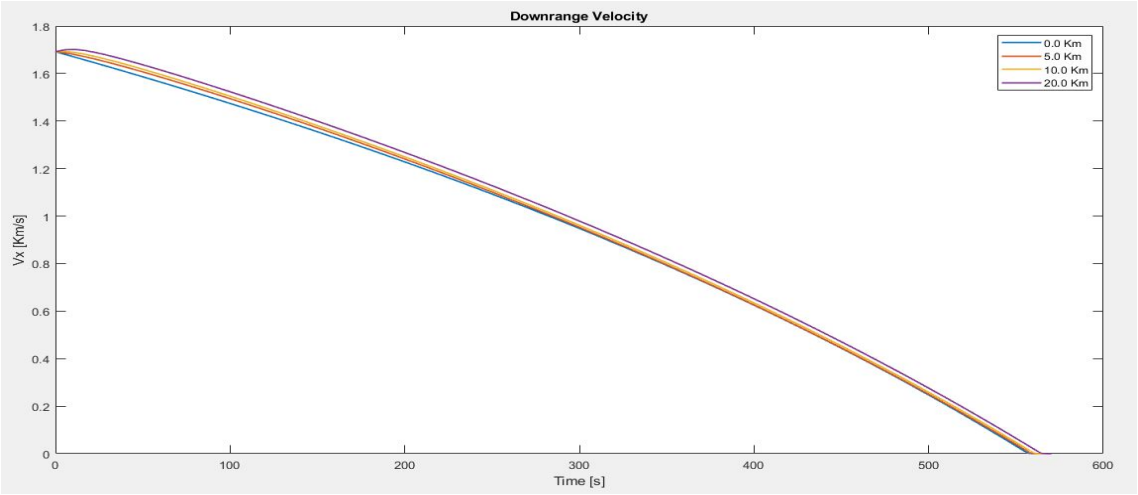


Figure 7.3.23: Dowrange Velocities for increasing landing site positions

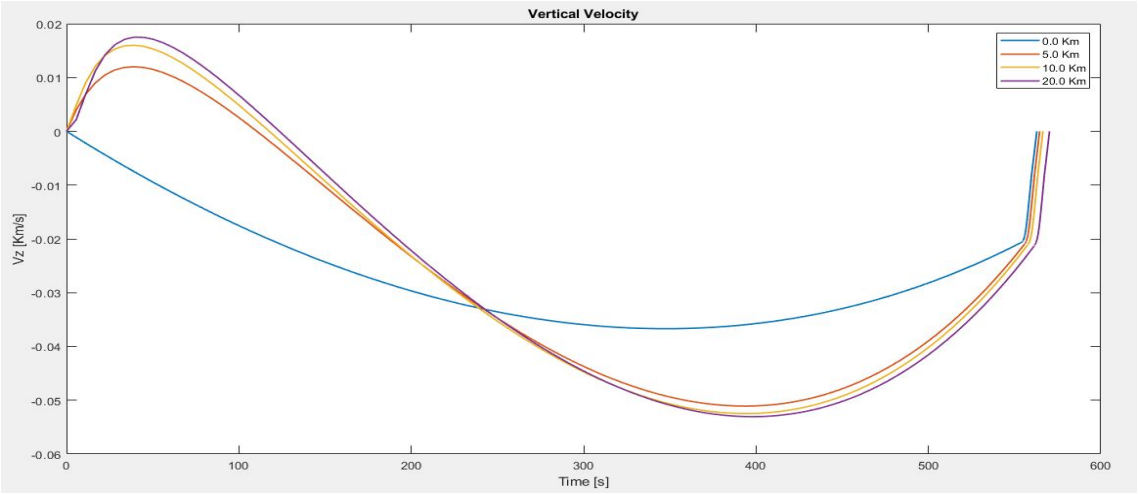


Figure 7.3.24: Vertical Velocities for increasing landing site positions

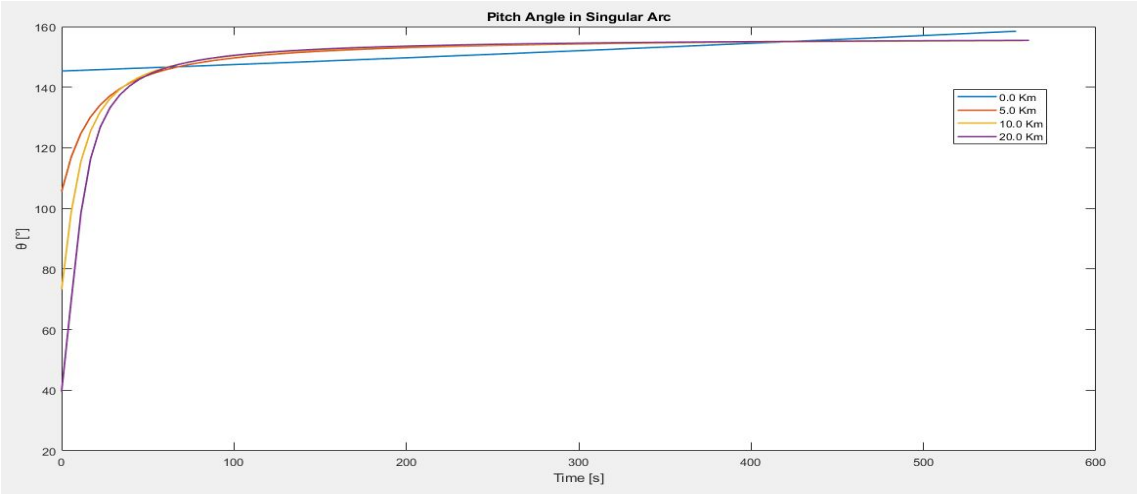


Figure 7.3.25: Pitch Angles for increasing landing site positions

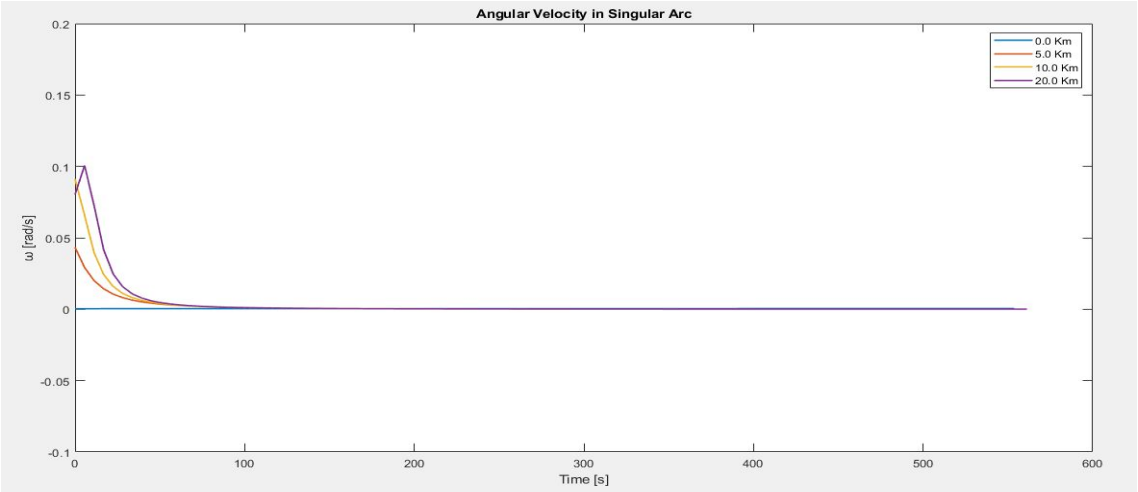


Figure 7.3.26: Angular Velocities for increasing landing site positions

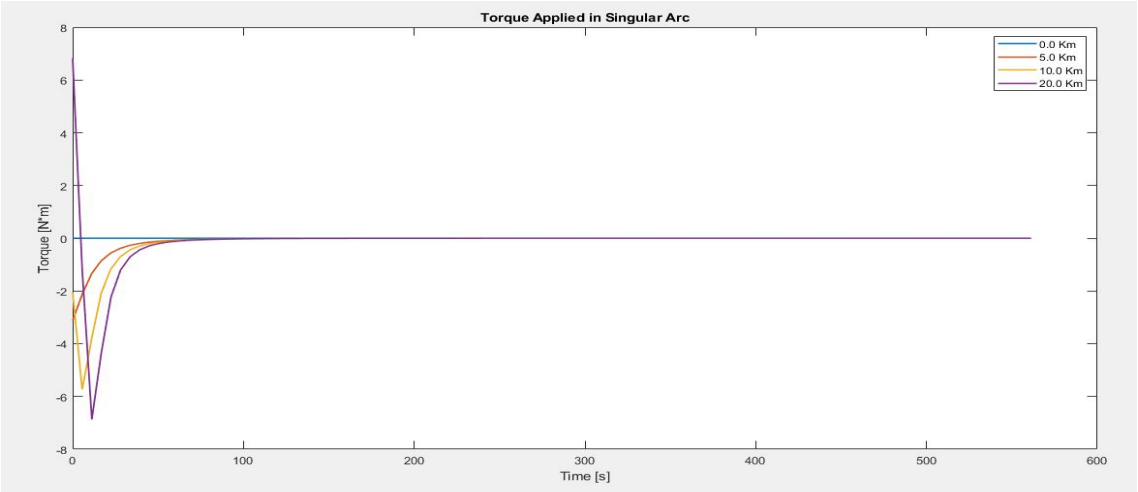


Figure 7.3.27: Torques Applied in Singular Arc for increasing landing site positions

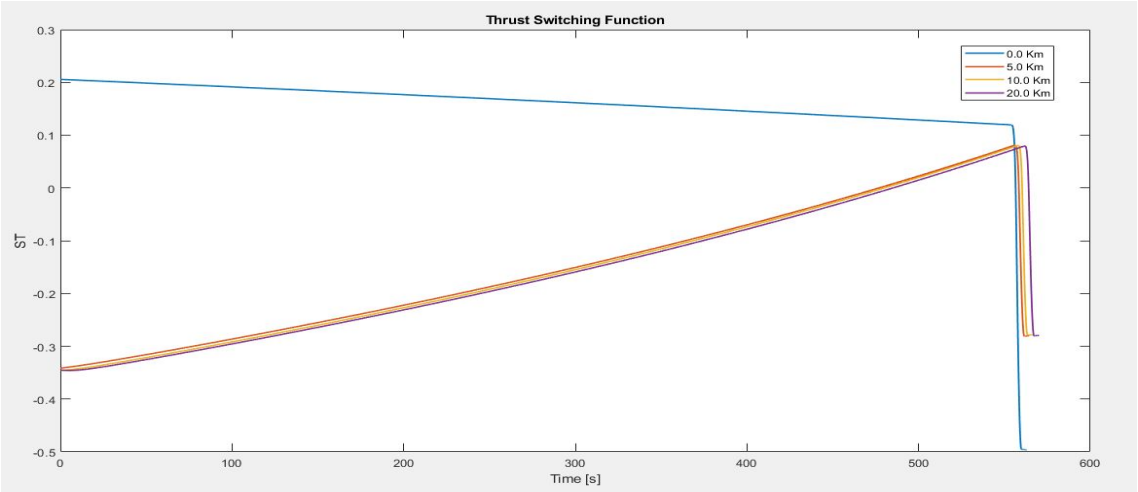


Figure 7.3.28: Thrust Switching Functions for increasing landing site positions

arcs will be similar since the angular velocities are negligible. On the other hand, the time of the singular arc increases linearly with the increase of the covered downrange. As expected, this results in a decrease in the final mass with which it is possible to land, since the propulsion system remains on for a longer time.

Nonetheless, the spacecraft is capable of correcting its trajectory and landing precisely at greater distances. On the other hand, attempts have been made to reduce the landing site position for the reasons previously explained. In this case, however, the code is able to converge correctly up to about 270 meters before the center of the reference system. As we try to shorten the downrange, a significant increase in errors has been noticed on the boundary conditions, which is not acceptable according to the assumptions made in Section 5.6. Additionally, for shorter distances, the trajectories found have a negative time for the vertical descent phase. This implies that the lander is no longer able to find a optimal solution, as it would pass through negative vertical distances. Nevertheless, it can be deduced that in the previous solution ($R_{x_f} = 0$) the vehicle departs with a thrust direction that avoids crashing into the ground. If oriented in a direction more parallel to the surface, it would acquire too much vertical velocity, which would no longer be able to be arrested with this level of available thrust. Since reducing consumption also means decreasing the ignition time of the main thruster, it is not possible to optimize the trajectory with this level of thrust because the time to land cannot be reduced. Therefore, to overcome this problem, possible solutions will be investigated in the next section.

7.3.2 Fixed R_{x_0} *and* R_{x_f} With Shutdown Case

From Figure 7.3.28 of the above analysis, it can be noted that even in this case the bang-bang control suggests a framework for thrust of $T_{off} - T_{on}$ type, since the initial values of the Switching Function are less than zero as the downrange increases. Furthermore, it was shown how increasing the final position of the landing site appropriately increases the amount of propellant to be used for the maneuver. Therefore, as discussed in paragraph 7.2.1, the possibility of a first free-fall phase (coasting arc) was evaluated with the aim of reducing fuel consumption. To this end, a comparison is reported below, with and without the possibility of turning off the main propulsion system, for the case in which it is desired to land 20 km further from the original site.

From the graphs shown above, it can be observed that ignition occurs a few moments after the start of the maneuver, specifically 5.64 seconds later. There is a short freefall phase that the lander uses to orient itself properly before braking. Figure 7.3.32 shows that the pitch angle at the intial point is similar to the reference case in Paragraph 7.1, only slightly lower. This is to allow the vehicle to withstand more against lunar gravity, so as not to make the descent too rapid. In this way, it will be able to land at a greater distance, as the horizontal braking will be less abrupt. Therefore, in the initial point, there is no need to have a high angular velocity, as shown in Figure 7.3.33, nor a torque for attitude correction. Instead, a small initial value for ω and a small positive moment of inertia are used to increase θ during the descent. The landing maneuver, therefore, is less demanding than the previous case, as also shown by the trend of the vertical velocity in Figure 7.3.31. In fact, along the entire trajectory, it will remain contained, as the direction of the thrust is optimized. On the contrary, in the case with continuously running thrusters, the lander initially pushed upward and then had to decelerate more in the vertical

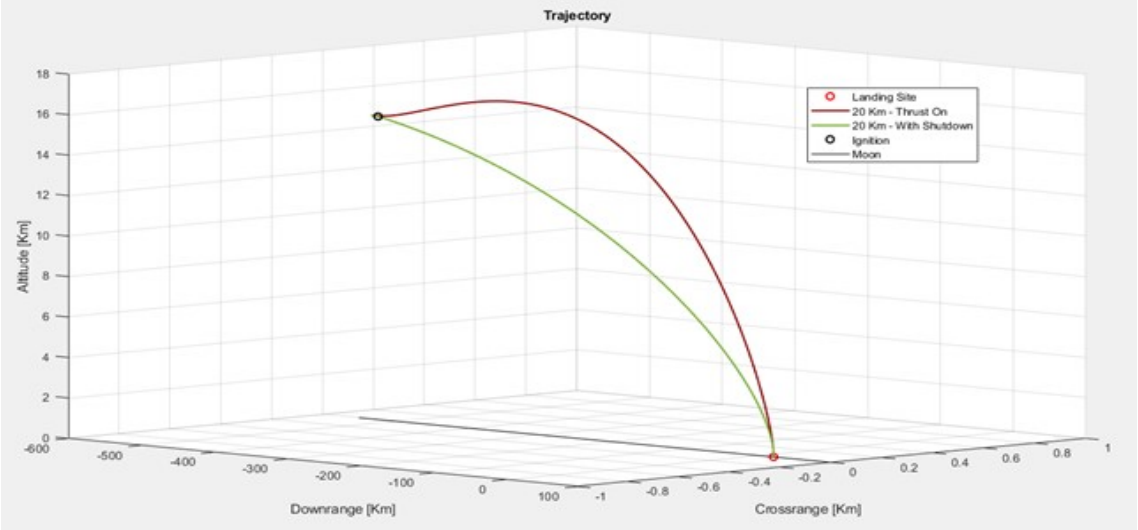


Figure 7.3.29: Trajectories Comparison

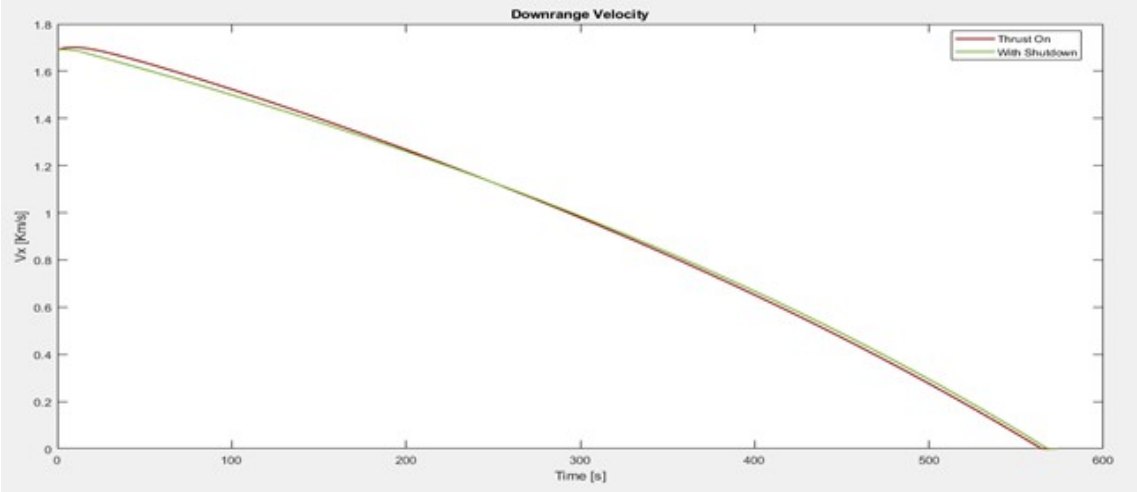


Figure 7.3.30: Dowrange Velocities Comparison

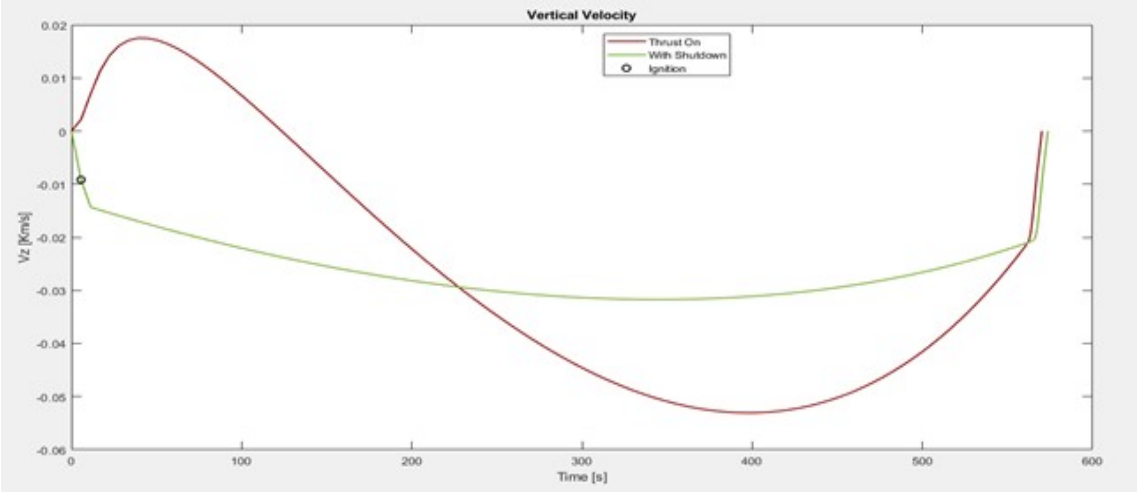


Figure 7.3.31: Vertical Velocities Comparison

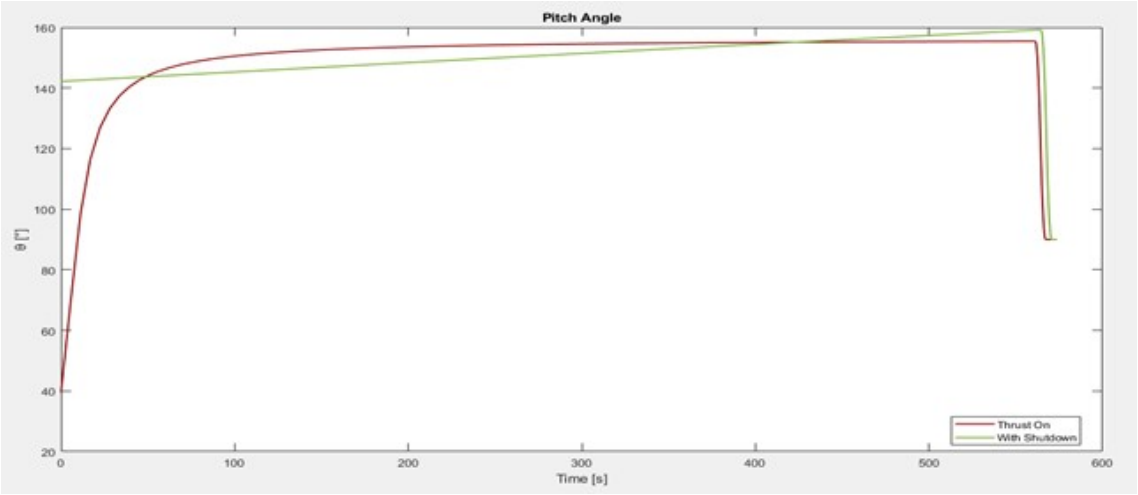


Figure 7.3.32: Pitch Angles Comparison

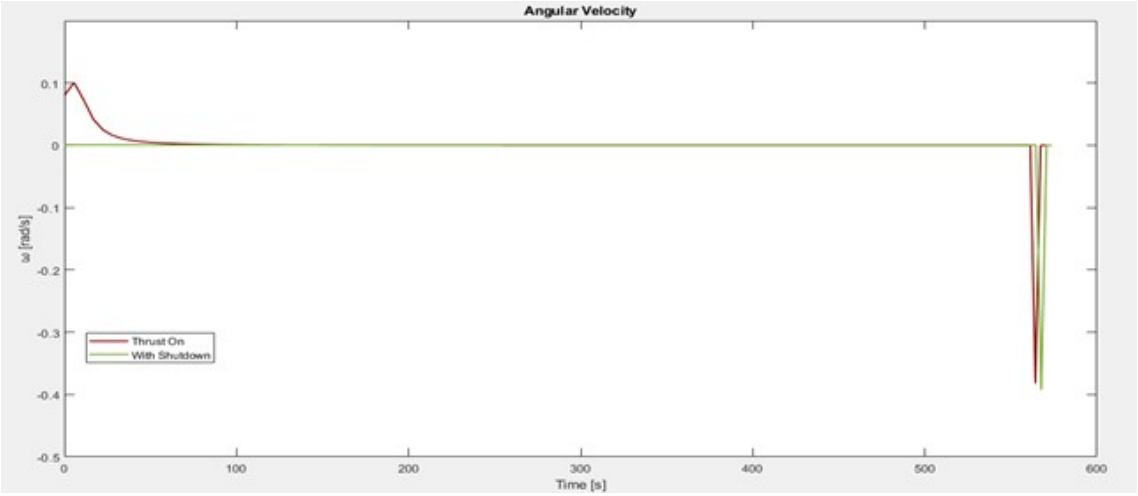


Figure 7.3.33: Angular Velocities Comparison

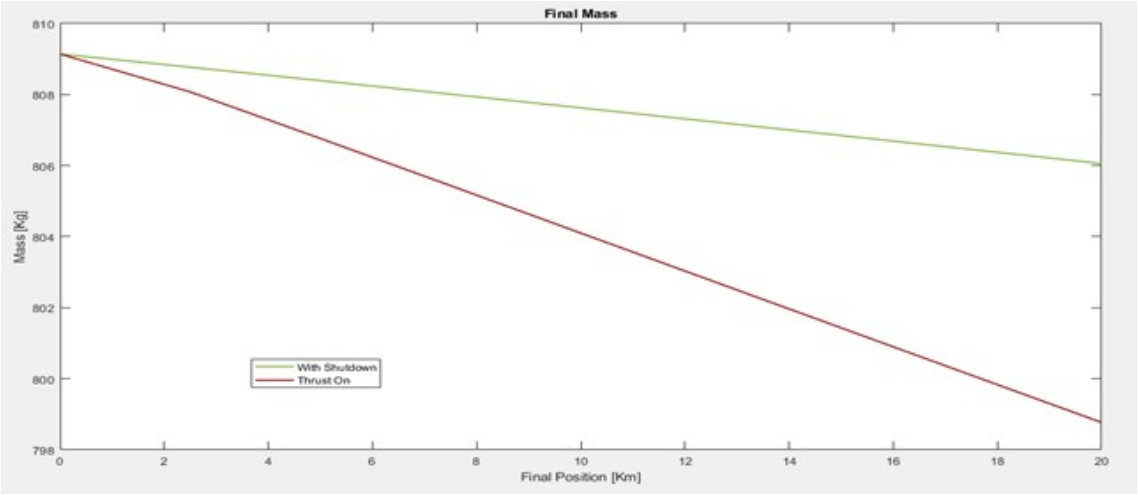


Figure 7.3.34: Final Masses Comparison

R_{x_f} [Km]	t_S [s]	t_2 [s]	t_3 [s]	t_4 [s]	θ_i [deg]	ω_i [rad/s]	θ_f [deg]	ω_f [rad/s]	M_f [Kg]
0.0	171.12	3.289	3.297	1.1239	180.00	-0.00108	169.54	-0.00104	857.33
-2.5	169.67	3.283	3.292	1.1239	180.69	-0.00119	169.27	-0.00114	857.32
-5.0	168.22	3.277	3.286	1.1239	181.36	-0.00131	169.01	-0.00123	857.30
-10.0	165.31	3.265	3.276	1.1239	182.63	-0.00154	168.51	-0.00141	857.24
-15.0	162.41	3.254	3.266	1.1238	183.82	-0.00177	168.04	-0.00157	857.16
-30.0	153.68	3.224	3.240	1.1235	186.82	-0.00247	166.78	-0.00198	856.76
-45.0	144.92	3.200	3.218	1.1230	188.95	-0.00318	165.75	-0.00230	856.19
-60.0	136.13	3.180	3.200	1.1226	190.17	-0.00384	164.91	-0.00256	855.51

Table 7.4.4: Results for decreasing landing site positions

direction. For this reason, the pitch angle at the end of the singular arc in this case will be greater. Additionally, it can be noted that the maneuver now lasts longer. With the same downrange, the total impulse needed to complete the mission does not change. Therefore, delaying the ignition of the thrusters will result in the trajectory being covered in a longer time. However, activating the thrusters a few seconds after landing allows the lander to save more than 7 Kg of propellant, returning to similar final mass values as the reference case. In this way, the algorithm is able to correct the maneuver more efficiently than in the previous case, while still ensuring the same level of safety in the operations.

7.4 Reduced Downrange Cases

As discussed in the previous section, contrary to what was predicted, the lander is not able to shorten trajectory with a thrust of 4000 Newton in order to anticipate landing. For this reason, in this study, we have decided to use the results from paragraph 7.2. By using a higher thrust, it was found that an initial coasting phase needs to be added to increase the efficiency of the maneuver. In this case, the aim is to exploit this coasting phase with the thrusters off in order to land earlier. Unlike before, the lander now has enough time to brake and, consequently, to find an optimal solution by decreasing the maneuver duration. By doing so, the code should be able to reduce the downrange by activating the thrusters a few moments before during the free fall. To this end, a thrust of 15000 Newton has been adopted, so that the lander begins the descent from a distance of 191.29 kilometers from the center of the reference system. The case discussed above where the landing site is at zero downrange is taken as reference. Below, the results and their analysis are presented.

From the trajectory graphs shown in Figure 7.4.35, it can be observed how it is actually possible to reduce the downrange by exploiting the coasting phase. In particular, it is noticed how the lander increasingly anticipates the ignition of the thrusters as the final required position decreases. This reduction is possible up to about 64 kilometers before the original landing site, that is the difference in downrange between the two different cases of continuously powered thrusters and thrusters with the ability to turn off for the adopted thrust value. To achieve this reduction, the initial pitch angle increases beyond 180 degrees, so there is a slight downward thrust component at ignition. The closer the site distance, the more θ will increase. Therefore, it can be seen in the Figure 7.4.37 that, after the first free-fall phase, the slope of the vertical speed curves gradually increases, reaching smaller negative peaks.

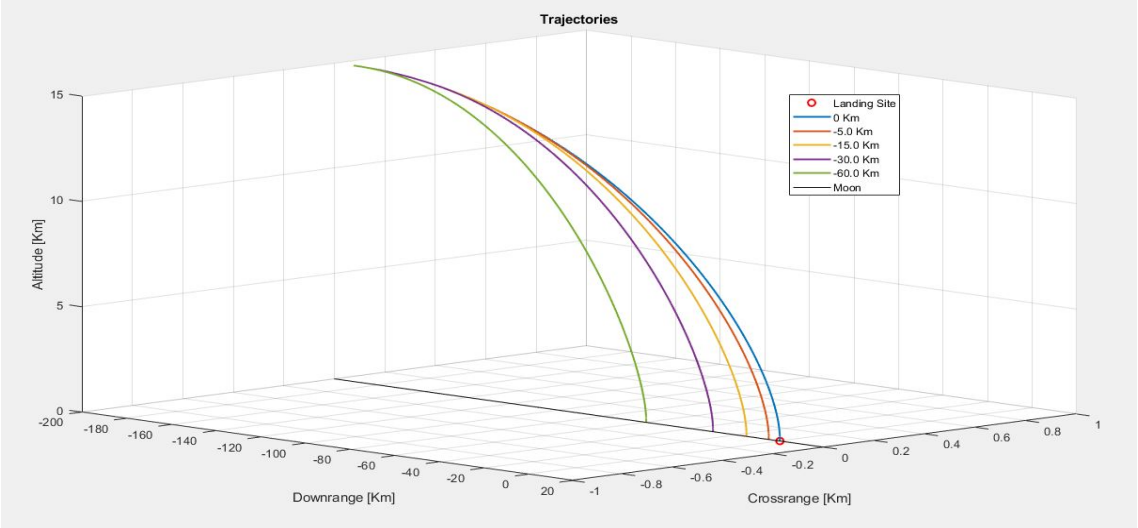


Figure 7.4.35: Trajectories for decreasing landing site positions

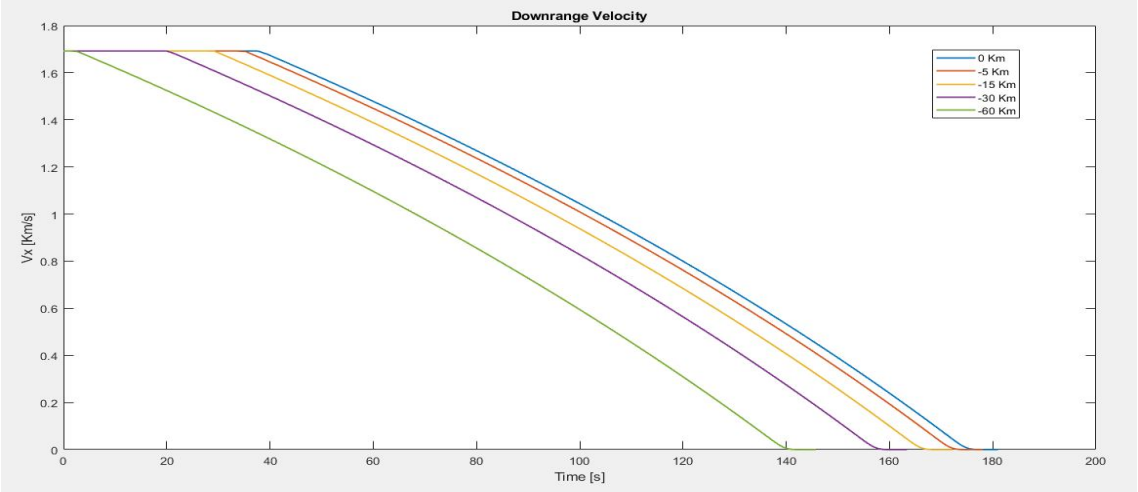


Figure 7.4.36: Dowrange Velocities for decreasing landing site positions

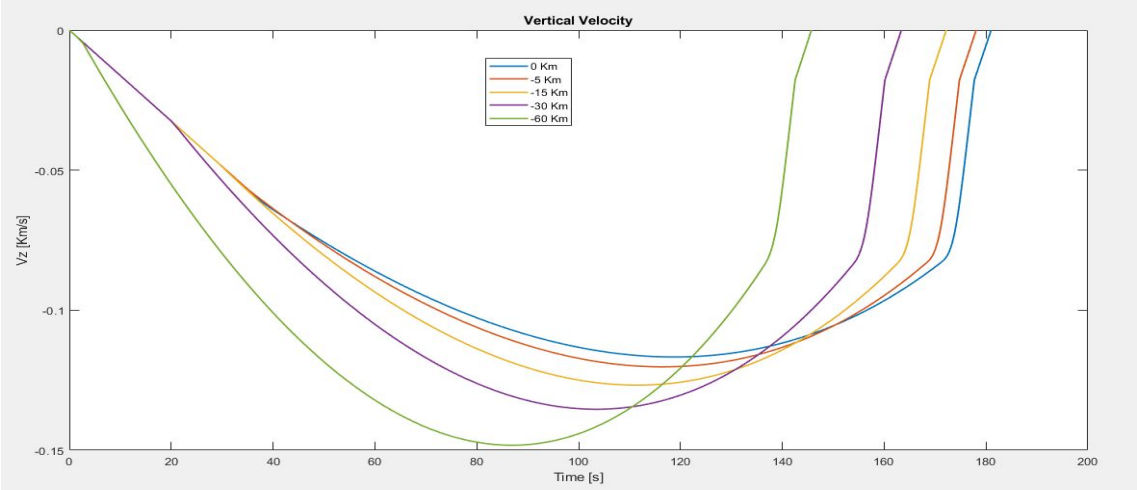


Figure 7.4.37: Vertical Velocities for decreasing landing site positions

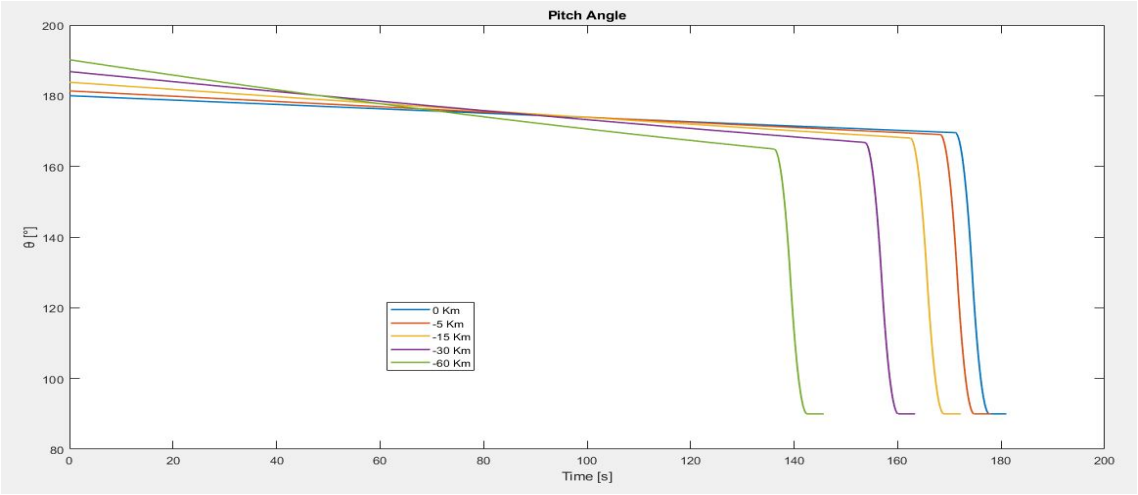


Figure 7.4.38: Pitch Angles for decreasing landing site positions

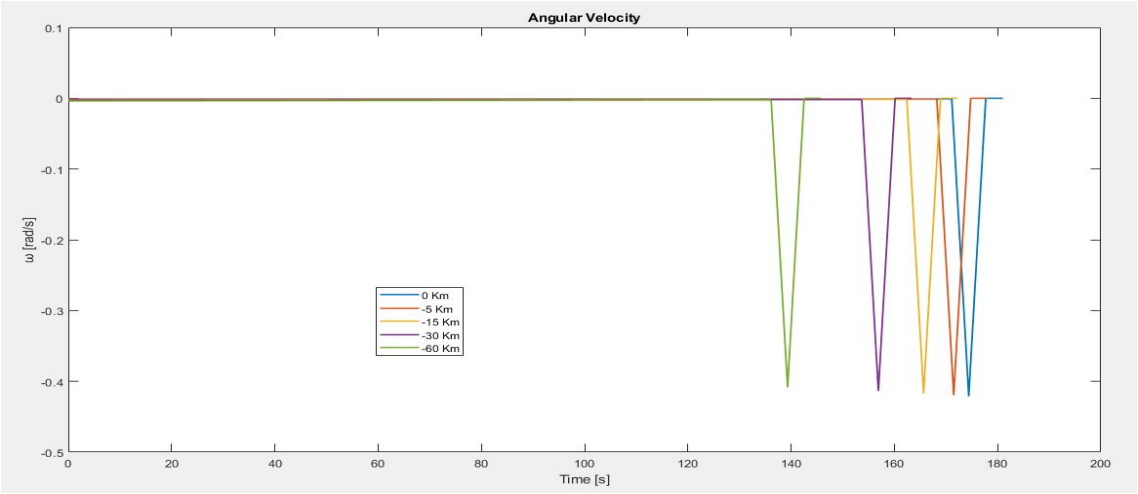


Figure 7.4.39: Angular Velocities for decreasing landing site positions

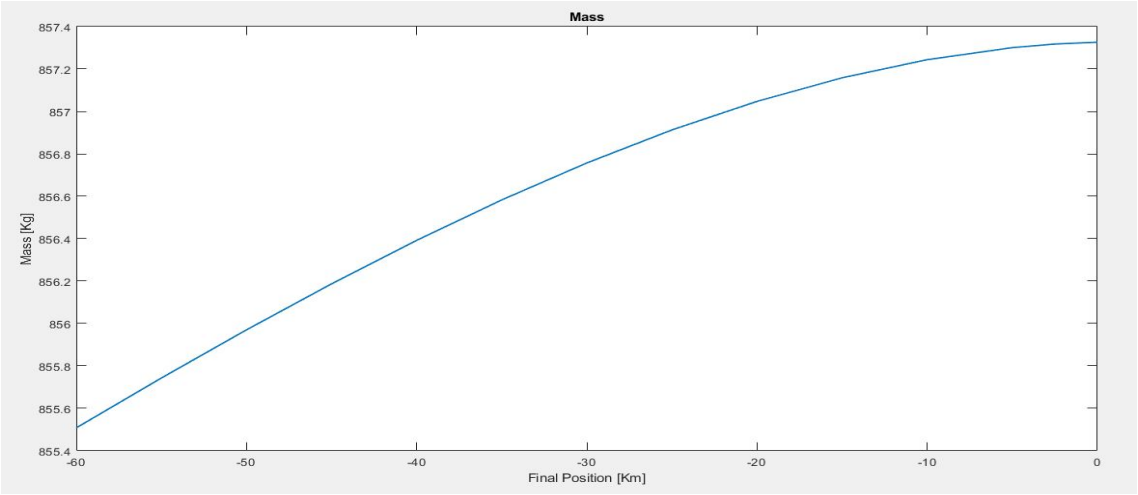


Figure 7.4.40: Mass in function of decreasing landing site positions

On the other hand, the angular speed with which the lander starts the maneuver gradually takes on larger values in module, allowing the vehicle to reach the end of the singular arc with a greater pitch angle. It is seen that for this reason, there are no major variations in the trend of the $V_x(t)$ curves. Moreover, this translates into shorter times for the second and third phases. In particular, it is noted that the difference between their durations increases in order to cancel out an always greater final ω. Instead, the overall duration of the maneuver decreases, but since the propulsion system is activated earlier, the propellant consumption associated with the trajectory increases. As a result, the mass with which the lander can land decreases, but only to a negligible extent. In fact, it can be seen from the Table 7.4.4 that landing 30 kilometers earlier requires approximately 600 grams more fuel. It can be concluded that this maneuver increases the efficiency of the landing compared to the case with a thrust of 4000 Newtons, as seen, while still leaving the spacecraft with a good safety margin for landing.

Chapter 8

Conclusions

For centuries, the Moon has fascinated us, and now humanity is ready to return with NASA's Artemis program, with the aim of staying there. Colonizing our satellite would not only be a globally impactful achievement, but it would also allow for up-close and in-depth study of the early stages that characterized the origin of the Solar System, while laying the groundwork for future planetary missions. As a result, lunar surface missions will increase in frequency in the coming years to support the Artemis Base Camp. For this reason, this thesis proposes a guidance algorithm that can efficiently and safely conduct the soft-landing maneuver, taking into account stringent constraints and a strategy for optimizing the spacecraft's final mass. Minimizing propellant consumption is crucial to maximize the payload mass to be carried on the surface, as well as to reduce the costs associated with each mission. To apply such a strategy, it was necessary to introduce the fundamental concepts of optimal control theory and apply them to the expected mission scenario, which is based on the mission architecture followed by the Apollo missions. Thus, by first defining the system dynamics and maneuver constraints, the optimization problem was solved using an indirect method to achieve a high degree of numerical accuracy and reduce computation time. By applying Pontryagin's Maximum Principle to the optimal control problem, a Boundary Value Problem formulation was obtained, which was solved using shooting techniques. At this point, the code was applied to obtain the optimal soft-landing trajectory at a specific point on the lunar surface for a reference case, starting from a certain thrust level and leaving the downrange and initial attitude of the lander free. From this study, better solutions were investigated by increasing the thrust in accordance with Pontryagin's Maximum Principle. It was observed that the highest final mass is obtained for T=10800 Newton. For higher thrust levels, however, the constraints imposed in the final phase of the trajectory to achieve a vertical descent and soft-landing result in a relative deterioration of performance. Nevertheless, it was possible to deduce that for these thrust levels, the guidance algorithm provides for an initial coasting phase with the thrusters off, in accordance with the Bang-Bang control applied to the thrust. These results were useful in the second part of the downrange analysis. In this case, in order to account for possible contingencies and ensure the possibility of real-time trajectory recalculation, a safety margin was sought for the mission, lengthening and shortening the downrange from a certain initial position. Increasing the distance of the landing site did not pose any problems in finding the optimal solution, and it was observed that, at a small additional cost, the trajectory could be correctly adjusted. On the other hand, reducing

the downrange did not provide enough time to break and, consequently, to avoid a hard landing or a crash of the lander on the surface in the final phases. For this reason, it was decided to exploit the initial coasting phase obtained by adopting a high thrust value. In this way, to reach a closer landing site, the lander can anticipate the ignition of the propulsion system in order to properly decelerate.

Although the model presented in this work adopts several simplifications, it still represents a good method for addressing the optimization problem for lunar soft-landing trajectories. Moreover, it can be a useful starting point for entry missions to Mars or asteroids with appropriate modifications. Since the code can be easily modified, it could be interesting to implement a six-degree-of-freedom dynamic system that also takes into account crossrange motion and rotational motions around the yaw and pitch axes. In this way, it would be possible to constrain the vehicle orientation to specified points of the trajectory in order to properly manage the visibility, taking into account, for example, possible tracking cameras for the Hazard Detection and Avoidance system. Furthermore, it would be interesting to improve the results obtained in this study, for instance by modulating the thrust in the final phases of the trajectory as the Lunar Module did in the Apollo missions. Alternatively, by adopting high levels of thrust, it would be possible to add a singular arc for the thrust at the beginning of the maneuver, increasing the size of the problem and investigating new possible constraints. In this way, in case of hazards, it would be possible to reduce the downrange even for lower thrust levels.

Bibliography

[1] Siddiqi, Asif A. (2018). *Beyond Earth: A Chronicle of Deep Space Exploration, 1958–2016*, Washington, D.C, NASA history series (2nd edition)

[2] NASA (1966). *Apollo Lunar Landing Mission Symposium*, Manned Spacecraft Center Houston, Texas

[3] NASA (1969). *Apollo 11 Fligh Plan*

[4] NASA (2020). *NASA's Lunar Exploration Program Overview*

[5] Floyd V, Bennett (1972). *Apollo Experience Report - Mission Planning for Lunar Module Descent and Ascent*, Manned Spacecraft Center Houston, Texas

[6] Jeff Delaune, Diego De Rosa Stephen Hobbs (2018). *Guidance and Control system design for Lunar Descent and Landing*, Toronto, Ontario Canada, AIAA Guidance, Navigation, and Control Conference

[7] Ingo Gerth , Erwin Mooij (2014). *Guidance for Autonomous Precision Landing on Atmosphereless Bodies*, National Harbor, Maryland, AIAA Guidance, Navigation, and Control Conference

[8] Donald E. Kirk (1998). *Optimal Control Theory- An Introduction*, San José, California, (1st edition)

[9] John T. Betts (1998). *Survey of Numerical Methods for Trajectory Optimization*, Washington, D.C, JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS Vol. 21, No. 2

[10] Daniel Liberzon (2012). *Calculus of Variations and Optimal Control Theory - A Concise Introduction*, Princeton, New Jersey, Princeton University Press (1st edition)

[11] D. G. Hull (2003). *Optimal Control Theory for Applications*, New York, NY

[12] Arthur E. Bryson, Yu-Chi Ho (1975). *Applied Optimal Control*, New York, NY, (1st edition)

[13] William E. Boyce, Richard C. DiPrima (2017). *Elementary Differential Equations and Boundary Value Problems*, Hoboken, New Jersey

[14] Guido Colasurdo, Lorenzo Casalino (2013). *Indirect Methods for the Optimization of Spacecraft Trajectories*, New York, NY, Modeling and Optimization in Space Engineering

[15] L. S. Pontryagin (1962). *Mathematical Theory of Optimal Processes*, New York, NY

[16] D. F. Lawden (1963). *Optimal Trajectories for Space Navigation*, Butterworths, London

[17] Bruce A. Conway (2010). *Spacecraft Trajectory Optimization*, Cambridge, England, Cambridge Aerospace Series