AN ALTERNATIVE SOLUTION METHOD FOR TRAJECTORY

OPTIMIZATION FOR MARTIAN DESCENT AND LANDING

by

Kyle Wianecki, B.S.

A thesis submitted to the Graduate Council of
Texas State University in partial fulfillment
of the requirements for the degree of
Master of Science
with a Major in Mathematics
May 2021

Committee Members:

Young Ju Lee, Chair

Gregory Passty

Raymond Treinen

FAIR USE AND AUTHOR'S PERMISSION STATEMENT

Fair Use

Duplication Permission

ACKNOWLEDGMENTS

To all of my friends, family, and loved ones who have helped me see this through to completion. Thank you.

TABLE OF CONTENTS

Page

CHAPTER

LIST OF FIGURES

I. INTRODUCTION

In this era of rapid technological advancement, space exploration is beginning to move toward autonomous missions; numerical algorithms and computation now govern our space vehicles. This rise in automation can be seen in a range of areas from thrust control, orbit determination, orbital maneuvers, to atmospheric re-entry, descent and landing. At present, one can witness such automation by pointing their attention towards the private American aerospace company Space Exploration and their fully automated vertical takeoff, vertical landing (VTVL) Falcon 9 rocket. This partially reusable rocket is capable of delivering cargo and humans safely to orbit and of re-entering the atmosphere and softly landing on the surface of the planet [10].

Another recent example of autonomous missions is the successful landing of NASA's Perseverance rover on the surface of Mars. Following the re-entry into the Martian atmosphere and the deployment of a parachute to reduce its descent velocity, the vehicle uses terrain relative navigation to determine the vehicle's location above the martian surface followed by the on-board computation to determine the optimal trajectory to the desired landing location. Great attention is required in the computation of such an optimal trajectory. The requirements of the mission and the satisfaction of all the physical constraints and bounds can lead to infeasible solutions. The speed and accuracy at which the computation is executed can also present complications, particularly when one is dealing with on-board, real-time calculations.

The focus of this thesis is in developing a unique robust method for solving the martian descent and landing algorithm under bounding constraints. This consists of simultaneously minimizing the space vehicle's fuel consumption and the perturbation error incurred in its distance from the targeted landing location. The problem contains nonconvex thrust bounding constraints and is governed

by nonlinear dynamics. We propose the use of the Moreau-Yosida regularization technique to simplify the state constraints and using Lagrange Multipliers to handle the control constraints. Additionally, the problem is presented as a decoupled final-time minimization problem where first the problem is minimized over the state and controls followed by using a line search algorithm to minimize over the final time.

The history of landing on Mars can be dated back to the Viking mission in 1976 where it was the first time a spacecraft was placed safely on Mars [4]. However the historical background relevant to the research contained in this thesis dates back to 2007 to the work done by Behçet Açikmeşe and Scott Ploen in their paper "Convex Programming Approach to Powered Descent Guidance for Mars Landing" [2]. Here the authors present an algorithm for solving the minimum fuel landing problem with a linear programming technique, particularly a second order cone programming approach. A major contribution in this paper is a proof that the non-convex constraint, $0 < \rho_1 \leq T(t) \leq \rho_2$, on the bounds of the thrust of the vehicle can be convexified and the solution to the convex version is equivalent to the solution of the nonconvex problem. The authors' methodology is to convexify the problem, reformulate the problem by simplifying the pieces of non-linearity, discretize the problem into a finite dimensional second-order cone problem and solve it using built in MATLAB solvers.

Other sources with heavy contributions to the framework of this thesis include the 2010 and 2008 papers by Lars Blackmore, Behçet Açikmeşe, and Daniel Scharf titled "Minimum-Landing-Error Powered-Descent Guidance for Mars Landing Using Convex Optimization"[3] and "Enhancements on the Convex Programming Based Powered Descent Guidance Algorithm for Mars Landing"[1]. In these papers the authors introduce the minimum landing error cost functional used to minimize the error in reaching the target location. In building upon the

work done in 2007, the authors present the minimum landing error problem as a finite dimensional second order cone problem which can be solved in the same manner as done in the work from 2007. In the 2010 paper the authors introduce a novel approach which consists of solving a prioritized algorithm that first solves the minimum fuel landing algorithm for the optimal final time, states and controls, followed next by solving the minimum landing error algorithm for its subsequent optimal final time, state and controls.

We reference the work done by John Pearson in his dissertation from the University of Oxford on "Fast Iterative Solvers for PDE-Constrained Optimization Problems" [12] for our implementation. We make use of his algorithms involving the Moreau-Yosida regularization technique and Lagrange multipliers to solve PDE-constrained problems. We adapt his methodology to solve our problem which contains much more challenging compound inequality constraints and nonlinear conic constraints.

Motivated by these current advancements in solution techniques by Blackmore, Açikmeşe, Scharf and others, we believe that our model will provide an even faster and more stable algorithm that can be solved in less time and with fewer iterations than previous methods. Our procedure is as follows. We begin with the problem formulation by introducing the governing physics, particularly the dynamic equation governing the position of the vehicle, and the dynamic equation governing the rate at which the mass of the rocket changes over time. Next is the determination of what constraints we should impose on the model. The constraints include upper and lower bounds on the thrust of the vehicle, a conic constraint to restrict the space in which the rocket can travel, a lower bound on the mass of the vehicle at the final time, a bound that keeps the vehicle from traveling subsurface, and the requirement that the position and the velocity of the vehicle must simultaneously reach zero. Next we alter the relaxation of the nonconvex

thrust constraint proven in 2007 [2]. Because the relaxation is feasible for values on the boundary of the thrust bounds, we introduce a squared slack variable. We introduce another change of variables needed to reformulate the original L^1 cost function for the minimization of the fuel as a squared L^2 cost function. From here we discretize our dynamical equations into a recursive Euler linear-time invariant equation with discretizations applied over a bounded interval $[0, t_f]$. We regularize the problem with the Moreau-Yosida regularization parameter and the introduction of Lagrange multipliers. The problem is then decoupled into two minimization problems, one minimizing over the state and the controls and the other minimizing over the final time. In [2] the authors derive upper and lower bounds on the final time t_f that further reduces the difficulty in determining the optimal final time. Those bounds are used in this thesis and the resulting novel formulation is solved using the quasi-Newton Broyden method.

Nomenclature

- t_0, t_f are the initial and final time, and $h = \Delta t$.

- $\bar{t} = t - t_f$ and t_k is a time step $\in [0, t_f]$ $k = 0, \ldots, N$.

- $m(t)$ is the mass function, $z(t) = \ln m(t)$.

- $\underset{\sim}{T}(t)$ is the thrust equation.

- $\Gamma^2(t)$ is the slack variable for the thrust.

- $\underset{\sim}{u}(t)$ and $\sigma^2(t)$ are the mass normalized thrust and slack functions.

- n is the number of engines.

- α is the rate at which fuel is consumed.

- λ is the Tikhonov regularization parameter.

- $\epsilon_1, \epsilon_2, \epsilon_3$ are Moreau-Yosida regularization parameters.

- $\underset{\sim}{g}$ is the planetary gravitational vector.

- θ is the cant angle on the thrusters.

- γ is the path bounding angle.

- ρ_1, ρ_2 are the lower bound and upper bound on the thrust.

- m_{wet} is the initial mass of the vehicle which includes the mass of the fuel.

- m_{dry} is the mass of the vehicle without the mass of the fuel.

- $\underset{\sim}{e}$ is a unit vector in $\mathbb{R}^3$ and $\underset{\approx}{A}$ is a matrix in $\mathbb{R}^{3 \times 3}$.

- $\|\cdot\|$ is the standard L_2 norm.

II.   PRELIMINARY WORK

Throughout this paper we assume that the force of gravity acting on the vehicle is uniformly distributed. We will also treat outside forces that affect the vehicle's aerodynamics as perturbations. Considering that the landing phase of a space vehicle takes place at low altitudes and at speeds much slower than atmospheric re-entry, this is a reasonable decision which also leads to a reduction in the problem's difficulty but has no loss in the solution's accuracy.

Here we begin to describe and formulate the optimal trajectory problem for planetary landing of space vehicles . Let $\underset{\sim}{r}(t) = [r_1(t)\, r_2(t)\, r_3(t)]^T \in \mathbb{R}^3$ be the vehicle's position vector at time t. This model is situated in a surface fixed reference frame where $r_1(t)$ is the outward facing normal vector i.e. the altitude. Take $\underset{\sim}{g} \in \mathbb{R}^3$ to be the planetary gravitational constant, let $\underset{\sim}{T}(t) \in \mathbb{R}^3$ be the vehicle's thrust profile, and let $m(t)$ be the mass of the vehicle as a function of time. The vehicle's flight dynamics are expressed as the following dynamic equations:

$$\ddot{\underset{\sim}{r}}(t) \quad = \quad \underset{\sim}{g} + \frac{\underset{\sim}{T}(t)}{m(t)} \tag{1a}$$

$$\dot{m}(t) \quad = \quad -\alpha \left\| (\underset{\sim}{T}(t)) \right\| \tag{1b}$$

where α is the rate at which the fuel is being exhausted:

$$\alpha = \frac{1}{gI_{sp}\cos(\theta)} \tag{2}$$

where I_{sp} is the specific impulse. The thrust equation is given as

$$\underset{\sim}{T}(t) = n\underset{\sim}{e}^T T \cos(\theta) \tag{3}$$

where $\underset{\sim}{e}^T$ is the canonical unit vector pointed in the direction of the thrust, and

6

θ is the thrusters cant angle, or the degree for which the thrusters are angled with respect to the body of the space vehicle. The number of thrusters is n, each assumed to contribute an equivalent amount of thrust T. Figure 1 provides an example of a generic thrust profile throughout the duration of the flight.
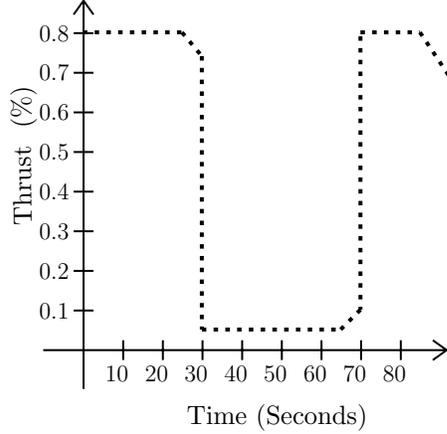
Figure 1.: Generic thrust profile.

In this framework the vehicle is represented as a dimensionless point rather than a geometric object and the rotational and translational dynamics are treated as separate. It is common practice to treat rotational and translational dynamics as separate because any translational maneuver can be done with such speed that any interaction with the attitude control is negligible.

It is assumed that once a vehicle's thrusters are ignited they cannot be extinguished, this is due to the safety concern that if an engine is extinguished, there is no guarantee of re-ignition. This places a lower bound on the thrust throughout the vehicle's flight. Denote T_{max} and T_{min} to be the respective allowable maximum thrust and minimum thrust. This gives rise to the following inequalities:

$$0 < \rho_1 = n \cos(\theta) T_{min} \le \|(T(t))\| \le n \cos(\theta) T_{max} = \rho_2. \tag{4}$$

Many landing models require that the trajectory of a vehicle have constraints imposed on the path the vehicle can take, whether it is due to the need for the vehicle to maintain a low velocity (a strong requirement for manned vehicles) or for it to avoid certain obstacles. There are several ways to express this constraint; the choice made in this paper uses a cone to enforce the requirement that the altitude of the rocket be constrained above the surface at all times during the descent and landing [5]. The cone constraint is:

$$\tan^2(\gamma)\|\underset{\approx}{E}_r \underset{\sim}{r}(t)\|^2 - \underset{\sim}{e}_1^T \underset{\sim}{r}(t)^2 \le 0. \tag{5}$$

The matrix $\underset{\approx}{E}_r$ is a projection matrix accessing the last two components of the position vector $\underset{\sim}{r}$. This matrix is give as

$$\underset{\approx}{E}_r = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6}$$

with $\tan(\gamma) > 0$ and $0 \le \gamma < \frac{\pi}{2}$ is the path bounding angle. With these constraints together, Figure 2 below depicts the geometry of the model. You may notice that simplification of the squares present in each term would provide an equivalent representation of the required constraint. The point in squaring all of the terms comes to light in the discretization section, where a finite element approximation will be applied to the continuous functions, and the square of this standard euclidean norm can be expressed as a simple quadratic expression. It should also be noted that the meaning of $\underset{\sim}{e}_1^T \underset{\sim}{r}(t)^2$ represents the component-wise squaring of vectors. Meaning,

$$\underset{\sim}{e}_1^T \underset{\sim}{r}(t)^2 = (\underset{\sim}{e}_1^T \underset{\sim}{r}(t))^2 = (\underset{\sim}{e}_1^T \underset{\sim}{r}(t))^T (\underset{\sim}{e}_1^T \underset{\sim}{r}(t)) = \underset{\sim}{e}_1^T (\underset{\sim}{r}(t))^T \underset{\sim}{r}(t))$$

This same notational meaning will hold throughout this thesis.

Another constraint is added to represent the lower bound physically present on the mass, $m_{t_f} \geq m_{dry}$, where m_{dry} is the mass of the vehicle without any fuel. This constraint follows from the finite availability of fuel in the vehicle. In [2], the following compound inequality

$$(m_{wet} - t\alpha\rho_2) \leq m(t) \leq (m_{wet} - t\alpha\rho_1)$$

is introduced to place naturally occurring physical bounds on the space vehicle's mass. This is equivalent to the constraint $m(t_f) \geq m_{dry}$. To see this, notice that the left-hand functions and the right hand functions are both decreasing. Because $\rho_2 > 0$ represents the maximum upper bound on the thrust, and $\rho_1 > 0$ is the lower bound, $m_{wet} - t\alpha\rho_2$ approaches m_{dry} at a rate faster than $m_{wet} - t\alpha\rho_1$ for all $t \in [0, t_f]$. Since $m(t)$ is a decreasing function it will *always* be decreasing at a rate equal to or less than $m_{wet} - t\alpha\rho_1$ thus there is no reason to require it as an upper bound because the only way $m(t)$ could decrease at a slower rate than $m_{wet} - t\alpha\rho_1$ is if the thrust is less than ρ_1 which violates the constraint bounding the thrust. Now, to see the equivalency between $m(t_f) \geq m_{dry}$ and $m_{wet} - t\alpha\rho_2 \leq m(t)$ notice that m_{dry} can be determined by some $t_* \in [0, t_f]$ such that $m_{wet} - t_*\alpha\rho_2 = m_{dry}$. Suppose that $t_* = 0$ then the equivalency holds trivially. If $0 < t_* < t_f$ then for the time intervals $(t_*, t_f]$ the problem is infeasible. If $t_* = t_f$ then the equivalency between the inequalities holds trivially.

It should be assumed that $\rho_1 \leq \left\| m_{dry}\underset{\sim}{g} \right\|$ and $\left\| m_0\underset{\sim}{g} \right\| \leq \rho_2$. It does not represent any additional constraints, but is a physical requirement for soft landing. It is also always satisfied for feasible solutions to the soft landing problem. Suppose that the lower bound on the thrust magnitude was less than the gravitational acceleration. If at anytime during the landing phase the thrusters transition to the lower bound magnitude, the vehicle will begin to accelerate. In turn this will lead to infeasible landing solutions, typically subsurface flight. Similarly, if the vehicle is,

at the outset, traveling at a speed greater than the maximum thrust, the trajectory again will result in subsurface flight because the vehicle will not decelerate.
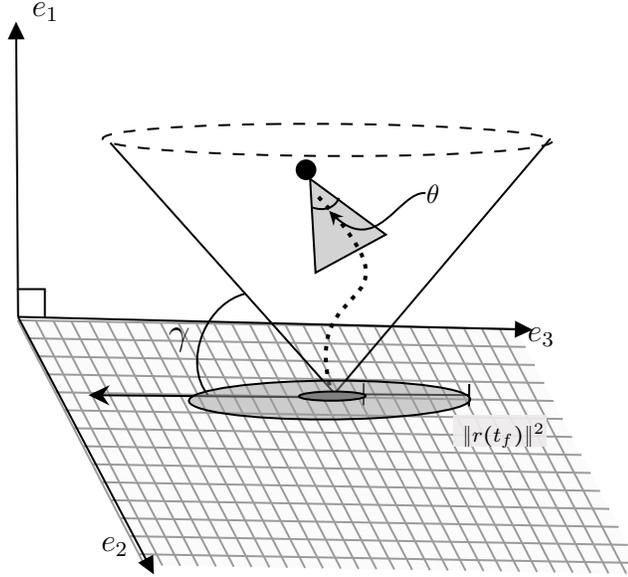
Figure 2.: Diagram with glide slope and thrust pointing constraint.

Problem Statement

Presented in [3], the planetary landing problem is stated as the following two optimization problems. That notation and structure are stated differently and not verbatim from their publication but has been restated to match the notation of this thesis. It is possible for there to exist multiple solutions that minimize the minimum fuel problem, hence the requirement for an additional step of calculating the minimum landing error solution that also minimizes the distance of the optimal landing location and the targeted landing location. This first problem is the Nonconvex Minimum Landing Error Problem (MLEP). The algorithm is sourced from [3].

Problem 1. $\min\limits_{t_f} \min\limits_{\rho_1 \leq \|\underset{\sim}{T}(t)\| \leq \rho_2} \|\underset{\sim}{r}(t_f)\|^2$

subject to:

$$\begin{cases} \ddot{r}(t) = g + \dfrac{T(t)}{m(t)}, & r(0) = r_0, \quad \dot{r}(0) = \dot{r}_0, \quad r_1(t_f) = \dot{r}(t_f) = 0, \\[2mm] \quad \tan^2(\gamma)\|E_r(r(t))\|^2 - e_1^T r(t)^2 \le 0, & \forall t \in [0, t_f] \end{cases}$$

$$\begin{cases} \dot{m}(t) = -\alpha\,\|(T(t))\|, & m(0) = m_{wet}, \quad m(t_f) \ge m_{dry} \end{cases}$$

The next algorithm represents the Nonconvex Minimum Fuel Problem (MFP). The only point of difference between the MFP from the MLEP is in the cost functional. Particularly, the MFP is concerned with the following sourced directly from [2, 3]:

Problem 2. $\displaystyle\min_{t_f}\ \min_{\rho_1 \le \|T(t)\| \le \rho_2}\ \int_0^{t_f} \alpha\,\|T(t)\|\,dt$

subject to:

$$\begin{cases} \ddot{r}(t) = g + \dfrac{T(t)}{m(t)}, & r(0) = r_0, \quad \dot{r}(0) = \dot{r}_0, \quad r_1(t_f) = \dot{r}(t_f) = 0, \\[2mm] \quad \tan^2(\gamma)\|E_r(r(t))\|^2 - e_1^T r(t)^2 \le 0, & \forall t \in [0, t_f] \end{cases}$$

$$\begin{cases} \dot{m}(t) = -\alpha\,\|(T(t))\|, & m(0) = m_{wet}, \quad m(t_f) \ge m_{dry} \end{cases}$$

Relaxation of Nonconvex Constraints

The first action in solving this problem is to begin by relaxing the constraints of the nonconvex optimal control problem In [2], the authors prove that the thrust bounding inequality constraint can be convexified with no loss in solution accuracy, this fact is very useful in our alternative approach, so the work done by Behçet Açikmeşe and Scott Ploen in their paper "Convex Programming Approach to Powered Descent Guidance for Mars Landing" [2], will be summarized and the techniques used to relax the nonconvex problem as well as the author's discretization techniques will also be presented here.

The authors relax the MFP by introducing a slack function $\Gamma(t)$ that replaces $\|(T(t))\|$. Consequently, this imposes the constraint

$$\|(T(t))\| \le \Gamma(t). \tag{7}$$

The main result of [2] is Lemma 1, which is the proof that states that if there exists a solution to the convex MFP, then there also exists a solution to the nonconvex MFP and it can be obtained directly from the solution to the convex problem. The question about the existence of an optimal solution to the relaxed MFP is also proven in [2].

The modified nonconvex MFP that includes the relaxation of the control magnitude constraint:

Problem 3. $\min\limits_{t_f} \min\limits_{T,\Gamma} \int_0^{t_f} \Gamma(t)dt$

subject to:

$$
\begin{cases}
\ddot{\underline{r}}(t) = \underline{g} + \dfrac{\underline{T}(t)}{m(t)}, \quad \underline{r}(0) = \underline{r}_0, \quad \dot{\underline{r}}(0) = \dot{\underline{r}}_0, \quad \underline{r}_1(t_f) = \dot{\underline{r}}(t_f) = 0, \\[2mm]
\tan^2(\gamma)\|\underline{\underline{E}}_r(\underline{r}(t))\|^2 - \underline{e}_1^T \underline{r}(t)^2 \le 0, \quad \forall t \in [0, t_f]
\end{cases}
$$

$$
\begin{cases}
\dot{m}(t) = -\alpha\Gamma(t), \quad m(0) = m_{wet}, \quad m(t_f) \ge m_{dry}
\end{cases}
$$

$$
\begin{cases}
\|\underline{T}(t)\| \le \Gamma(t), \quad \rho_1 \le \Gamma(t) \le \rho_2 \quad \forall t \in [0, t_f]
\end{cases}
$$

Nondimensionalization

After the convexification of the control magnitude constraint, $\|(\underline{T}(t))\|$, the authors proceed to perform a change of variables. The purpose is the standard process of nondimensionalization. Change of variables:

$$
\sigma(t) \triangleq \frac{\Gamma(t)}{m(t)}, \quad \underline{u}(t) \triangleq \frac{\underline{T}(t)}{m(t)} \tag{8a}
$$

By the change of variables, the additional constraint imposed after the relaxation of the thrust constraint becomes

$$
\|\underline{u}(t)\| \le \sigma(t), \quad \forall t \in [0, t_f],
$$

12

and the lower and upper bounds on the thrust profile become

$$\frac{\rho_1}{m(t)} \le \sigma(t) \le \frac{\rho_2}{m(t)} \quad \forall t \in [0, t_f].$$

Notice that $\frac{1}{m}$ now posses an additional point of nonconvexity. In particular $\sigma(t)$ and $\underset{\sim}{u}(t)$ defined above are bi-linear functions because the treatment of $m(t)$ is that of a problem variable. To handle this case of minor nonconvexity the authors of [2] introduce the following change of variable: $z(t) \triangleq \ln m(t)$. Now the differential equation for $\dot{m}(t)$ is reformulated as:

$$\frac{\dot{m}(t)}{m(t)} = -\alpha\sigma(t) \implies \dot{z}(t) = -\alpha\sigma(t)$$

This substitution physically poses no issue seeing as the mass of the vehicle is always much larger than zero. Since $\alpha > 0$ is the rate of fuel being consumed, minimizing the amount of fuel used is the same as minimizing this integral equation

$$\int_0^{t_f} \sigma(t)dt.$$

The inequality constraints on the control input has a lower and upper bound that can be stated as:

$$\rho_1 e^{-z(t)} \le \sigma(t) \le e^{-z(t)}\rho_2 \quad \forall t \in [0, t_f] \tag{9}$$

The authors of [2, 3] rewrite (9) involving a Taylor expansion. Particularly, for the lower thrust bound the authors use the first three terms of the Taylor expansion and the first two terms of the Taylor expansion for the upper thrust bound. Define

$$\mu_1 = \rho_1 e^{-z(t)}, \quad \mu_2 = \rho_2 e^{-z(t)} \tag{10}$$

such that

$$\mu_1(t)\left[1-(z(t)-z_0(t))+\frac{(z(t)-z_0(t))^2}{2}\right]\le\sigma(t)\le\mu_2(t)\left[1-(z(t)-z_0(t))\right]$$

for all time steps t in $[0,t_f]$, where the initial value of $z(t)$ is

$$z_0(t)=\ln\left(m_{wet}-\alpha\rho_2 t\right).$$

The following optimization problem represents the Relaxed MFP with the necessary change of variables:

Problem 4. $\min\limits_{t_f}\min\limits_{T,\Gamma}\int_0^{t_f}\Gamma(t)dt$

subject to:

$$\begin{cases}\ddot{\underline{r}}(t)=\underline{g}+\underline{u}(t),\quad \underline{r}(0)=\underline{r}_0,\quad \dot{\underline{r}}(0)=\dot{\underline{r}}_0,\quad \underline{r}_1(t_f)=\dot{\underline{r}}(t_f)=0,\\ \qquad\tan^2(\gamma)\|\underline{E}_r(\underline{r}(t))\|^2-\underline{e}_1^T\underline{r}(t)^2\le 0,\quad \forall t\in[0,t_f]\end{cases}$$

$$\begin{cases}\dot{z}(t)=-\alpha\sigma(t),\quad z(0)=\ln m_{wet},\quad z(t_f)\ge\ln m_{dry}\end{cases}$$

$$\begin{cases}\|\underline{u}(t)\|\le\sigma(t)\\ \mu_1(t)\left[1-(z(t)-z_0(t))+\frac{(z(t)-z_0(t))^2}{2}\right]\le\sigma(t)\\ \sigma(t)\le\mu_2(t)\left[1-(z(t)-z_0(t))\right]\end{cases}$$

Discretization and Solution Algorithm

In [2, 3, 5], the authors use a standard discretization process to reduce the problem to a finite dimensional problem by fixing the final time t_f, discretizing it into evenly space intervals and inducing the state and control constraints at each node, where N represents the number of nodes. For an in-depth look at the details of previous standard discretizations, the reader is referred to [2], otherwise the reader is directed to an alternative discretization derivation which is carried out in the proceeding chapter.

14

There are known upper and lower bounds that can be placed on the final

time t_f. These bounds are presented in [2]. In our alternative approach these

bounds will also be relevant.

$$\frac{m_{wet} - m_{dry} \|\dot{\underline{r}}(0)\|}{\rho_2} \leq t_f \leq \frac{m_{fuel}}{\alpha \rho_1}. \tag{11}$$

With these known bounds, the authors of [2, 3, 1]argue that the problem expressed

as a decoupled minimization problem can be solved by solving the first the problem

is minimized over the state and controls, followed by a line search algorithm used

to minimize over the final time, t_f. In [2], the authors experimentally prove that

the first cost functional $\int_0^{t_f} \sigma(t)dt$ has only one minimizer. The authors of [3] prove

experimentally that the cost functional $\|\underline{r}(t_f)\|_2^2$ contains only one minimizer. In

both cases the authors of [2] and [3] make use of the unimodality and the new

upper and lower bounds on the final time to implement the golden line search

algorithm to their respective cost functionals to minimize over the final time.

The method for obtaining a numerical solution to the minimum-fuel landing

problem, presented in [2, 3, 1] begins by defining a space that consists of all the

feasible thrust values that satisfy all of the aforementioned constraints for every for

$t_k \in [0, t_f]$. Call this set $\mathbb{F}$. This more commonly can be thought of as the active

set, or within the framework of this problem, the set of all σ values that live inside

the feasible region defined by the given state and control constraints. If the set $\mathbb{F}$

is empty then the minimum fuel problem has no feasible solution where the issue

arises in the thrust. If the set is not empty, solve $f(t_f) = \min\limits_{\sigma \in \mathbb{F}} \int_0^{t_f} \sigma(t)dt$. Now

apply a line search on t_f over $f(t_f)$ to generate the optimal pair (t_f^*, σ^*). Finally,

perform a final validity test by checking if the optimal pair (t_f^*, σ^*) satisfy the

fuel availability constraint, that is some form of $m(t_f) \geq m_{dry}$. If the constraint

is satisfied the the pair (t_f^*, σ^*) is a feasible solution. Otherwise if the constraint

is not satisfied, the minimum fuel problem has no feasible solution caused by a

limitation in the fuel.

A similar approach is taken to solve the minimum-fuel landing problem, for which the authors of [3] develop an algorithm to solve the minimum-fuel and minimum-landing error problems in a prioritized fashion. Begin by performing the method as outlined above the generate a feasible optimal solution pair for the MFP (t_f^*, σ^*). If no solution exists, terminate the process. Else, take the value t_f^* and now solve the MLEP using the same outlined method for the cost functional defined as $\left\| \underset{\sim}{r}(t_f^*) \right\|_2^2$ to generate an optimal solution pair $(t_f^\dagger, \sigma^\dagger)$.
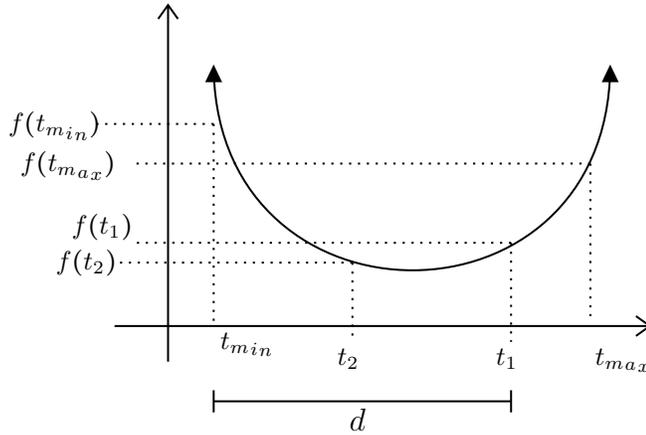
Figure 3.: Simple example of the golden search method.

The software used to simulate the results in [2] is a primal dual path following interior point method called SeDuMi, or Self-Dual Minimization. The software is compatible with a MATLAB interface.

Remark: The golden search method is an iterative method used to find the minimizer of a function where an upper and lower bound must be known for what you are optimizing over. The golden search method works by using the golden ratio $\phi = \frac{1}{2}(\sqrt{5} - 1)$ to determine a value denoted $d = \phi(t_{min} - t_{max})$. Now, consider for some function $f(t)$, where $t_1 = t_{min} + d$ and $t_2 = t_{max} - d$, as depicted in Figure 3, such that $t_{min} < t_2 < t_1 < t_{max}$. These values t_2, t_1 corresponds to output values $f(t_2), f(t_1)$ and if $f(t_2) < f(t_1)$ then $t_2 \triangleq t_{min}$ and $t_1 \triangleq t_2$ with no change to t_{max}.

Finding the updates for the case when $f(t_1) < f(t_2)$ follows mutatis mutandis. This process is repeated until convergence is achieved. One attributable benefit to such a line search method is the main to requirements are (1) upper and lower bounds and (2) a unimodal cost function.

III. A NOVEL APPROACH

A point of separation between the work stated above and the work of this thesis is with the handling of the slack function $\Gamma(t)$, and with the inequality constraints on the thrust bounds. We opt to square the right-hand side to make for better implementability later down the line.

$$\|\underset{\sim}{T}(t)\| \leq \Gamma^2(t). \tag{12}$$

The typical nature of the thrust function is one that is max-min-max. This follows from the application of Pontrygan's maximum principle. The proof of such a claim is omitted here but can be found in the main result of [2]. This very useful fact is used in justifying that the convexificaiton term is lossless. Additionally by Pontrygan, the points of concern regarding the validity of a solution exists on the boundary. Lemma 2 in [2] proves that the convexification is lossless. The lemma is stated as follows and the reader is referred to [2] for the poof.

Lemma 5. *Consider a solution of Problem 3 given by $[t_f^*, \underset{\sim}{T}^*(\cdot), \Gamma^*(\cdot)]$. Then $[t_f^*, T^*(\cdot)]$ is also a solution to Problem 1 and $\|\underset{\sim}{T}^*(t)\| = \rho_1$ or $\|\underset{\sim}{T}^*(t)\| = \rho_2$ for $t \in [0, t_f^*]$.*

Therefore, because the convexification is lossless and that any solution to the trajectory optimization problem involving the slack variable Γ^2 is also a solution to original non-convexified problem with $\|\underset{\sim}{T}\|$, *even on the boundary*, rewriting this constraint as an inequality constraint poses no concern.

This new convexification affects the inequality constraint on the upper and lower bounds on the thrust magnitude and appears in the dynamical equation modeling the mass. All other constraints remain the same. Careful scrutiny is required with the change of variables. The consequences of this relaxation has little

effect on the constraints seen in [2] except that the convex representation of the thrust bounds takes the form

$$\rho_1 \le \Gamma^2(t) \le \rho_2 \quad \forall t \in [0, t_f]. \tag{13}$$

As an alternative to computing the optimization problems in a prioritized but separate fashion as done in [3], we propose the introduction of $\lambda > 0$, the regularization parameter commonly known as the Tikhonov parameter. When $\lambda \gg 0$, emphasis is placed on minimizing the MLEP. When λ is small, emphasis is placed on minimizing the fuel. Our relaxation of the nonconvex formulation results in the following optimization problem.

Problem 6. $\min\limits_{t_f} \min\limits_{\Gamma, r} \int_0^{t_f} \Gamma^2(t)dt + \lambda \|\underset{\sim}{r}(t_f)\|_2^2$

subject to:

$$\begin{cases} \underset{\sim}{\ddot{r}}(t) = \underset{\sim}{g} + \dfrac{\underset{\sim}{T}(t)}{m(t)}, \quad \underset{\sim}{r}(0) = \underset{\sim}{r}_0, \quad \underset{\sim}{\dot{r}}(0) = \underset{\sim}{\dot{r}}_0, \quad \underset{\sim}{r}_1(t_f) = \underset{\sim}{\dot{r}}(t_f) = 0, \\ \tan^2(\gamma)\|\underset{\approx}{E}_r(\underset{\sim}{r}(t))\|^2 - \underset{\sim}{e}_1^T \underset{\sim}{r}(t)^2 \le 0, \quad \forall t \in [0, t_f] \end{cases}$$

$$\begin{cases} \dot{m}(t) = -\alpha \Gamma^2(t), \quad m(0) = m_{wet}, \quad m(t_f) \ge m_{dry} \end{cases}$$

$$\begin{cases} \|\underset{\sim}{T}(t)\| \le \Gamma^2(t), \quad \rho_1 \le \Gamma^2(t) \le \rho_2 \quad \forall t \in [0, t_f] \end{cases}$$

New Nondimensionalization

In order to implement our chosen fast numerical solver using the Moreau-Yosida regularization, the cost functionals both must be expressed as an L_2 norm. To satisfy this requirement we redefine the normalized slack variable as a square function. Considering Γ and the mass are both always greater than zero, this change of variables poses no issue. Define the following:

$$\sigma^2(t) \triangleq \frac{\Gamma^2(t)}{m(t)}, \quad \underset{\sim}{u}(t) \triangleq \frac{\underset{\sim}{T}(t)}{m(t)}. \tag{14}$$

19

Rewriting the relaxed equality constraint on the thrust profile in terms of these nondimensionalized variables yields:

$$\|\underset{\sim}{u}(t)\| \le \sigma^2(t), \quad \forall t \in [0, t_f]. \tag{15}$$

Recall that $z(t) = \ln m(t)$, and after repeating the steps outlined in the summary of the previous work, the inequality constraint on the thrust magnitude becomes

$$\rho_1 e^{-z(t)} \le \sigma^2(t) \le \rho_2 e^{-z(t)} \quad \forall t \in [0, t_f]. \tag{16}$$

The updated nondimensionalized MLEP can now be stated as the following minimization problem

Problem 7. $\underset{t_f}{min} \ \underset{\sigma, r}{min} \ \int_0^{t_f} \sigma^2(t) dt + \lambda \|\underset{\sim}{r}(t_f)\|_2^2$

subject to:

$$\begin{cases} \ddot{\underset{\sim}{r}}(t) = \underset{\sim}{g} + \underset{\sim}{u}(t), \quad \underset{\sim}{r}(0) = \underset{\sim}{r}_0, \quad \dot{\underset{\sim}{r}}(0) = \dot{\underset{\sim}{r}}_0, \quad \underset{\sim}{r}_1(t_f) = \dot{\underset{\sim}{r}}(t_f) = 0, \\ \tan^2(\gamma)\|\underset{\approx}{E}_r \underset{\sim}{r}(t)\|^2 - \underset{\sim}{e}_1^T \underset{\sim}{r}(t)^2 \le 0, \quad \forall t \in [0, t_f] \end{cases}$$

$$\begin{cases} \dot{z}(t) = -\alpha \sigma^2(t), \quad z(0) = \ln m_{wet}, \quad z(t_f) \ge \ln m_{dry} \end{cases}$$

$$\begin{cases} \|\underset{\sim}{u}(t)\| \le \sigma^2(t), \quad \rho_1 e^{-z(t)} \le \sigma^2(t) \le \rho_2 e^{-z(t)} \quad \forall t \in [0, t_f] \end{cases} \cdot$$

New Discretization

The first step on the discretization process is to convert the second-order differential equation for $\ddot{\underset{\sim}{r}}(t)$ into a system of first-order differential equations by performing a reduction of order. Then the differential equations for the spacecraft dynamics are

$$\begin{aligned} \ddot{\underset{\sim}{r}}(t) &= \underset{\sim}{g} + \underset{\sim}{u}(t) \\ \dot{z}(t) &= -\alpha \sigma^2(t) \end{aligned}$$

. Let $\underset{\sim}{v}(t) = \dot{\underset{\sim}{r}}(t)$ and perform the substitution

$$
\begin{aligned}
\dot{\underset{\sim}{r}}(t) &= \underset{\sim}{v}(t) \\
\dot{\underset{\sim}{v}}(t) &= \underset{\sim}{g} + \underset{\sim}{u}(t) \\
\dot{z}(t) &= -\alpha\sigma^2(t)
\end{aligned}
$$

Now applying the explicit Euler Method where $h = \Delta t$, and $k = [0, \ldots, N]$ represents the number of time steps:

$$
\begin{aligned}
\underset{\sim}{r}_{k+1} - \underset{\sim}{r}_k &= h\underset{\sim}{v}_k \\
\underset{\sim}{v}_{k+1} - \underset{\sim}{v}_k &= h(\underset{\sim}{g} + \underset{\sim}{u}_k) \\
z_{k+1} - z_k &= h(-\alpha\sigma^2(t)).
\end{aligned}
$$

For each of these components $\underset{\sim}{r}, \underset{\sim}{v}, z$, etc., lump all iterates of each variable into a single vector where $\hat{\underset{\sim}{r}}, \hat{\underset{\sim}{v}}, \hat{\underset{\sim}{u}}, \hat{\underset{\sim}{g}}$ are in $\mathbb{R}^{3(N+1)}$, to account for each spacial direction. The vectors $\underset{\sim}{z}, \underset{\sim}{\sigma}$ are in $\mathbb{R}^{N+1}$. For example,

$$
\hat{\underset{\sim}{r}} = \begin{bmatrix} \underset{\sim}{r}_0 & \underset{\sim}{r}_1 & \cdots & \underset{\sim}{r}_N \end{bmatrix}^T, \tag{17a}
$$

$$
\hat{\underset{\sim}{\sigma}} = \begin{bmatrix} \sigma_0^2 & \sigma_1^2 & \cdots & \sigma_N^2 \end{bmatrix}^T. \tag{17b}
$$

With this new representation, each ODE can be rewritten as these matrix equations

$$
\underset{\approx}{A}\hat{\underset{\sim}{r}} - \underset{\approx}{B}\hat{\underset{\sim}{v}} = \underset{\sim}{c}_1 \tag{18a}
$$

$$
\underset{\approx}{A}\hat{\underset{\sim}{v}} - \underset{\approx}{B}\hat{\underset{\sim}{u}} = \underset{\sim}{c}_2 \tag{18b}
$$

$$
\underset{\approx}{D}\hat{\underset{\sim}{z}} - \underset{\approx}{F}\hat{\underset{\sim}{\sigma}} = 0. \tag{18c}
$$

The matrices $\underset{\approx}{A}, \underset{\approx}{B}$ are in $\mathbb{R}^{3N \times 3(N+1)}$. Also, here the matrix $\underset{\approx}{h} \in \mathbb{R}^3$ is a diagonal

21

matrix with scalar value h along the diagonal. These matrices have the structure

$$
\underset{\approx}{A} = \begin{bmatrix} -\underset{\approx}{I} & \underset{\approx}{I} & 0 & \cdots & 0 \\ 0 & -\underset{\approx}{I} & \underset{\approx}{I} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\underset{\approx}{I} & \underset{\approx}{I} \end{bmatrix}, \quad \underset{\approx}{B} = \begin{bmatrix} \underset{\approx}{h} & 0 & 0 & \cdots \\ 0 & \underset{\approx}{h} & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \underset{\approx}{h} \end{bmatrix}. \tag{19a}
$$

Recall that the mass is not vectorized, as the direction the vehicle is traveling does not influence the rate of change of the mass. So $\underset{\approx}{D}, \underset{\approx}{F}$ are in $\mathbb{R}^{N \times N+1}$, given my

$$
\underset{\approx}{D} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix}, \quad \underset{\approx}{F} = \begin{bmatrix} -\alpha h & 0 & 0 & \cdots \\ 0 & -\alpha h & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\alpha h \end{bmatrix}. \tag{20a}
$$

Furthermore, $\underset{\sim}{c}_1$ and $\underset{\sim}{c}_2$ are meticulously chosen to include the initial and final conditions for the respective initial and final time steps. Here $\underset{\sim}{a}, \underset{\sim}{d}$ are the placeholder vectors for the remaining unknown values of $\underset{\sim}{r}, \underset{\sim}{v}$, respectively, for all other time steps. The matrix $\bar{\underset{\approx}{M}}$ is the resulting mass matrix from the finite element discretization addressed in this next subsection. We find

$$
\underset{\sim}{c}_1 = \begin{bmatrix} (\bar{\underset{\approx}{M}} \underset{\sim}{r}_0 + \underset{\sim}{a}) & \underset{\sim}{a} & \cdots & \underset{\sim}{a} & (\bar{\underset{\approx}{M}} \underset{\sim}{r}_N + \underset{\sim}{a}) \end{bmatrix}^T, \tag{21a}
$$

$$
\underset{\sim}{c}_2 = \begin{bmatrix} (\bar{\underset{\approx}{M}} \underset{\sim}{v}_0 + \underset{\sim}{d}) & \underset{\sim}{d} & \cdots & \underset{\sim}{d} & (\bar{\underset{\approx}{M}} \underset{\sim}{v}_N + \underset{\sim}{d}) \end{bmatrix}^T. \tag{21b}
$$

Now with each differential equation written as a matrix equation, it is worth showing that each of these matrix equations is invariant with respect to time.

Consider matrix differential equation of the form

$$\dot{x}(t) \;=\; A x(t) + B u(t) \tag{22a}$$

$$x_0 \;=\; \beta, \tag{22b}$$

where $x_0 \in \mathbb{R}^n$, $u_0 \in \mathbb{R}^m$ and $A, B \in \mathbb{R}^{n \times m}$. Notice that:

$$\dot{x}(t) - A x(t) \;=\; B u(t)$$

$$e^{-At}[\dot{x}(t) - A x(t)] \;\Longrightarrow\; \frac{d}{dt}[e^{-At} x(t)] = e^{-At} B u(t).$$

Now integrate from the known initial time t_1 to an instant $t > t_1$:

$$\int_{\tau=t_1}^{\tau=t} \frac{d}{d\tau}[e^{-A\tau} x(\tau)] \, d\tau = e^{-At} x(t) - e^{-At_1} x(t_1) = \int_{\tau=t_1}^{\tau=t} e^{-A\tau}(B u(\tau)) \, d\tau.$$

Next multiplying by the matrix exponential e^{At} yields the exact solution to (22):

$$x(t) = e^{A(t-t_1)} x(t_1) + \int_{\tau=t_1}^{\tau=t} e^{A(t-\tau)}(B u(\tau)) \, d\tau. \tag{23}$$

To convert this problem from a continuous model to a discrete model, we seek to evaluate $x(t)$ at $x(t_k) \triangleq x(t_{k-1} + \Delta t)$ for $k = 0, 1, 2, \ldots$. Begin with $t = t_1$ then integrate from t_1 to $t_2 = t_1 + \Delta t$. Next, integrate from t_2 to $t_3 = t_2 + \Delta t$ and continue iteratively from t_{k-1} to t_k. This gives

$$x(t_k) = e^{A(t_k - t_{k-1})} x(t_{k-1}) + \int_{\tau=t_{k-1}}^{\tau=t_k} e^{A(t_k-\tau)}(B u(\tau)) \, d\tau. \tag{24}$$

Until now it should be stated that the solution is exact. To introduce the approximation, assume that $u(\tau) \approx u(t_{k-1})$ for $t_{k-1} \le \tau < t_k$. With this, we can

express the integral above as:

$$\underset{\sim}{x}(t_k) = e^{\underset{\approx}{A}\Delta t}\underset{\sim}{x}(t_{k-1}) + \left[\int_{\tau=t_{k-1}}^{\tau=t_{k-1}+\Delta t} e^{\underset{\approx}{A}(t_{k-1}+\Delta t-\tau)}\, d\tau\right](\underset{\approx}{B}\underset{\sim}{u}(t_{k-1})). \tag{25}$$

We make the following two substitutions:

$$\xi = t_{k-1} + \Delta t - \tau. \tag{26a}$$

$$e^{\underset{\approx}{A}\xi} \approx I + \underset{\approx}{A}\xi. \tag{26b}$$

The Taylor approximation for the matrix exponential is used here and the integral in (25) with these substitutions becomes,

$$\begin{aligned}
\int_{\xi=\Delta t}^{\xi=0} e^{\underset{\approx}{A}\xi}\, d(-\xi) &= \int_{\xi=0}^{\xi=\Delta t} e^{\underset{\approx}{A}\xi}\, d\xi \\
&\approx \int_{\xi=0}^{\xi=\Delta t} I + \underset{\approx}{A}\xi\, d\xi \\
&= I\Delta t + \underset{\approx}{A}\frac{\Delta t^2}{2} \\
&= \Delta t(I + \underset{\approx}{A}\frac{\Delta t}{2}).
\end{aligned}$$

Substituting this above expression in for the integral, we have

$$\underset{\sim}{x}(t_k) \approx (I + \underset{\approx}{A}\Delta t)\underset{\sim}{x}(t_{k-1}) + \Delta t(I + \frac{\underset{\approx}{A}\Delta t}{2})\underset{\approx}{B}\underset{\sim}{u}(t_{k-1}). \tag{27}$$

By letting $\Phi = (I + \underset{\approx}{A}\Delta t)$ and $\Psi = \Delta t(I + \frac{\underset{\approx}{A}\Delta t}{2})\underset{\approx}{B}$, we can arrive at the following explicit recursive representation for the discrete dynamics of the space vehicle:

$$\underset{\sim}{x}(t_k) \approx \Phi\underset{\sim}{x}(t_{k-1}) + \Psi\underset{\sim}{u}(t_{k-1}). \tag{28}$$

Which, melds with the derivation of each of our dynamic equations, and (28) can be rewritten in a identical manner to (18).

Discretization of Cost Functions and Constraints

The cost function for the trajectory optimization problem must also be discretized. To discretize the cost functional, represented as $\int_0^{t_f} \sigma^2(t)dt$ and $\|r(t_f)\|^2$ for the respective minimum fuel landing problem and the minimum landing error problems proceeds as follows. For the MFLP cost functional, the authors of [2] use piece-wise constant step functions and the left-endpoint rule to approximate the cost function. For the MLEP cost functional, a standard direct discretization approach is taken.

We choose to approximate the cost function, dynamic equations and constraints using a finite element method. Before we get started, a few definitions. The inspiration for this is from [12], where such techniques are applied to PDE type problems. We consider the time-space domain $\Omega \times [0, t_f]$ where $\Omega \subset \mathbb{R}^d$ for $d \in \mathbb{N}$. To simplify the notation as much as possible we will simply use Ω to represent this time-space domain. Define

$$L_2(\Omega) = \left\{ \underset{\sim}{x} : \int_\Omega |\underset{\sim}{x}|^2 \, d\Omega < \infty \right\}$$

and consider a subspace of finite dimension $V_N \subset L_2(\Omega)$ consisting of the span of finite dimensional basis functions $\{\phi_1, \ldots, \phi_N\}$. Then any $x_n \in V_N$ can be given as

$$x_n = \sum_{i=0}^N x_i \phi_i$$

We are now set to express a norm on $L_2(\Omega)$ as

$$\|\underset{\sim}{x}\|^2_{L_2(\Omega)} = \sum_{i=0}^N \sum_{j=0}^N x_i x_j \int_\Omega \phi_i(t)\phi_j(t)dt := \underset{\sim}{x}^T \bar{\underset{\approx}{M}} \underset{\sim}{x} \tag{29}$$

Where $\bar{\underset{\approx}{M}}$ is the mass matrix defined to be $\bar{M}_{ij} = \int_\Omega \phi_i(t)\phi_j(t)dt$. Let us see how this can be applied to our cost function. Recall that our cost functional, as stated

prior to our updated discretization, was

$$J(\underline{r}, \sigma) = \int_0^{t_f} \sigma^2(t)\, dt + \lambda \left\| \underline{r}(t_f) \right\|_2^2. \tag{30}$$

Given the new definitions of $\hat{\underline{r}}$ and $\hat{\underline{\sigma}}$, defined in (17), define the sparse matrix $\underline{\underline{E}}_N \in \mathbb{R}^{3N \times 3(N+1)}$ to access the final-time entry for the position vector $\underline{r}$. The cost functional transforms into

$$J(\underline{r}, \underline{\sigma}) = \|\hat{\underline{\sigma}}\|_{L_2(\Omega)}^2 + \lambda \left\| \underline{\underline{E}}_N \hat{\underline{r}} \right\|_{L_2(\Omega)}^2.$$

Let's multiply each term by $\frac{1}{2}$ to make for simpler computation later down the line. Naturally this is no issue as the minimizer for the original cost function is equivalent to the minimizer with a constant multiplied to it. Now with our definition (29), we have our discretization, and an updated representation of the trajectory optimization problem. Define $\underline{\underline{\Lambda}}_\sigma \in \mathbb{R}^{N \times N+1}$ to be the diagonal matrix with the entries of $\hat{\underline{\sigma}}$ entered along the diagonal. In similar fashion as for the position vector, define $\hat{\underline{\underline{E}}}_N \in \mathbb{R}^{N \times N+1}$ to access the value at the final time for the mass as of the vehicle.

Problem 8. $\min\limits_{N} \min\limits_{r,\sigma} \frac{1}{2} \|\hat{\underline{\sigma}}\|_{L_2(\Omega)}^2 + \frac{\lambda}{2} \left\| \underline{\underline{E}}_N \hat{\underline{r}} \right\|_{L_2(\Omega)}^2$

    subject to:

$$\begin{cases} \underline{\underline{A}}\hat{\underline{r}} - \underline{\underline{B}}\hat{\underline{v}} = \underline{c}_1, & \text{(Governing Dynamics)} \\[4pt] \underline{\underline{A}}\hat{\underline{v}} - \underline{\underline{B}}\hat{\underline{u}} = \underline{c}_2, & \text{(Governing Dynamics)} \\[4pt] \underline{\underline{D}}\hat{\underline{z}} - \underline{\underline{F}}\hat{\underline{\sigma}} = 0, & \text{(Governing Dynamics)} \\[4pt] \hat{\underline{\underline{E}}}_N \hat{\underline{z}} \geq \ln m_{dry}, & \text{(Lower Bound on Mass)} \\[4pt] \tan^2(\gamma)\|\underline{\underline{E}}_r \hat{\underline{r}}\|^2 - \underline{\underline{E}}_1 \hat{\underline{r}}^2 \leq 0, & \text{(Cone Constraint on Position)} \end{cases}$$

$$\begin{cases} \|\hat{\underline{u}}\|^2 \leq \underline{\underline{\Lambda}}_\sigma \hat{\underline{\sigma}}, & \text{(Convexification Slack Variable)} \\[4pt] (\underline{1} - \hat{\underline{z}})\rho_1 \leq \hat{\underline{\sigma}} \leq (\underline{1} - \hat{\underline{z}})\rho_2, & \text{(Thrust Bound)} \end{cases}$$

This updated representation of the convexificaiton slack variable will be addressed here. Squaring both sides allows for promising results once the optimality conditions for this problem are found and the resulting matrix equation of optimality conditions can take on a symmetric form. Let's see how we arrived at the above representation. We began with the constraint having the form

$$\|\underset{\sim}{u}(t)\| \le \sigma^2(t).$$

After the discretization and reformulation of the components of this constraint into a lumped matrix involving $\hat{\underset{\sim}{u}}$ and $\hat{\underset{\sim}{\sigma}}$ we have

$$\|\hat{\underset{\sim}{u}}\| \le \hat{\underset{\sim}{\sigma}}.$$

Once the square of both sides is applied, notice that the left hand side can be rewritten using some of the results above from the finite element method, $(\hat{\underset{\sim}{u}})^T M \underset{\sim}{u}$ and the right hand side is similar in nature to the square of of a function, but in this instance it is more in line with component-wise squaring, which in turn can be expressed as $\underset{\approx}{\Lambda}_\sigma \hat{\underset{\sim}{\sigma}}$ where the matrix $\underset{\approx}{\Lambda}_\sigma$ is a diagonal matrix with the entries of $\hat{\underset{\sim}{\sigma}}$ along its diagonal. Because $\bar{\underset{\approx}{M}}$ is a diagonal matrix, it can be rewritten as $\bar{\underset{\approx}{M}} = \underset{\approx}{G}^T \underset{\approx}{G}$ (say) where $\underset{\approx}{G} = \underset{\approx}{G}^T$ is a matrix with the entries of $\bar{\underset{\approx}{M}}$ along its diagonal. Now then,

$$\hat{\underset{\sim}{u}}^T \bar{\underset{\approx}{M}} \hat{\underset{\sim}{u}} - \underset{\approx}{\Lambda}_\sigma \hat{\underset{\sim}{\sigma}} \;\le\; 0$$
$$\hat{\underset{\sim}{u}}^T \underset{\approx}{G}^T \underset{\approx}{G} \hat{\underset{\sim}{u}} - \underset{\approx}{\Lambda}_\sigma \hat{\underset{\sim}{\sigma}} \;\le\; 0$$

By the properties of the transpose and diagonal matrices,

$$(\underset{\approx}{G}\,\underset{\sim}{u})^T \underset{\approx}{G}\,\underset{\sim}{u} - \underset{\approx}{\Lambda}_\sigma \underset{\sim}{\hat{\sigma}} \;\leq\; 0$$

$$(\underset{\approx}{G}\,\underset{\sim}{u})^2 - \underset{\approx}{\Lambda}_\sigma \underset{\sim}{\hat{\sigma}} \;\leq\; 0$$

$$\underset{\approx}{\bar{M}}\,\underset{\sim}{u}^2 - \underset{\approx}{\Lambda}_\sigma \underset{\sim}{\hat{\sigma}} \;\leq\; 0$$

$$\underset{\approx}{\bar{M}}\,\underset{\approx}{\Lambda}_u\,\underset{\sim}{\hat{u}} - \underset{\approx}{\Lambda}_\sigma \underset{\sim}{\hat{\sigma}} \;\leq\; 0$$

where $\underset{\approx}{\Lambda}_u$ is the diagonal matrix with the entries of $\underset{\sim}{u}$ being along the diagonal.

Alternative Solution Approach

For a moment let us ignore the inequality constraints presented in the model and focus on the equality constraints. This action is taken to identify the discrete Lagrangian. Following this, we will return to handle the inequality constraints with the application of the Moreau-Yosida regularization. The discrete Lagrangian consists of the sum of the cost functional $J(\underset{\sim}{\hat{\sigma}},\underset{\sim}{\hat{r}})$ and Lagrange multiplier terms $\underset{\sim}{p}_1, \underset{\sim}{p}_2, \underset{\sim}{p}_3$ which will be used to enforce the constraints in the problem. By applying (29), our finite element discretization of our cost functional:

$$J(\underset{\sim}{\hat{\sigma}},\underset{\sim}{\hat{r}}) = \tfrac{1}{2}\underset{\sim}{\hat{\sigma}}^T \underset{\approx}{\bar{M}}\,\underset{\sim}{\hat{\sigma}} + \tfrac{\lambda}{2}(\underset{\approx}{E}_N \underset{\sim}{\hat{r}})^T \underset{\approx}{\hat{M}}\,\underset{\approx}{E}_N \underset{\sim}{\hat{r}}.$$

The discrete Lagrangian is found by rewriting the equality constraints equal to zero and multiplying the resulting expression by a Lagrange multiplier. This new Lagrange multiplier term is added to the original cost functional. For instance, consider our cost functional and only one equality constraint. Let's use the dynamics of the mass of the vehicle and consider the Lagrange multiplier $\underset{\sim}{p}_3$. The dynamics of the mass vehicle is $\underset{\approx}{D}\,\underset{\sim}{\hat{z}} = \underset{\approx}{F}\,\underset{\sim}{\hat{\sigma}}$. First, set it equal to zero and multiply by a Lagrange multiplier $\underset{\sim}{p}_3^T(\underset{\approx}{D}\,\underset{\sim}{\hat{z}} - \underset{\approx}{F}\,\underset{\sim}{\hat{\sigma}}) = 0$. Combining this with the cost functional

gives the discrete Lagrangian for one equality constraint as

$$L(\hat{\sigma}, \hat{r}, \hat{z}, \underset{\sim}{p}_3) = \tfrac{1}{2}\hat{\sigma}^T \bar{M}\hat{\sigma} + \tfrac{\lambda}{2}(E_N\hat{r})^T \hat{M} E_N\hat{r} + \underset{\sim}{p}_3^T(D\hat{z} - F\hat{\sigma}).$$

Extending this for all such equality constraints will yield this discrete Lagrangian

$$L(\hat{\sigma}, \hat{r}, \hat{v}, \hat{u}, \hat{z}, \underset{\sim}{p}_1, \underset{\sim}{p}_2, \underset{\sim}{p}_3) = \tfrac{1}{2}\hat{\sigma}^T \bar{M}\hat{\sigma} + \tfrac{\lambda}{2}(E_N\hat{r})^T \hat{M} E_N\hat{r} + \underset{\sim}{p}_1^T(A\hat{r} - B\hat{v} - \underset{\sim}{c}_1)$$

$$+ \underset{\sim}{p}_2^T(A\hat{v} - B\hat{u} - \underset{\sim}{c}_2) + \underset{\sim}{p}_3^T(D\hat{z} - F\hat{\sigma}). \quad (31)$$

To determine the optimality conditions for this problem, we take the derivative of L with respect to $\hat{r}, \hat{\sigma}, \hat{v}, \hat{u}, \hat{z}, \underset{\sim}{p}_1, \underset{\sim}{p}_2, \underset{\sim}{p}_3$ set to zero. These first order necessary conditions represent the Karush-Kuhn Tucker, or KKT conditions.

$$\frac{\partial L}{\partial \hat{r}} = \lambda E_N^T \hat{M} E_N \hat{r} + \underset{\sim}{p}_1^T A = 0 \tag{32a}$$

$$\frac{\partial L}{\partial \hat{\sigma}} = \bar{M}\hat{\sigma} - \underset{\sim}{p}_3^T F = 0 \tag{32b}$$

$$\frac{\partial L}{\partial \hat{v}} = -\underset{\sim}{p}_1^T B + \underset{\sim}{p}_2^T A = 0 \tag{32c}$$

$$\frac{\partial L}{\partial \hat{u}} = -\underset{\sim}{p}_2^T B = 0 \tag{32d}$$

$$\frac{\partial L}{\partial \hat{z}} = \underset{\sim}{p}_3^T D = 0 \tag{32e}$$

$$\frac{\partial L}{\partial \underset{\sim}{p}_1} = A\hat{r} - B\hat{v} - \underset{\sim}{c}_1 = 0 \tag{32f}$$

$$\frac{\partial L}{\partial \underset{\sim}{p}_2} = A\hat{v} - B\hat{u} - \underset{\sim}{c}_2 = 0 \tag{32g}$$

$$\frac{\partial L}{\partial \underset{\sim}{p}_3} = D\hat{z} - F\hat{\sigma} = 0 \tag{32h}$$

Combining these optimality conditions into a single matrix equation gives

$$
\begin{bmatrix}
\lambda E_N^T \hat{M} E_N & 0 & 0 & 0 & 0 & A^T & 0 & 0 \\
0 & \bar{M} & 0 & 0 & 0 & 0 & 0 & -F^T \\
0 & 0 & 0 & 0 & 0 & -B^T & A^T & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -B^T & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & D^T \\
A & 0 & -B & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & A & -B & 0 & 0 & 0 & 0 \\
0 & -F & 0 & 0 & D & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\hat{r} \\ \hat{\sigma} \\ \hat{v} \\ \hat{u} \\ \hat{z} \\ p_1 \\ p_2 \\ p_3
\end{bmatrix}
=
\begin{bmatrix}
0 \\ 0 \\ 0 \\ 0 \\ 0 \\ c_1 \\ c_2 \\ 0
\end{bmatrix}.
\tag{33}
$$

Moreau–Yosida Regularization

To begin to handle the inequality constraints, a digression is needed to discuss the theory overlaying the Moreau-Yosida regularization. This discussion is heavily influenced by [7], where I refer the reader who is seeking a more in-depth understanding of Moreau-Yosida Regularization.

Definition 9. From [7]. Let $f : \mathbb{R}^n \to R \cup \{+\infty\}$ be a proper closed convex function. The Moreau-Yosida regularization of a given function f, associated to the metric defined by M, denoted by F, is defined as follows and can otherwise be termed an *infimal convolution*:

$$
F(x) := \min_{y \in \mathbb{R}^n} f(y) + \frac{1}{2} \|y - x\|_M^2.
\tag{34}
$$

Theorem 10. *The infimal convolution of a convex function is a convex function.*

The proof of this theorem is found on p.50 of [13].

Proposition 11. *The infimal convolution, as defined in (34), is always differentiable.*

The proof is found in [9].

Proposition 12. *The following statements are equivalent:*

1. x^* is the minimizer for f.   *3. $\nabla F(x^*) = 0$.*

2. x^* is the minimizer for F.   *4. $F(x) = f(x)$ for all such x.*

I refer the reader to [6, 7, 9] for the proof of this proposition.

With some of the mechanics of the Moreau-Yosida regularization established, we may return to trajectory optimization problem at hand. To begin applying the regularization technique, the active sets associated with our bounded constraints must be determined. Define

$$
\mathscr{A}^+ \;=\; \{i : \hat{\underset{\sim}{\sigma}}_i > (\underset{\sim}{1} - \underset{\sim}{z})\rho_2\}, \quad \mathscr{A}^- = \{i : \hat{\underset{\sim}{\sigma}}_i < (\underset{\sim}{1} - \underset{\sim}{z})\rho_1\} \tag{35a}
$$

$$
\mathscr{B} \;=\; \{i : \tan^2(\gamma)\,\big\|\underset{\sim}{E}_r \hat{\underset{\sim}{r}}_i\big\|^2 > \underset{\approx}{E}_1 \hat{r}_i^2\} \tag{35b}
$$

$$
\mathscr{C} \;=\; \{\hat{\underset{\approx}{E}}_N \hat{\underset{\sim}{z}}_i < \ln m_{dry}\} \tag{35c}
$$

$$
\mathscr{D} \;=\; \{\hat{\underset{\sim}{u}}_i^T \bar{\underset{\approx}{M}} \hat{\underset{\sim}{u}}_i > \underset{\approx}{\Lambda}_\sigma \hat{\underset{\sim}{\sigma}}_i\} \tag{35d}
$$

With these, introduce $\chi_{\mathscr{A}^+}$, $\chi_{\mathscr{A}^-}$, $\chi_{\mathscr{B}}$, $\chi_{\mathscr{C}}$, and $\chi_{\mathscr{D}}$ to be the characteristic functions for the respective indices of $\hat{\underset{\sim}{\sigma}}, \hat{\underset{\sim}{r}}$ and $\hat{\underset{\sim}{z}}$. Let $\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_4 > 0$ be regularization parameters that penalize any dissatisfaction of what will be their respective inequality constraint. $(\underset{\sim}{1} - \hat{\underset{\sim}{z}})\rho_1 \leq \hat{\underset{\sim}{\sigma}} \leq (\underset{\sim}{1} - \hat{\underset{\sim}{z}})\rho_2$, the terms needed for the Moreau-Yosida regularization are

$$
\frac{1}{2\epsilon_1}\big\|min\{0, \hat{\underset{\sim}{\sigma}} - (\underset{\sim}{1} - \hat{\underset{\sim}{z}})\rho_1\}\big\|^2 \tag{36a}
$$

$$
\frac{1}{2\epsilon_1}\big\|max\{0, \hat{\underset{\sim}{\sigma}} - (\underset{\sim}{1} - \hat{\underset{\sim}{z}})\rho_2\}\big\|^2 . \tag{36b}
$$

Here ϵ_1 acts to provide a high penalty when the lower bound on the thrust is violated. A similar situation for ϵ_2. For the cone constraint, the replacement

Moreau-Yosida expression is

$$\frac{1}{2\epsilon_2}\left\|max\{0,\tan^2(\gamma)(\underset{\sim}{E}_r\hat{\underset{\sim}{r}})^T\hat{\underset{\approx}{M}}(\underset{\sim}{E}_r\hat{\underset{\sim}{r}})-\underset{\sim}{E}_1\hat{\underset{\sim}{r}}^2\}\right\|^2.\tag{37a}$$

Again, for $\hat{\underset{\approx}{E}}_N\hat{\underset{\sim}{z}}\geq\ln m_{dry}$, we introduce

$$\frac{1}{2\epsilon_3}\left\|max\{0,\ln m_{dry}-\hat{\underset{\approx}{E}}_N\hat{\underset{\sim}{z}}\}\right\|^2.\tag{38}$$

This leaves the convexification variable $\hat{\underset{\sim}{u}}^T\bar{\underset{\approx}{M}}\hat{\underset{\sim}{u}}\leq\underset{\approx}{\Lambda}_\sigma\hat{\underset{\sim}{\sigma}}$, its representation is

$$\frac{1}{2\epsilon_4}\left\|max\{0,\hat{\underset{\sim}{u}}^T\bar{\underset{\approx}{M}}\hat{\underset{\sim}{u}}-\underset{\approx}{\Lambda}_\sigma\hat{\underset{\sim}{\sigma}}\}\right\|^2.\tag{39}$$

Combining this with (31), the Lagrangian, including the Moreau-Yosida regularization parameters becomes

$L(\hat{\underset{\sim}{\sigma}},\hat{\underset{\sim}{r}},\hat{\underset{\sim}{v}},\hat{\underset{\sim}{u}},\hat{\underset{\sim}{z}},\underset{\sim}{p}_1,\underset{\sim}{p}_2,\underset{\sim}{p}_3)=\frac{1}{2}\hat{\underset{\sim}{\sigma}}^T\bar{\underset{\approx}{M}}\hat{\underset{\sim}{\sigma}}+\frac{\lambda}{2}(\underset{\approx}{E}_N\hat{\underset{\sim}{r}})^T\hat{\underset{\approx}{M}}\underset{\approx}{E}_N\hat{\underset{\sim}{r}}+\underset{\sim}{p}_1^T(\underset{\approx}{A}\hat{\underset{\sim}{r}}-\underset{\approx}{B}\hat{\underset{\sim}{v}}-\underset{\sim}{c}_1)$

$+\underset{\sim}{p}_2^T(\underset{\approx}{A}\hat{\underset{\sim}{v}}-\underset{\approx}{B}\hat{\underset{\sim}{u}}-\underset{\sim}{c}_2)+\underset{\sim}{p}_3^T(\underset{\approx}{D}\hat{\underset{\sim}{z}}-\underset{\approx}{F}\hat{\underset{\sim}{\sigma}}+\frac{1}{2\epsilon_1}max\{0,\hat{\underset{\sim}{\sigma}}-(\underset{\sim}{1}-\hat{\underset{\sim}{z}})\rho_2\}^T\underset{\approx}{M}'max\{0,\hat{\underset{\sim}{\sigma}}-(\underset{\sim}{1}-\hat{\underset{\sim}{z}})\rho_2\}$

$+\frac{1}{2\epsilon_1}min\{0,\hat{\underset{\sim}{\sigma}}-(\underset{\sim}{1}-\hat{\underset{\sim}{z}})\rho_1\}^T\underset{\approx}{M}'min\{0,\hat{\underset{\sim}{\sigma}}-(\underset{\sim}{1}-\hat{\underset{\sim}{z}})\rho_1\}$

$+\frac{1}{2\epsilon_3}max\{0,\ln m_{dry}-\hat{\underset{\approx}{E}}_N\hat{\underset{\sim}{z}}\}^T\underset{\approx}{M}''max\{0,\ln m_{dry}-\hat{\underset{\approx}{E}}_N\hat{\underset{\sim}{z}}\}$

$+\frac{1}{2\epsilon_4}max\{0,\hat{\underset{\sim}{u}}^T\bar{\underset{\approx}{M}}\hat{\underset{\sim}{u}}-\underset{\approx}{\Lambda}_\sigma\hat{\underset{\sim}{\sigma}}\}\underset{\approx}{M}'''max\{0,\hat{\underset{\sim}{u}}^T\bar{\underset{\approx}{M}}\hat{\underset{\sim}{u}}-\underset{\approx}{\Lambda}_\sigma\hat{\underset{\sim}{\sigma}}\}$

$+\frac{1}{2\epsilon_2}max\{0,\tan^2(\gamma)(\underset{\sim}{E}_r\hat{\underset{\sim}{r}})^T\hat{\underset{\approx}{M}}(\underset{\sim}{E}_r\hat{\underset{\sim}{r}})-\underset{\approx}{E}_1\hat{\underset{\sim}{r}}^2\}^T\hat{\underset{\approx}{M}}max\{0,\tan^2(\gamma)(\underset{\sim}{E}_r\hat{\underset{\sim}{r}})^T\hat{\underset{\approx}{M}}(\underset{\sim}{E}_r\hat{\underset{\sim}{r}})-\underset{\approx}{E}_1\hat{\underset{\sim}{r}}^2\}.$

Differentiating each Moreau-Yosida parameter for placement into the KKT system results in these three additional equations that can be added to (32).

$$\underset{\sim}{c}_3=\frac{\partial}{\partial\underset{\sim}{r}}\quad=\quad-2\epsilon_2^{-1}\chi_{\mathscr{B}}\hat{\underset{\approx}{M}}max\{0,\tan^2(\gamma)(\underset{\sim}{E}_r\hat{\underset{\sim}{r}})^T\hat{\underset{\approx}{M}}(\underset{\sim}{E}_r\underset{\sim}{r})-\underset{\approx}{E}_1\hat{\underset{\sim}{r}}^2\}(\tan^2(\gamma)\underset{\sim}{E}_r\hat{\underset{\approx}{M}}\underset{\approx}{E}_r\hat{\underset{\sim}{r}}-\underset{\approx}{E}_1\hat{\underset{\sim}{r}})$$

$$\underset{\sim}{c}_4=\frac{\partial}{\partial\underset{\sim}{z}}\quad=\quad\epsilon_3^{-1}\chi_{\mathscr{C}}\underset{\approx}{M}''max\{0,\ln m_{dry}-\hat{\underset{\approx}{E}}_N\hat{\underset{\sim}{z}}\}\hat{\underset{\approx}{E}}_N-\frac{\rho_2}{\epsilon_1}\chi_{\mathscr{A}^+}\underset{\approx}{M}'$$

$$\underset{\sim}{c}_5=\frac{\partial}{\partial\underset{\sim}{\sigma}}\quad=\quad-\epsilon_1^{-1}\chi_{\mathscr{A}^+}\underset{\approx}{M}'max\{0,\hat{\underset{\sim}{\sigma}}-(\underset{\sim}{1}-\hat{\underset{\sim}{z}})\rho_2\}-\epsilon_1^{-1}\chi_{\mathscr{A}^-}\underset{\approx}{M}'min\{0,\hat{\underset{\sim}{\sigma}}-(\underset{\sim}{1}-\hat{\underset{\sim}{z}})\rho_1\}$$

$$\underset{\sim}{c}_6=\frac{\partial}{\partial\hat{\underset{\sim}{u}}}\quad=\quad-\epsilon_4^{-1}\chi_{\mathscr{D}}\underset{\approx}{M}'''max\{0,\hat{\underset{\sim}{u}}^T\bar{\underset{\approx}{M}}\hat{\underset{\sim}{u}}-\underset{\approx}{\Lambda}_\sigma\hat{\underset{\sim}{\sigma}}\}\bar{\underset{\approx}{M}}\hat{\underset{\sim}{u}}$$

The KKT system is now given to be

$$
\begin{bmatrix}
\lambda \underset{\approx}{E}_N^T \hat{\underset{\sim}{M}} \underset{\approx}{E}_N & 0 & 0 & 0 & 0 & \underset{\approx}{A}^T & 0 & 0 \\
0 & \bar{\underset{\sim}{M}} & 0 & 0 & 0 & 0 & 0 & -\underset{\approx}{F}^T \\
0 & 0 & 0 & 0 & 0 & -\underset{\approx}{B}^T & \underset{\approx}{A}^T & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -\underset{\approx}{B}^T & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \underset{\approx}{D}^T \\
\underset{\approx}{A} & 0 & -\underset{\approx}{B} & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & \underset{\approx}{A} & -\underset{\approx}{B} & 0 & 0 & 0 & 0 \\
0 & -\underset{\approx}{F} & 0 & 0 & \underset{\approx}{D} & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\hat{\underset{\sim}{r}} \\ \hat{\underset{\sim}{\sigma}} \\ \hat{\underset{\sim}{v}} \\ \hat{\underset{\sim}{u}} \\ \hat{\underset{\sim}{z}} \\ \underset{\sim}{p}_1 \\ \underset{\sim}{p}_2 \\ \underset{\sim}{p}_3
\end{bmatrix}
=
\begin{bmatrix}
\underset{\sim}{c}_3 \\ \underset{\sim}{c}_5 \\ 0 \\ \underset{\sim}{c}_6 \\ \underset{\sim}{c}_4 \\ \underset{\sim}{c}_1 \\ \underset{\sim}{c}_2 \\ 0
\end{bmatrix} .
\tag{40}
$$

Notice that the matrix on the left-hand-side can be expressed as the form

$$
\underset{\approx}{P} =
\begin{bmatrix}
\underset{\approx}{M} & \underset{\approx}{K}^T \\
\underset{\approx}{K} & \underset{\approx}{0}
\end{bmatrix} .
\tag{41}
$$

The dimension of the entire matrix $\underset{\approx}{P}$ in is $(18n+11) \times (18n+11)$. Where the size of M is $11(n+1) \times 11(n+1)$, K is $11(n+1) \times 7n$ and the size of $\underset{\approx}{0}$ is $7n \times 7n$.

Well-Posedness

At face value, it is unclear if such a matrix $\underset{\approx}{P}$ in (41) is cosistant, particularly with zeros along the diagonal in $\underset{\approx}{M}$. In order to ensure that this matrix equation can be inverted, two things must be established. The matrix $\underset{\approx}{K}^T$ must be injective, and the intersection of the null space of $\underset{\approx}{K}$ and the null space of $\underset{\approx}{M}$ must be empty. To better understand the first requirement, a dissection into the structure of $\underset{\approx}{K}^T$ is a good place to begin. Let's analyze the base case for $N = 1$.

The expanded matrix $\underset{\approx}{K}^{T}$ is expressed as

$$
\begin{pmatrix}
-1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
-h & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & -h & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & -h & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \alpha h \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & -h & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -h & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & -h & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}.
$$

To place this into reduced row echelon form, the following actions are taken:

1. $R_1 \leftarrow R_1 \div -1$

2. $R_4 \leftarrow R_4 - R_1$

3. $R_7 \leftarrow R_7 + hR_1$

Now repeat this pattern involving the next respective row at each step, namely

4. $R_2 \leftarrow R_2 \div -1$

5. $R_5 \leftarrow R_5 - R_2$

6. $R_8 \leftarrow R_8 + hR_2$

and we can begin to see a pattern emerging. Continue this process iteratively until complete and it is easy to check that the resulting process will result in

$$
\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}.
$$

So as we can see for the base case, the $rank(\underset{\sim}{K}^{T}) = dim(col(\underset{\sim}{K}^{T})) = 7$ and therefore the matrix $\underset{\sim}{K}^{T}$, for $N = 1$ is injective and the columns of the matrix are linearly independent. With this exercise in mind, it would suffice to prove that for all $N \in \mathbb{N}$, the columns of the resulting matrix $\underset{\sim}{K}^{T}$ are always linearly independent.

Theorem 13 (The Rank Nullity Theorem). *For any $m \times n$ matrix $\underset{\sim}{A}$,*

$$
rank(\underset{\sim}{A}) + null(\underset{\sim}{A}) = n
$$

I refer the reader to any number of undergraduate textbooks for a proof of this common theorem.

Proposition 14. *For all $N \in \mathbb{N}$, The matrix $\underset{\sim}{K}^{T} \in \mathbb{R}^{11(N+1) \times 7N}$ as defined above is an injective matrix.*

Proof. Recall that the matrix $\underset{\approx}{K}^T = \begin{pmatrix} \underset{\approx}{A}^T & 0 & 0 \\ 0 & 0 & \underset{\approx}{F}^T \\ -\underset{\approx}{B}^T & \underset{\approx}{A}^T & 0 \\ 0 & \underset{\approx}{B}^T & 0 \\ 0 & 0 & \underset{\approx}{D}^T \end{pmatrix}$ and $\underset{\sim}{x} = \begin{pmatrix} \underset{\sim}{p}_1 \\ \underset{\sim}{p}_2 \\ \underset{\sim}{p}_3 \end{pmatrix}$. Also,

$$\underset{\approx}{A} = \begin{bmatrix} -\underset{\approx}{I} & \underset{\approx}{I} & 0 & \cdots & 0 \\ 0 & -\underset{\approx}{I} & \underset{\approx}{I} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -\underset{\approx}{I} & \underset{\approx}{I} \end{bmatrix}, \quad \underset{\approx}{B} = \begin{bmatrix} \underset{\approx}{h} & 0 & 0 & \cdots \\ 0 & \underset{\approx}{h} & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \underset{\approx}{h} \end{bmatrix}$$

$$\underset{\approx}{D} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 1 \end{bmatrix}, \quad \underset{\approx}{F} = \begin{bmatrix} -\alpha h & 0 & 0 & \cdots \\ 0 & -\alpha h & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -\alpha h \end{bmatrix}$$

To see that $\underset{\approx}{K}^T$ is injective it must be that if $\underset{\approx}{K}^T \underset{\sim}{x} = \underset{\sim}{0}$ then $\underset{\sim}{x} = \underset{\sim}{0}$. Notice that $\underset{\approx}{F}^T \underset{\sim}{p}_3 = \underset{\sim}{0}$ only when $\underset{\sim}{p}_3 \equiv \underset{\sim}{0}$. Also $\underset{\approx}{B}^T \underset{\sim}{p}_2 = \underset{\sim}{0}$ only when $\underset{\sim}{p}_3 \equiv \underset{\sim}{0}$. This follows because $0 < h < 1$ is imposed for the time-step. Now $\underset{\approx}{A}^T \underset{\sim}{p}_2 - \underset{\approx}{B}^T \underset{\sim}{p}_1 = \underset{\sim}{0}$ this is because it is already shown that $\underset{\sim}{p}_2 \equiv \underset{\sim}{0}$ and it must be that $\underset{\sim}{p}_1 \equiv \underset{\sim}{0}$ because of the restriction on h. Thus. $\underset{\sim}{p}_1, \underset{\sim}{p}_2, \underset{\sim}{p}_3$ must be equal to zero to satisfy $\underset{\approx}{K}^T \underset{\sim}{x} = \underset{\sim}{0}$. Therefore $\underset{\approx}{K}^T \in \mathbb{R}^{11(N+1)\times 7N}$ as defined above is an injective matrix. This completes the proof. $\square$

Returning to the case for $N = 1$ can provide the basis for the argument that the intersection of the null spaces of $\underset{\approx}{K}$ and $\underset{\approx}{M}$ is in fact the zero vector. Through a lengthy computation, the span of the null space for $\underset{\approx}{K}$ defined as $N(\underset{\approx}{K}) = \mathbb{S}$ where $\mathbb{S}$

is

$$\mathbb{S} = span\left\{
\begin{pmatrix}1\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{pmatrix}
\begin{pmatrix}0\\1\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{pmatrix}
\begin{pmatrix}0\\0\\1\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{pmatrix}
\begin{pmatrix}0\\0\\0\\0\\0\\0\\1\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{pmatrix}
\begin{pmatrix}-h\\0\\0\\0\\0\\0\\0\\0\\1\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{pmatrix}
\begin{pmatrix}0\\-h\\0\\0\\0\\0\\0\\0\\0\\0\\1\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\end{pmatrix}
\begin{pmatrix}0\\0\\-h\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\\1\\0\\0\\0\\0\\0\\0\\0\\0\end{pmatrix}
\begin{pmatrix}h^2\\0\\0\\0\\0\\0\\0\\-h\\0\\0\\0\\0\\0\\0\\1\\1\\0\\0\\0\\0\\0\\0\end{pmatrix}
\begin{pmatrix}0\\h^2\\0\\0\\0\\0\\0\\0\\-h\\0\\0\\0\\0\\0\\0\\0\\1\\1\\0\\0\\0\\0\end{pmatrix}
\begin{pmatrix}0\\0\\h^2\\0\\0\\0\\0\\0\\0\\-h\\0\\0\\0\\0\\0\\0\\0\\0\\1\\1\\0\\0\end{pmatrix}
\begin{pmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\\0\end{pmatrix}
\begin{pmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\end{pmatrix}
\begin{pmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\\tfrac{1}{\alpha h}\\0\\0\\0\\0\\1\\0\end{pmatrix}
\begin{pmatrix}0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\\tfrac{-1}{\alpha h}\\0\\0\\0\\0\\0\\1\end{pmatrix}
\right\}$$

The mass matrix $\underset{\approx}{M}$, for the case $N = 1$ is a diagonal matrix in $\mathbb{R}^{22\times22}$ with values along the diagonal corresponding only to the first component of the final time of $\hat{\underset{\sim}{r}}$ and the two values corresponding to $\hat{\underset{\sim}{\sigma}}$.

$$\underset{\approx}{M} = \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & \lambda\hat{m}_{r2} & 0 & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & \lambda\hat{m}_{r3} & 0 & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & \bar{m}_{\sigma_1} & 0 & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & \bar{m}_{\sigma_2} & 0 & \cdots \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \cdots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{bmatrix}$$

Computing $\underset{\approx}{M}(\mathbb{S})$ will generate a collection of vectors. Within this collection, the zero vectors present will correspond to vectors $\underset{\sim}{\xi}_i$ in $\mathbb{S}$ that satisfy $\underset{\approx}{M}\underset{\sim}{\xi}_i = 0$ and as such, the zero vectors correspond to the vectors living in the intersection of these

null spaces. Now, $\underset{\approx}{M}(\mathbb{S}) =$

$$\left\{ \begin{pmatrix} 0\\0\\0\\0\\\lambda\hat{m}_{r2}\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\\lambda\hat{m}_{r3}\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\m_{\sigma 2}\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\\frac{\bar{m}_{\sigma 1}}{\alpha h}\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\\frac{-\bar{m}_{\sigma 1}}{\alpha h}\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \right\}$$

From this collection we can determine that,

$$N(\underset{\approx}{M}) \cap \mathbb{S} = span \left\{ \begin{pmatrix} 0\\1\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\1\\0\\0\\1\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} -h\\0\\0\\0\\0\\0\\0\\1\\0\\1\\0\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\-h\\0\\0\\0\\0\\0\\0\\1\\0\\1\\0\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\-h\\0\\0\\0\\0\\0\\0\\0\\0\\1\\0\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} h^2\\0\\0\\0\\0\\0\\0\\0\\-h\\0\\0\\0\\1\\0\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\h^2\\0\\0\\0\\0\\0\\0\\0\\-h\\0\\0\\0\\1\\0\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\h^2\\0\\0\\0\\0\\0\\0\\0\\-h\\0\\0\\0\\1\\0\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1\\0 \end{pmatrix} \begin{pmatrix} 0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\0\\1 \end{pmatrix} \right\} \tag{42}$$

when $N = 1$. With all of these pieces in place, it appears as though the intersection is non-empty and as such poses a problem for the invertibility of the entire block matrix $\underset{\approx}{P}$. However, a deeper look into the physical interpretation behind these null spaces can shed some new light. Looking back at (), notice that any linear combination of any of those vectors will result in $\underset{\approx}{K}\underset{\sim}{\xi}_i = 0$. These even involve free values for z_1 and z_2 depicted in the last two entries of each vector- namely the last two vectors in $\mathbb{S}$. With this in mind, notice that within the intersection of $\underset{\approx}{M}$ and $\underset{\approx}{K}$, the vectors containing non-zero values for z_1 and z_2 do not appear. This is of significant importance because we see that in order for a vector to be within the intersection of the $\underset{\approx}{M}$ and $\underset{\approx}{K}$ null spaces, seen in (42), the values for z_1 and z_2

38

must be zero. Recall that during the nondimensionalization step, $z = \ln m$. This means z is undefined at $m = 0$ and is zero when $m = 1$. Physically, this is infeasible because the mass of any space vehicle will never be 1. In fact, you would expect the minimum mass value to be the mass of the vehicle without any fuel (m_{dry}) and values of m_{dry} are in the thousands of kilograms. Therefore it is reasonable to conclude that

$$N(\underset{\approx}{M}) \cap \mathbb{S} = \{\underset{\sim}{0}\}$$

because, physically, there are no such vectors in the span of the intersection to account for non-zero values of z_1 and z_2. A similar argument holds for values of $\hat{\underset{\sim}{u}}$ which is bounded below by $\rho_1 > 0$.

To address this concern regarding $\hat{\underset{\sim}{z}} > 0$, the following adjustment to the KKT system can be made. Redefine the values $\underset{\sim}{c}_3, \underset{\sim}{c}_4, \underset{\sim}{c}_5, \underset{\sim}{c}_6$ to be

$$\underset{\sim}{c}_4 \triangleq \epsilon_3^{-1} \underset{\approx}{G}_{\mathscr{C}} \hat{\underset{\approx}{E}}_N \underset{\approx}{M}'' \ln m_{dry} - \frac{\rho_2}{\epsilon_1} \underset{\approx}{G}_{\mathscr{A}^+} \underset{\approx}{M}'$$

$$\underset{\sim}{c}_3 \triangleq -2\epsilon_2^{-1} \underset{\approx}{G}_{\mathscr{B}} \hat{\underset{\approx}{M}} max\{0, \tan^2(\gamma)(\underset{\approx}{E}_r \hat{\underset{\sim}{r}})^T \hat{\underset{\approx}{M}}(\underset{\approx}{E}_r \underset{\sim}{r}) - \underset{\approx}{E}_1 \hat{\underset{\sim}{r}}^2\}(\tan^2(\gamma)\underset{\approx}{E}_r \hat{\underset{\approx}{M}} \underset{\approx}{E}_r \hat{\underset{\sim}{r}} - \underset{\approx}{E}_1 \hat{\underset{\sim}{r}})$$

$$\underset{\sim}{c}_5 \triangleq -\epsilon_1^{-1} \underset{\approx}{G}_{\mathscr{A}^+} \underset{\approx}{M}' max\{0, \hat{\underset{\sim}{\sigma}} - (\underset{\sim}{1} - \hat{\underset{\sim}{z}})\rho_2\} - \epsilon_1^{-1} \underset{\approx}{G}_{\mathscr{A}^-} \underset{\approx}{M}' min\{0, \hat{\underset{\sim}{\sigma}} - (\underset{\sim}{1} - \hat{\underset{\sim}{z}})\rho_1\}$$

$$\underset{\sim}{c}_6 \triangleq -\epsilon_4^{-1} \underset{\approx}{G}_{\mathscr{D}} \underset{\approx}{M}''' max\{0, \hat{\underset{\sim}{u}}^T \bar{\underset{\approx}{M}} \hat{\underset{\sim}{u}} - \underset{\approx}{\Lambda}_\sigma \hat{\underset{\sim}{\sigma}}\} \bar{\underset{\approx}{M}} \hat{\underset{\sim}{u}}$$

Where $\underset{\approx}{G}, \underset{\approx}{G}_{\mathscr{C}}, \underset{\approx}{G}_{\mathscr{A}^+}, \underset{\approx}{G}_{\mathscr{A}^-}, \underset{\approx}{G}_{\mathscr{B}}, \underset{\approx}{G}_{\mathscr{D}}$ are projection matrices onto the respective active sets $\mathscr{C} \cup \mathscr{C}^c, \mathscr{C}, \mathscr{A}^+, \mathscr{A}^-, \mathscr{B}$, and $\mathscr{D}$. This means that each $\underset{\approx}{G}$ is a matrix with 1's along the diagonal corresponding to indices where the constraint is violated. As a result of this, the other term now left out of $\underset{\sim}{c}_4$, the term involving $\hat{\underset{\sim}{z}}$, is moved to the right-hand side. By the nature of these Moreau-Yosida regularization terms, this constraint, when placed in the large block matrix $\underset{\approx}{M}$, ensures that all values of $\underset{\sim}{z}$ remain greater than zero and therefore ensures that $N(\underset{\approx}{M}) \cap N(\underset{\approx}{K}) = \{\underset{\sim}{0}\}$. These

39

alterations result in this KKT system.

$$
\begin{bmatrix}
\lambda E_N^T \hat{M} E_N & 0 & 0 & 0 & 0 & A^T & 0 & 0 \\
0 & \bar{M} & 0 & 0 & 0 & 0 & 0 & -F^T \\
0 & 0 & 0 & 0 & 0 & -B^T & A^T & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -B^T & 0 \\
0 & 0 & 0 & 0 & \epsilon^{-1}\hat{E}_N^T G M'' G \hat{E}_N & 0 & 0 & D^T \\
A & 0 & -B & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & A & -B & 0 & 0 & 0 & 0 \\
0 & -F & 0 & 0 & D & 0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
\hat{r} \\ \hat{\sigma} \\ \hat{v} \\ \hat{u} \\ \hat{z} \\ p_1 \\ p_2 \\ p_3
\end{bmatrix}
=
\begin{bmatrix}
c_3 \\ c_5 \\ 0 \\ c_6 \\ c_4 \\ c_1 \\ c_2 \\ 0
\end{bmatrix}
\quad (43)
$$

Theorem 15. *Given* $A = \begin{pmatrix} M & B^T \\ B & 0 \end{pmatrix}$. *Define* $V_0 := \{u \in V | Bu = 0\}$. *If B is surjective and if* $(Mv, v) > 0, \quad \forall v \in V_0$ *then the matrix is invertible.*

Proof. Let V be the space for u and Q be the space for p. From the assumption that $B : V \mapsto Q$ is onto. We see that $B : V_0^\perp \mapsto Q$ is isormorphic, i.e., it is invertible and also we see that $B^\top : Q \mapsto R(B^\top) = N(B)^\perp = V_0^\perp$ is isormorphic. Under the additional condition that M is positive on V_0, we shall now show that the columns of A are linearly independent, which is equivalent to the invertibility. Basically, we assume that $Ax = 0$. It is enough to show $x = 0$. With $x = (u, p)^\top$, we assume that

$$
Mu + B^\top p = 0 \quad \text{and} \quad Bu = 0.
$$

Since $B : V_0^\perp \mapsto Q$ is an isomorphism, we see that since $Bu = 0$, it holds $u \in V_0$. We now see that with $\Pi : V \mapsto V_0$,

$$
0 = \Pi(Mu + B^\top p) = \Pi M u. \quad (44)
$$

This gives that $u = 0$ since $M_0 = \Pi M : V_0 \mapsto V_0$ is isomorphic. Lastly, we see that

40

$B^{\top}\underset{\sim}{p} = \underset{\sim}{0}$ gives $\underset{\sim}{p} = \underset{\sim}{0}$. This completes the proof. $\qquad\qquad\qquad\qquad\qquad$ $\square$

With the problem condensed to a matrix equation shown to be consistent, fast solvers can be used to compute the solution. We intend to use a Broyden class quasi- Newton method. This choice rests on the fact that the Broyden method avoids computing derivative matrices at each iteration in the numerical solver. This next section covers the Broyden method with rank-two update. Once a solution set is found for our matrix equation, a line search algorithm will be used to compute the optimal final time, t_f^*.

Broyden's Method

Now that this problem boils down to solving the matrix equation(43), we use a version of a quasi-Newton (secant) method to solve it. This section will be dedicated to describing the method. A major advantage of quasi-Newton methods is that it does not require the computation of the Jacobian subsequently denoted B_k for k iterations. Instead, these quasi-Newton methods use an approximate form of the Jacobian is updated each iteration. The following material provides some introductory theory for the Broyden method. It is sourced from [11] and the reader is directed there for a more in-depth discussion.

Theorem 16. *Given a vector $s \neq 0, v \in \mathbb{R}^n$, $C \in \mathbb{R}^{n \times n}$, there is a unique matrix $B \in \mathbb{R}^{n \times n}$ such that*

$$Bs = v, \quad Bz = Cz, \quad \forall z \text{ such that } z^{\top}s = 0.$$

Then B is defined to be:

$$B = C + \frac{1}{s^T s}(v - Cs)s^T.$$

Let $w = \frac{v - Cs}{s^T s}$, and notice that B is an rank one update of C. Namely,

$$B = C + ws^T.$$

The proof of this theorem can be found in [11] Chapter 6.

With this representation for the Jacobian using a rank-one update, the algorithm for Broyden's method is given as:

Algorithm 17. *Choose $x_0 \in \mathbb{R}^n$ and $B_0 \in \mathbb{R}^{n \times n}$.*

For $i = 0, 1, \ldots$

$$
\begin{aligned}
x_{i+1} &= x_i + B_i^{-1}(-\nabla J(x_i)) \\
s_i &= x_{i+1} - x_i \\
v_i &= \nabla J(x_{i+1}) - \nabla J(x_i) \\
B_{i+1} &= B_i + \frac{1}{s_i^T s_i}(v_i - B_i s_i)s_i^T.
\end{aligned}
$$

Symmetric Rank Two Update

A useful theorem that will be applied below is the Sherman-Morrison-Woodbury theorem which provides a computationally cheap and efficient way to compute the inverse of a matrix. The theorem is presented here, sourced from [11].

Theorem 18 (Generalized Sherman-Morrison-Woodbury Formula). *Let $B = C + UV^T$ with $C \in \mathbb{R}^{n \times n}$ is invertible , $U \in \mathbb{R}^{n \times k}$, and $V \in \mathbb{R}^{n \times k}$. Assume also that $I + V^T C^{-1} U$ is invertible. Then, B is invertible and the inverse of B is given as:*

$$B^{-1} = C^{-1} - C^{-1}U(I_k + V^T C^{-1} U)^{-1} V^T C^{-1}. \tag{45}$$

where I_k is the identity matrix in $\mathbb{R}^{k \times k}$.

The symmetric rank two update for the Broyden method, known as the

Davison-Fletcher-Powerll (DFP) method is chosen for the implementation of our particular problem. The symmetric rank two update of a matrix can be given as:

$$B = C + \alpha u u^T + \beta w w^T. \tag{46}$$

We impose the condition $v = Bs = Cs$ which is known to be the quasi-Newton condition [11]. This condition together with (46) gives

$$\alpha u u^T s + \beta w w^T s = v - Cs. \tag{47}$$

Let $\alpha u^T s = 1$ and $\beta w^T s = 1$, we have

$$u + w = v - Cs.$$

Set $u = v$ and $w = -Cs$, then

$$\alpha = 1/v^T s \quad \text{and} \quad \beta = -1/(s^T Cs).$$

Therefore, the rank 2 correction is given as the following equation:

$$B = C + \frac{vv^T}{v^T s} - \frac{Css^T C}{s^T Cs}. \tag{48}$$

If we define the vectors $U = [\alpha u \; \beta w] \in \mathbb{R}^{n \times 2}$ and $V = [u \, w] \in \mathbb{R}^{n \times 2}$. Then, we can find out the formula of B^{-1} in terms of C^{-1}, u and w. With the following identities in mind : $u = v$ and $w = -Cs$, $\alpha = 1/u^T s = -1/u^T C^{-1} w$ and $\beta = -1/(w^T C^{-1} w)$, we can derive the following:

By applying Theorem 18 above to (48), we begin with

$$
\begin{aligned}
B^{-1} &= C^{-1} - [\alpha C^{-1}u \ \beta C^{-1}w]\begin{bmatrix} 1+\alpha u^T C^{-1}u & \beta u^T C^{-1}w \\ \alpha w^T C^{-1}u & 1+\beta w^T C^{-1}w \end{bmatrix}^{-1}\begin{bmatrix} u^T \\ w^T \end{bmatrix}C^{-1} \\
&= C^{-1} - [\alpha C^{-1}u \ \beta C^{-1}w]\begin{bmatrix} 1+\alpha u^T C^{-1}u & \beta u^T C^{-1}w \\ \alpha w^T C^{-1}u & 1+\beta w^T C^{-1}w \end{bmatrix}^{-1}\begin{bmatrix} u^T C^{-1} \\ w^T C^{-1} \end{bmatrix} \\
&= C^{-1} - \frac{1}{\beta w^T C^{-1}u}[\alpha C^{-1}u \ \beta C^{-1}w]\begin{bmatrix} 1+\beta w^T C^{-1}w & -\beta u^T C^{-1}w \\ -\alpha w^T C^{-1}u & 1+\alpha u^T C^{-1}u \end{bmatrix}\begin{bmatrix} u^T C^{-1} \\ w^T C^{-1} \end{bmatrix} \\
&= C^{-1} + \frac{w^T C^{-1}w}{u^T C^{-1}w}[\alpha C^{-1}u \ \beta C^{-1}w]\begin{bmatrix} 0 & -\beta u^T C^{-1}w \\ 1 & 1+\alpha u^T C^{-1}u \end{bmatrix}\begin{bmatrix} u^T C^{-1} \\ w^T C^{-1} \end{bmatrix} \\
&= C^{-1} + \frac{w^T C^{-1}w}{u^T C^{-1}w}[\alpha C^{-1}u \ \beta C^{-1}w]\begin{bmatrix} -\beta u^T C^{-1}ww^T C^{-1} \\ u^T C^{-1}+w^T C^{-1}+\alpha u^T C^{-1}uw^T C^{-1} \end{bmatrix} \\
&= C^{-1} + \frac{-1}{u^T C^{-1}w}\left[C^{-1}uw^T C^{-1}+C^{-1}wu^T C^{-1}+C^{-1}ww^T C^{-1}+\alpha C^{-1}wu^T C^{-1}uw^T C^{-1}\right] \\
&= C^{-1} + \frac{1}{u^T s}\left[-C^{-1}us^T - su^T C^{-1}+ss^T+\frac{u^T C^{-1}u}{u^T s}(ss^T)\right] \\
&= C^{-1} + \left[1+\frac{u^T C^{-1}u}{u^T s}\right]\frac{ss^T}{u^T s}-\frac{[C^{-1}us^T+su^T C^{-1}]}{u^T s} \\
&= C^{-1} + \left[1+\frac{v^T C^{-1}v}{v^T s}\right]\frac{ss^T}{v^T s}-\frac{[C^{-1}vs^T+sv^T C^{-1}]}{v^T s}.
\end{aligned}
$$

An important theorem for Broyden class quasi-Newton methods is that the Broyden method converges locally and at least linearly [11]. Let x_i and B_i denote the sequence of vectors and matrices produced by Broyden's method. Set

$$
e_i = x_i - x_*, \quad \text{and} \quad M_i = B_i - H(J)(x_i).
$$

and take note of the following equivalencies:

$$
e_{i+1} = -B_i^{-1}[\nabla J(x_i)-\nabla J(x_*)-H(J)(x_*)(x_i-x_*)]+B_i^{-1}M_i e_i
$$

$$M_{i+1} \quad = \quad M_i \left(I - \frac{1}{s_i^T s_i} s_i s_i^T \right) + \frac{1}{s_i^T s_i} (v_i - H(J)(x_*)s_i)s_i^T.$$

Theorem 19. *Let ∇J be differentiable in a ball Ω about a root $x_* \in \mathbb{R}^n$ whose derivative has a Lipschitz constant γ on the ball. Suppose that $H(J)(x_*)$ is invertible, with $\|H(J)(x_*)^{-1}\| \le \beta$. Let $x_0 \in \Omega$ and $B_0 \in \mathbb{R}^{n \times n}$ be given satisfying*

$$\|M_0\| + 2\gamma \|e_0\| \le \frac{1}{8\beta}.$$

Then the iterates x_i, B_i given by Broyden's method are well defined, and the errors satisfy

$$\|e_{i+1}\| \le \frac{\|e_i\|}{2}, \quad \forall i = 0, 1, \dots.$$

This theorem is proven by Broyden, Dennis, and More [8].

Theorem 20. *Considering B_{i+1}, the DFP symmetric rank two update of B_i is positive if B_i is positive and an exact line search is used to obtain the step length λ_i, i.e.,*

$$x_{i+1} = x_i + \lambda_i B_i^{-1}(-\nabla F(x_i)).$$

IV. CONCLUDING REMARKS

In this thesis, a novel solution method is provided to solve the trajectory optimization problem for Martian descent and landing. In chapter one a summary of previous solution methods involving the relaxation of non-convex constraints, the linearization of components, the discretization and the solution method using interior-point linear programming techniques is given. Chapter two is the bulk of the novel contributions. Here, the decision is made to simultaneously minimize the minimum fuel and minimum landing error problems, rather than treat them as separate but prioritized problems. Additional alterations are made to the nondimensionalization which involve the introduction of squared slack variables. A reworking of the discretization of the cost functionals, dynamical equations and the constraints are made. When working with a system of differential equations, as done in prior work such as [2, 3, 1], for which only a few components are needed to be optimized, the resulting KKT system becomes exceptionally sparse and singular. By leaving the differential equations separated during the discretization rather than lumping them into a system of differential equations removes several obstacles that would arise in the determination of the optimality conditions.

Our method expands upon the research done in [2, 3] and implements a new method for handling these mixed inequality constrained trajectory problems through the application of Moreau-Yosida regularization techniques. Some introductory material of Moreau-Yosida regularization is given to merit its use. For the rest of the equality constraints, Lagrange multipliers are used and the first order necessary conditions, optimality conditions, or better known KKT conditions are found and arranged into a large block matrix that takes the form of a saddle-point matrix.

Some analysis of the well-posedness of the problem is undertaken to provide some reasoning as to why the matrix system can be solved. Within the large block

matrix, because the matrix $\underset{\approx}{B}$ is shown to to surjective, and because $N(\underset{\approx}{M}) \cap$ $N(\underset{\approx}{K}) = \underset{\sim}{0}$, then the matrix $\underset{\approx}{P}$ is proven to be invertible. With invertibility shown, the quasi- Newton Broyden method will be used to quickly solve the resulting optimization problem. The symmetric rank two update is used to reduce the computation time needed to solve the optimization problem. This is because the symmetric rank two updated provides a fast means of computing the the next iteration of the inverse of a matrix. Some theorems related to the super-linear convergence of the Broyden method is stated to justify to the reader that the resulting problem and solution method will guarantee a solution under appropriate conditions.

To completely establish this novel approach to solving the trajectory optimization problem, there are a few additional tasks needed to be undertaken. Further analysis of the nonlinear constraints seen on the right-hand side of (43) will be done. This solution technique also needs to be implemented and tested against the current benchmarks. Additionally, adding a thrust-pointing constraint such as

$$\Gamma^2(t) \le -\underset{\sim}{e}_1 \underset{\sim}{T}(t) \sec(\theta)$$

would enhance this model's on-board implementability and would provide additional safety measures that would accommodate human cargo. Having a thrust-pointing constraint ensures that the direction of the space vehicle is always pointing in the direction of the landing surface. This further ensures that unorthodox trajectories are avoided that would, say, place the vehicle on a horizontal trajectory close to the planetary surface for extended periods of time. Such a trajectory is in some instances a feasible solution to the problem. This constraint is also a requirement for terrain relative navigation (TRN). TRN is software that uses real-time image data procured from downward facing cameras and sensors on the space vehicle. These images are then used to provide the vehicle

with location and other telemetry data needed to locate the vehicle in relation with the targeted landing location. This software is needed especially when landing on a surface with no active/real-time GPS data.

REFERENCES

[1] B. Acikmese, L. Blackmore, D.P. Scharf, and A. Wolf. Enhancements on the convex programming based powered descent guidance algorithm for mars landing. In *AIAA/AAS Astrodynamics Specialist Conference and Exhibit*, number AIAA/AAS Astrodynamics Specialist Conference and Exhibit, Jet Propulsion Laboratory, California Institute of Technology, M/S 198-326, 2008.

[2] Behcet Acikmese and Scott R. Ploen. Convex programming approach to powered descent guidance for mars landing. *Journal of Guidance, Control & Dynamics*, 30(5):1353 – 1366, 2007.

[3] Lars Blackmore, Behcet Acikmese, and Daniel P. Scharf. Minimum-landing-error powered-descent guidance for mars landing using convex optimization. *Journal of Guidance, Control & Dynamics*, 33(4):1161 – 1171, 2010.

[4] William R. Corliss. *The Viking mission to Mars.* NASA SP: 334. Scientific and Technical Information Office, National Aeronautics and Space Administration, 1974.

[5] Daniel Dueri, Behcet Acikmese, Daniel P. Scharf, and Matthew W. Harris. Customized real-time interior-point methods for onboard powered-descent guidance. *Journal of Guidance, Control & Dynamics*, 40(2):197 – 212, 2017.

[6] A. Fuduli. Metodi numerici per la minimizzazione di funzioni convesse nondifferenziabili, phd thesis. 1998.

[7] Milanka Gardasevic-Filipovic and Nada Duranovic-Milicic. An algorithm using moreau-yosida regularization for minimization of a nondifferentiable convex function. *Filomat*, 27(1):15 – 22, 2013.

[8] Chi-Ming Ip and Jerzy Kyparisis. Local convergence of quasi-newton methods for b-differentiable equations. *Mathematical Programming*, 56(1-3):71, 1992.

[9] C. Lemarechal and C. Sagastizabal. Practical aspects of the moreau-yosida regularization: Theoretical preliminaries. *SIAM Journal on Optimization*, 7(2):367–385, 1997.

[10] Olivia Lowenberg. Falcon 9 sticks second sea landing, another victory for spacex's reusable rockets. *The Christian Science Monitor*, 2016.

[11] Jorge Nocedal and Stephen J. Wright. *Numerical optimization.* Springer series in operations research and financial engineering. Springer, 2006.

[12] John W. Pearson. A dissertation titled: Fast iterative solvers for pde-constrained optimization problems. Oxford University, 2013.

[13] R. Tyrrell Rockafellar. *Convex analysis.* Princeton landmarks in mathematics and physics. Princeton University Press, 1997.