INDIRECT METHODS FOR SPACE TRAJECTORY OPTIMIZATION
Lorenzo Casalino
Indirect Optimization Methods

Outline

• introduction
• indirect methods: how to make them
  – optimization of discrete systems
  – optimization of continuous systems
    * calculus of variations and indirect optimization
    * solution of boundary value problems
  – indirect trajectory optimization
• indirect methods: how to make them work
  – general rules
  – hints and tips
Trajectory Design

- Trajectory design usually requires maximization or minimization of a performance index (e.g., maximize final mass or payload, minimize, $\Delta V$, propellant consumption or trip time) for given mission requirements.

- Typical missions
  - Transfer between specified orbits
  - Rendezvous mission (transfer from a specified orbit to a definite time-dependent position on a target orbit)
  - Interplanetary missions (rendezvous missions with planets' gravity)
  - Station keeping
Trajectory Model

• spacecraft state defined by state variables \( x_i, i = 1, n \) 
  (position, velocity, mass, angular position and velocity for attitude control)

• a propulsion system can produce thrust (or angular momentum) to modify the rate of change (derivative) of the spacecraft velocity

• thrust components are typically the control variables \( u_j, j = 1, m \)
Indirect Optimization Methods

Optimization Problem

- system of equations \( \dot{x}_i = f_i[x_1(t), ..., x_n(t), u_1(t), ..., u_m(t), t], \) for \( i = 1, n \)
- boundary conditions concerning state (and possibly control) variables at initial, final and possibly intermediate points
- find the control variables \( u_j(t) (j = 1, m) \) and the corresponding trajectory \( x_i(t) (i = 1, n) \) that maximize a given performance index \( J \) while satisfying differential equations and boundary conditions
Trajectory Assumptions

- preliminary design usually (but not necessarily) admits simplifications
  - two-body problem
  - patched-conic approximation for interplanetary transfers
- it is often convenient to split a complex trajectory into elementary legs that are then patched together
Solution Methods (1)

- direct methods
  - discretization of trajectory and controls
  - start from tentative solution
  - evaluate performance index and constraints and respective gradients
  - change tentative solution with the aim of improving performance index and constraint fulfillment
Solution Methods (2)

- indirect methods
  - continuous system described by differential equations with assigned boundary conditions
  - derive first-order necessary conditions for optimality and define a boundary value problem
  - start from tentative solution
  - evaluate error on boundary conditions and error gradient
  - change tentative solution with the aim of reducing the error
Solution Methods (3)

- evolutionary algorithms
  - more suitable to deal with a low number of parameters (impulsive missions)
  - solution described by finite set of variable values
  - random initialization of a population of different solutions
  - new solutions are created by combining old solutions (typically, using some random process) with the aim of creating new populations with improved performance index
Direct Methods

• pros
  – capability of treating complex problems
  – easier treatment of constraint “structure”
  – higher robustness

• cons
  – scarce accuracy (may require solution refinement)
  – high computational cost
  – solution may depend on tentative values and be suboptimal
Indirect Optimization Methods

Indirect Methods

- **pros**
  - high accuracy
  - low computational cost and time
  - theoretical insight

- **cons**
  - difficult treatment of complex equations or constraints and necessary preliminary assumption of constraint “structure” (an “a posteriori” analysis may suggest changes)
  - lower robustness
  - find stationary solutions, possibly suboptimal, and dependent on tentative values
Evolutionary Algorithms

• pros
  – no tentative solution
  – higher robustness
  – “global” optimization

• cons
  – (possibly) scarce accuracy (may require solution refinement)
  – difficult treatment of constraints
  – heuristic methods: no proof of finding actual optimum
Indirect Optimization Methods

Introduction

Shooting Method

- both direct and indirect methods present unknown initial values and constraints or optimality conditions to be satisfied
- shooting method
  - choose initial tentative values for the unknowns
  - compute error on constraints
  - change tentative values
- errors on constraints may be very sensitive to the initial values, affecting convergence
Multiple Shooting

- split the trajectory into phases
- introduce the variable values at the start of each phase as additional unknowns
- enforce conditions for trajectory consistency (e.g., variable continuity) at the phase junctions
- reduced sensitivity, improved convergence
- higher computational cost? not always
Discrete Systems

Summary

• notation and assumptions
• one-variable optimization
• unconstrained multivariable optimization
• constrained optimization
  – equality constraints
  – inequality constraints
Vectorial Notation (1)

- a column vector is indicated by a bold character $\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \cdots \\ a_n \end{pmatrix}$

- a row vector is written as the transpose of a column vector $\mathbf{a}^T = (a_1, a_2, \cdots, a_n)$

- the scalar product of two vectors (e.g., $\mathbf{a}$ and $\mathbf{b}$, same dimension) is written as $\mathbf{a} \cdot \mathbf{b} = \mathbf{a}^T \mathbf{b} = \mathbf{b}^T \mathbf{a}$
Vectorial Notation (2)

- the derivative of a column vector (e.g., $a$, $n$ components) with respect to a scalar variable (e.g., $t$) is a column vector with components equal to the derivatives of the components of $a$ with respect to $t$

$$\frac{da}{dt} = \dot{a} = \begin{pmatrix} \frac{da_1}{dt} \\ \frac{da_2}{dt} \\ \vdots \\ \frac{da_n}{dt} \end{pmatrix}$$

- the derivative of a scalar quantity (e.g., $\phi$) with respect to a column vector (e.g., $b$, $m$ components) is defined as a row vector with components equal to the derivatives of $\phi$ with respect to the components of $b$

$$\frac{d\phi}{db} = (\frac{d\phi}{db_1}, \frac{d\phi}{db_2}, \cdots, \frac{d\phi}{db_m})$$
Vectorial Notation (3)

- by extension, the derivative of a column vector (e.g., $a$) with respect to another column vector (e.g., $b$) is a matrix
- the $i$-th row contains the derivatives of the $i$-th component of $a$ with respect to $b$
- the $j$-th column contains the derivatives of the constraints $a$ with respect to the $j$-th component of $b$

\[
\frac{da}{db} = \begin{bmatrix}
\frac{da_1}{db_1} & \frac{da_1}{db_2} & \cdots & \frac{da_1}{db_m} \\
\frac{da_2}{db_1} & \frac{da_2}{db_2} & \cdots & \frac{da_2}{db_m} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{da_n}{db_1} & \frac{da_n}{db_2} & \cdots & \frac{da_n}{db_m}
\end{bmatrix}
\]
Assumptions

• performance index and boundary conditions are continuous and differentiable with respect to the variables
• existence of an optimal solution is usually assumed
One-Variable Unconstrained Optimization (1)

- \( \phi(\hat{x}) \) is maximum if \( \phi(\hat{x}) > \phi(x) \) (strong) or \( \phi(\hat{x}) \geq \phi(x) \) (weak) for any \( x \neq \hat{x} \)

- \( \phi(\hat{x}) \) is maximum (strong) if \( d\phi = \phi(\hat{x} + dx) - \phi(\hat{x}) < 0 \) for any choice of \( dx \neq 0 \) (weak maximum if \( d\phi \leq 0 \))

- if \( dx \) is small (local variation) use second-order Taylor’s expansion of \( \phi \)
  \[
  \phi(x + dx) = \phi(x) + gdx + \frac{1}{2}Hdx^2
  \]

- the first derivative of \( \phi \) is \( g = \frac{\partial \phi}{\partial x} \)

- the second derivative of \( \phi \) is \( H = \frac{\partial^2 \phi}{\partial x^2} \)
Indirect Optimization Methods

Optimization of discrete systems

One-Variable Unconstrained Optimization (2)

- a stationary point is characterized by \( g = 0 \)
- necessary conditions for a (local) maximum \( g = 0, \ H \leq 0 \)
- sufficient conditions for a (local) maximum \( g = 0, \ H < 0 \)
- \( H = 0 \) requires computation of additional derivatives to define the nature of the stationary point - if the first nonzero derivative is of order \( j \)
  - neither max nor min if \( j \) is odd
  - max if \( j \) is even and the derivative is negative
  - min if \( j \) is even and the derivative is positive
- the sign of \( H \) must be reversed for a minimum
Unconstrained Optimization (1)

- maximize $\phi(x)$
- $x$ $n$-component vector of variables $x^T = (x_1, x_2, ..., x_n)$
- $\phi$ is maximum if $d\phi \leq 0$ for any choice of $dx$
- second-order Taylor’s expansion of $\phi$ (small variations)
  \[
  \phi(x + dx) = \phi(x) + g^T dx + \frac{1}{2} dx^T H dx
  \]
- the gradient of $\phi$ is $g$ ($n$-component vector)
  \[
  g^T = \frac{\partial \phi}{\partial x} = (\partial \phi/\partial x_1, \partial \phi/\partial x_1, ..., \partial \phi/\partial x_n)
  \]
Unconstrained Optimization (2)

- the Hessian of $\phi$ is $[H]$ ($n \times n$ symmetric matrix)

$$[H] = \frac{\partial \left( \frac{\partial \phi}{\partial x} \right)^T}{\partial x} = \frac{\partial g}{\partial x} = \begin{bmatrix} \frac{\partial^2 \phi}{\partial x_1 \partial x_1} & \frac{\partial^2 \phi}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 \phi}{\partial x_1 \partial x_n} \\ \frac{\partial^2 \phi}{\partial x_1 \partial x_2} & \frac{\partial^2 \phi}{\partial x_2 \partial x_1} & \cdots & \frac{\partial^2 \phi}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 \phi}{\partial x_1 \partial x_n} & \frac{\partial^2 \phi}{\partial x_2 \partial x_n} & \cdots & \frac{\partial^2 \phi}{\partial x_n \partial x_n} \end{bmatrix}$$
Unconstrained Optimization (3)

- first-order expansion of $\phi$ provides $d\phi = \frac{\partial \phi}{\partial x} dx = g^T dx$
- $\phi$ being a local maximum (or minimum) requires $g = 0$ (necessary condition for a stationary point)
- an additional condition is required to determine the nature of the stationary point
- considering the second variation of $\phi$ at a stationary point ($g = 0$), one has $d\phi = dx^T H dx$
- for a maximum $dx^T H dx \leq 0$ for any $dx$, that is $H$ negative semidefinite for a maximum (non-positive eigenvalues)
- sufficient conditions for a maximum $g = 0$ and $dx^T H dx < 0$ for any $dx \neq 0$, that is $H$ negative definite for a maximum (negative eigenvalues)
Equality Constrained Optimization (1)

- maximize $\phi(x)$ ($x$ is a $n$-component vector of variables) with $c(x) = 0$ ($c$ is a $m$-component vector of constraints, $c^T = [c_1, c_2, ..., c_m]$, $m < n$)
- feasible points are those which satisfy $c(x) = 0$
- at a feasible point $\phi$ is maximum if $d\phi \leq 0$ for any admissible (i.e., that verifies $c(x + dx) = 0$) choice of $dx$
- at a feasible point, admissible variations require $dc = c(x + dx) - c(x) = 0$
Indirect Optimization Methods

Optimization of discrete systems

Equality Constrained Optimization (2)

- small variations are considered (local maximum)
- first order expansion of $c$ gives $c(x + dx) = c(x) + [G]dx$
- $[G]$ is the Jacobian ($m \times n$ matrix)

$$
[G] = \left[ \frac{\partial c_i}{\partial x_j} \right] =
\begin{bmatrix}
\frac{\partial c_1}{\partial x_1} & \frac{\partial c_1}{\partial x_2} & \cdots & \frac{\partial c_1}{\partial x_n} \\
\frac{\partial c_2}{\partial x_1} & \frac{\partial c_2}{\partial x_2} & \cdots & \frac{\partial c_2}{\partial x_n} \\
\vdots & \vdots & \ddots & \vdots \\
\frac{\partial c_m}{\partial x_1} & \frac{\partial c_m}{\partial x_2} & \cdots & \frac{\partial c_m}{\partial x_n}
\end{bmatrix}
$$

- the $i$-th row of $G$ is the gradient of the $i$-th constraint $c_i$
- the $j$-th column of $G$ contains the derivatives of the constraints with respect to the $j$-th variable $x_j$
- admissible (small) variations require $[G]dx = 0$ ($m$ equations)
Equality Constrained Optimization (3)

- $n$ variables, $m$ equations allow for $n-m$ free parameters
- define $y^T = [x_1, x_2, ..., x_m]$ as $m$ state variables and $u^T = [x_{m+1}, x_{m+2}, ..., x_n]$ as $n-m$ control (or decision) variables (the choice is usually not unique)
- obtain admissible variations $dy$ as a function of $du$ from $[G_y]dy + [G_u]du = 0$
- first order expansion $\phi(u + du) = \phi(u) + g_y^T dy + g_u^T du$
- $[G_y] = \partial c/\partial y$, $[G_u] = \partial c/\partial u$, $g_y^T = \partial \phi/\partial y$, $g_u^T = \partial \phi/\partial u$
- $\phi$ is maximum if $d\phi \leq 0$ for any choice of $du$
- necessary condition for a stationary point $-g_y^T [G_y]^{-1} [G_u] + g_u^T = 0$
Equality Constrained Optimization (4)

- alternatively, maximize the augmented function
  \[ \phi^* = \phi(x) + \lambda^T c \] (\( \lambda \) is a \( m \)-component vector of adjoint parameters to be determined)

- \( \phi \) and \( \phi^* \) coincide for any choice of \( \lambda \) if the constraints are verified

- second-order Taylor’s expansion of \( \phi^* \) (local maxima are sought)
  \[ \phi^*(x + dx) = \phi^*(x) + (g^T + \lambda^T [G])dx + \frac{1}{2} dx^T [H^*] dx \]

- augmented Hessian \( [H^*] = [H] + \partial[\partial(\lambda^T c/\partial x)^T] \partial x \), that is
  \[ H^*_{ij} = H_{ij} + \sum_{k=1}^{m} \lambda_k \frac{\partial^2 c_k}{\partial x_i \partial x_j} \]
Equality Constrained Optimization (5)

- $\lambda$ can be chosen to nullify the first variation of $\phi^*$ for any choice of $dx$ (also those that do not satisfy $c = 0$)

- First-order necessary condition for maximum $\phi$ is $g^T + \lambda^T [G] = 0$, that is, $g + [G]^T \lambda = 0$ ($n$ conditions), and $c = 0$ ($m$ conditions) for $n + m$ unknowns $x$ and $\lambda$

- Second-order necessary condition for a maximum $dx^T [H^*] dx < 0$ for any admissible $dx$ ($[H^*]$ is not required to be negative semidefinite)

- Sufficient conditions for a maximum $g^T + \lambda^T [G] = 0$ and $dx^T [H^*] dx < 0$ for any admissible $dx$ ($[H^*]$ is not required to be negative definite)
Equality Constrained Optimization (6)

- $g + [G]^T \lambda = 0$ relates the gradient of $\phi$ to the gradients of the constraints.
- With only one constraint ($m = 1$), $\frac{\partial \phi}{\partial x} = -\lambda \frac{\partial c}{\partial x}$, that is, the gradient of $\phi$ must be parallel to the gradient of $c$.
- Since the gradient of $c$ is perpendicular to the constraint $c = 0$, this means that the gradient of $\phi$ must be perpendicular to $c = 0$.
- For more than one constraint ($m \geq 2$), the gradient of $\phi$ must be a linear combination of the constraints gradients.
Equality Constrained Optimization - Example

- \( \phi = -x_1^2 - x_2^2, \ c = x_1 + x_2 - 1 = 0 \ (n = 2, \ m = 1) \)
  \[
  \phi^* = -x_1^2 - x_2^2 + \lambda(x_1 + x_2 - 1) \\
  g^T = [-2x_1, -2x_2], \ [G] = (1, 1), \\
  g^T + \lambda^T[G] = (-2x_1 + \lambda, -2x_2 + \lambda) \\
  \]
  stationary point requires \(-2x_1 + \lambda = 0, -2x_2 + \lambda = 0, \ x_1 + x_2 - 1 = 0, \) that is \(x_1 = x_2 = 1/2\) and \(\lambda = 1\)

- \([H^*] = [H] = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix}\), admissible variation is
  \[
  (dx)^T = (-a, a) \\
  dx^T[H^*]dx = -4a^2 < 0 \text{ therefore } x^T = (1/2, 1/2) \text{ is a maximum (in this case, } [H^*] \text{ is negative definite and } dx^T[H^*]dx < 0 \text{ for any } dx)\]
Indirect Optimization Methods

Optimization of discrete systems

Inequality Constrained Optimization (1)

- maximize $\phi(x)$ ($x$ is a $n$-component vector of variables) with $c(x) \geq 0$ ($c$ is a $m$-component vector of constraints, $c^T = [c_1, c_2, ..., c_m]$, $m$ may be larger than $n$)
- feasible points are those which satisfy $c(x) \geq 0$
- $\phi$ is maximum if $d\phi \leq 0$ for any admissible (i.e., that verifies $c(x + dx) \geq 0$) choice of $dx$
- at a feasible point a constraint may be inactive ($c_j > 0$) and can be neglected or active ($c_j = 0$) and can be (at least, initially) treated as an equality constraint
- at a feasible point admissible variations require $dc_a \geq 0$ ($c_a$ only contains the active constraints, $m_a$ components)
Inequality Constrained Optimization (2)

- main issue: determine the active set of constraints
- for a given active set, an equality constrained optimization problem must be solved
  - if the solution violates an inactive constraint, the constraint must be added to the active set and an augmented equality constrained problem must be solved
  - if the solution is feasible, the requirement of the active constraints must be checked and unnecessary constraints must be removed producing a reduced-dimension problem
Indirect Optimization Methods

Inequality Constrained Optimization (3)

- given the active set solve the equality constrained optimization
- maximize the augmented function \( \phi^* = \phi(x) + \lambda_a^T c_a \) (with \( c_a = 0 \) active set of constraints and \( \lambda_a \) vector of adjoint parameters to be determined, \( m_a \) components)
- \( \phi \) and \( \phi^* \) coincide for any choice of \( \lambda_a \) if the active constraints are verified
- second-order Taylor’s expansion of \( \phi^* \) (local maxima are sought)
  \[
  \phi^*(x + dx) = \phi^*(x) + (g^T + \lambda_a^T [G_a])dx + \frac{1}{2}dx^T[H^*]dx
  \]
- augmented Hessian \([H^*] = [H] + \partial[\partial(\lambda_a^T c_a/\partial x)^T]\partial x\), that is
  \[
  H^*_{ij} = H_{ij} + \sum_{k=1}^{m} \lambda_{ak} \frac{\partial^2 c_{ak}}{\partial x_i \partial x_j}
  \]
Inequality Constrained Optimization (4)

- for the equality constrained optimization given the active set
  - \( \lambda_a \) can be chosen to nullify the first variation of \( \phi^* \) for any choice of \( dx \)
  - first-order necessary condition \( g^T + \lambda_a^T [G_a] = 0 \), that is, \( g + [G_a]^T \lambda_a = 0 \) (\( n \) conditions) and \( c_a = 0 \) (\( m_a \) conditions) for \( n + m_a \) unknowns \( x \) and \( \lambda_a \)
  - second-order necessary condition for a maximum \( dx^T [H^*] dx \leq 0 \) for any admissible \( dx \) (\([H^*]\) is not required to be negative semidefinite)
  - \( dx^T [H^*] dx < 0 \) for any admissible \( dx \) is a sufficient condition
Indirect Optimization Methods
Optimization of discrete systems

Inequality Constrained Optimization (5)

- at a feasible stationary point \( (d\phi^* = 0) \) of the equality constrained problem \( (c_a = 0) \), admissible variations of the inequality constrained problem require \( dc_a \geq 0 \)
- \( d\phi = d\phi^* - \lambda_a^T dc_a \) and at a maximum \( (d\phi \leq 0) \) \( \lambda_a^T dc_a \geq 0 \) is required for any admissible \( dc_a \)
- this implies that all the components of \( \lambda_a \) must be positive
- the constraint \( c_j = 0 \) must be removed from the active set when \( \lambda_{aj} < 0 \)
- generalizing, necessary conditions are \( c \geq 0, g + [G]^T \lambda = 0, \lambda_j \geq 0 \) when \( c_j = 0, \lambda_j = 0 \) when \( c_j > 0 \), \( dx^T[H^*]dx \leq 0 \) for any admissible (with respect to the active constraints) \( dx \)
- if \( dx^T[H^*]dx < 0 \) for any admissible \( dx \) is enforced, a set of sufficient conditions is obtained
Indirect Optimization Methods

Optimization of discrete systems

Inequality Constrained Optimization (6)

- $g + [G]^T \lambda = 0$ relates the gradient of $\phi$ to the gradients of the constraints
- with only one constraint ($m = 1$) $\frac{\partial \phi}{\partial x} = -\lambda \frac{\partial c}{\partial x}$, that is, the gradient of $\phi$ must be parallel and opposite ($-\lambda \leq 0$) to the gradient of $c$
- since the gradient of $c$ is perpendicular the constraint $c = 0$, this means that the gradient of $\phi$ must be perpendicular to $c = 0$ pointing toward constraint violation
- for more than one constraint ($m \geq 2$), the gradient of $\phi$ must be a negative linear combination of the constraints gradients
- $- \lambda_i$ represents the change in $\phi$ for a unit change in $c_i$ (assuming a linear approximation)
Indirect Optimization Methods
Optimization of discrete systems

Inequality Constrained Optimization (7)

• a strict equality constraint $c_i = 0$ requires the simultaneous fulfillment of both $c_i \geq 0$ and $-c_i \geq 0$ and therefore, at a feasible point, an equality constraint is always active

• $\lambda_i > 0$ indicates that at a maximum ($d\phi < 0$) admissible $dc_i$ must be positive, i.e., $c_i \geq 0$ is active (the index would be increased for $c_i < 0$)

• $\lambda_i < 0$ indicates that at a maximum ($d\phi < 0$) admissible $dc_i$ must be negative, i.e., $c_i \leq 0$ is active (the index would be increased for $c_i > 0$)
Inequality Constrained Optimization

Examples (1)

- $\phi = -x_1^2 - x_2^2$, $c = x_1 + x_2 - 1 \geq 0$
- Assume $c$ being active, define $\phi^* = -x_1^2 - x_2^2 + \lambda(x_1 + x_2 - 1)$
- Stationary point requires $-2x_1 + \lambda = 0$, $-2x_2 + \lambda = 0$, $x_1 + x_2 - 1 = 0$, that is $x_1 = x_2 = 1/2$ and $\lambda = 1$
- $x^T = (1/2, 1/2)$ is a maximum as $\lambda > 0$ and, for admissible variations $dx^T = (-a, a)$, $dx^T[H]^*dx = -a^2 < 0$
Inequality Constrained Optimization

Examples (2)

• \( \phi = -x_1^2 - x_2^2 \), \( x_1 + x_2 - 1 \leq 0 \), that is, \( c = -x_1 - x_2 + 1 \geq 0 \)

• assume \( c \) being active, define
  \[
  \phi^* = -x_1^2 - x_2^2 + \lambda(-x_1 - x_2 + 1)
  \]

• stationary point requires \(-2x_1 - \lambda = 0, -2x_2 - \lambda = 0, x_1 + x_2 - 1 = 0 \)
  that is \( x_1 = x_2 = 1/2 \) and \( \lambda = -1 \)

• since \( \lambda < 0 \) the constraint is to be removed and the unconstrained optimization provides \( x^T = (0, 0) \) as the maximum
Example

\[ \frac{\partial}{\partial x_1} (x_1 + x_2 > 1) \]

\[ \frac{\partial}{\partial x_2} (x_1 + x_2 < 1) \]

\[ \phi = -3 \]

\[ \phi = 0 \]

\[ \phi = 0.5 \]

\[ \phi = 1 \]

\[ \frac{\partial c}{\partial x} \]

\[ g \]

\[ x_1, x_2 \]

L. Casalino 2015
Continuous Systems

Summary

- necessary conditions for a stationary solution
- Hamilton-Jacobi-Bellman equation
- necessary and sufficient conditions for optimality
Continuous Systems

• $t$ independent variable
• $x$ state variables ($n$-component vector)
• $u$ control variables ($m$-component vector)
• state equations $\dot{x}(t) = \frac{dx}{dt} = f(x, u, t)$ ($n$-component vector of differential equations)
• boundary conditions $\psi(x_0, x_f, t_0, t_f) = 0$ ($q$-component vector of algebraic equations, $q \leq n + 2$)
Bolza Problem

- determine the *extremal path* \( x(t) \) and the corresponding optimal control law \( u(t) \) satisfying
  - differential equations \( \dot{x}(t) = f(x, u, t) \)
  - boundary conditions \( \psi(x_0, x_f, t_0, t_f) = 0 \)

To maximize (or minimize)
\[
J = \phi(x_0, x_f, t_0, t_f) + \int_{t_0}^{t_f} \Phi(x, \dot{x}, t)dt
\]

- Meyer's formulation \( J = \phi(x_0, x_f, t_0, t_f) \) and \( \Phi = 0 \)
- Lagrange's formulation \( J = \int_{t_0}^{t_f} \Phi(x, \dot{x}, t)dt \) and \( \phi = 0 \)
- calculus of variations studies the behavior of functionals \( J \)
- optimal control theory applies CoV to the maximization or minimization of functionals
Equivalence of Formulations

- Lagrange’s formulation can be reduced to Meyer’s formulation by introducing the new state variable $x'$ with $\dot{x}' = \Phi$, $x'_0 = 0$ and $\phi = x'_f$

- Meyer’s formulation can be reduced to Lagrange’s formulation if $\phi = \phi(x_f, t_f)$ observing that

$$\phi(x_f, t_f) = \int_{t_0}^{t_f} [(\partial \phi / \partial x_f) \dot{x} + \partial \phi / \partial t_f] dt$$

or by opportunely introducing additional variables in the general case

- minimization can be turned into maximization by changing the sign of $\phi$ and $\Phi$

- the maximization problem is here considered
Adjoint Variables

- Adjoint variables \( \lambda \) (\( n \)-component vector) and constants \( \mu \) (\( q \)-component vector)
- Augmented index \( J^* = \phi + \mu^T \psi + \int_{t_0}^{t_f} [\Phi + \lambda^T (f - \dot{x})] dt \)
- \( J^* = J \) for any choice of \( \lambda \) and \( \mu \) if state equations and boundary conditions are satisfied
Index Variation (1)

• using a first-order expansion of $J^*$ one has

$$
\begin{align*}
\frac{dJ^*}{dt} &= \left( \frac{\partial \varphi}{\partial t_f} + \mu^T \frac{\partial \psi}{\partial t_f} + \Phi_f + \lambda_f^T (f_f - \dot{x}_f) \right) dt_f + \\
&+ \left( \frac{\partial \varphi}{\partial t_0} + \mu^T \frac{\partial \psi}{\partial t_0} - \Phi_0 - \lambda_0^T (f_0 - \dot{x}_0) \right) dt_0 + \\
&+ \left( \frac{\partial \varphi}{\partial x_f} + \mu^T \frac{\partial \psi}{\partial x_f} \right) dx_f + \left( \frac{\partial \varphi}{\partial x_0} + \mu^T \frac{\partial \psi}{\partial x_0} \right) dx_0 + \\
&+ \int_{t_0}^{t_f} \left[ \frac{\partial (\Phi + \lambda^T f)}{\partial \delta x} + \frac{\partial (\Phi + \lambda^T f)}{\partial \delta u} \right] dt
\end{align*}
$$
Index Variation (2)

- at a given (constant) time the variable variation is $\delta x$
- at the initial (or final) point the total variation $dx = \delta x + \dot{x}dt$ is the sum of variation at constant $t$, i.e., $\delta x$, and variation due to initial (or final) time variation $\dot{x}dt = f dt$
- the Hamiltonian is introduced $H = \Phi + \lambda^T f$
- term in $\delta \dot{x}$ is integrated by parts

$$\int_{t_0}^{t_f} -\lambda^T \delta \dot{x} dt = -\lambda_f^T \delta x_f + \lambda_0^T \delta x_0 + \int_{t_0}^{t_f} \dot{\lambda}^T \delta x dt$$
Index Variation (3)
Index Variation (4)

\[
dJ^* = \left( \frac{\partial \varphi}{\partial t_f} + \mu^T \frac{\partial \psi}{\partial t_f} + H_f \right) dt_f + \\
+ \left( \frac{\partial \varphi}{\partial t_0} + \mu^T \frac{\partial \psi}{\partial t_0} - H_0 \right) dt_0 + \\
+ \left( -\lambda_f^T + \frac{\partial \varphi}{\partial x_f} + \mu^T \frac{\partial \psi}{\partial x_f} \right) dx_f + \\
+ \left( \lambda_0^T + \frac{\partial \varphi}{\partial x_0} + \mu^T \frac{\partial \psi}{\partial x_0} \right) dx_0 + \\
+ \int_{t_0}^{t_f} \left[ \left( \frac{\partial H}{\partial \delta x} + \lambda^T \right) \delta x + \frac{\partial H}{\partial \delta u} \delta u \right] dt
\]
Stationary Point

• necessary condition for $J$ being stationary is $dJ = 0$ for any admissible choice of $\delta x$, $\delta u$, $dx_f$, $dx_0$, $dt_f$, $dt_0$ (that is, differential equations and boundary conditions must be satisfied)

• $\lambda$ and $\mu$ can be chosen to nullify $dJ^*$ for any choice of $\delta x$, $\delta u$, $dx_f$, $dx_0$, $dt_f$, $dt_0$ by nullifying their multiplying coefficients in $dJ^*$

• since $J$ and $J^*$ coincide when the constraints are satisfied, $dJ^* = 0$ for any variation implies $dJ = 0$ for any admissible variation
Equations for Adjoint and Control Variables

• Euler-Lagrange equations ($n$ differential equations for the adjoint variables)

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

• optimal control equations ($m$ algebraic equations for the control variables)

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

• the control equations do not formally depend on the performance index
Boundary Conditions for Optimality

- boundary conditions for optimality ($n + n$ algebraic equations at initial and final point)
  \[
  \lambda_0^T + \frac{\partial \varphi}{\partial x_0} + \mu^T \frac{\partial \psi}{\partial x_0} = 0 \quad -\lambda_f^T + \frac{\partial \varphi}{\partial x_f} + \mu^T \frac{\partial \psi}{\partial x_f} = 0
  \]

- transversality conditions ($1 + 1$ algebraic equations at initial and final time)
  \[
  -H_0 + \frac{\partial \varphi}{\partial t_0} + \mu^T \frac{\partial \psi}{\partial t_0} = 0 \quad H_f + \frac{\partial \varphi}{\partial t_f} + \mu^T \frac{\partial \psi}{\partial t_f} = 0
  \]

- with the imposed boundary conditions $\psi = 0$ ($q$ equations) one has $2n + q + 2$ equations which implicitly determine $q$ adjoint constants ($\mu$), 2 times ($t_0$ and $t_f$) and the initial values for $2n$ differential equations (for $x$ and $\lambda$)
Hamiltonian Time-Derivative

- \[ \frac{dH}{dt} = \frac{\partial H}{\partial t} + \frac{\partial H}{\partial x} \dot{x} + \frac{\partial H}{\partial \lambda} \dot{\lambda} + \frac{\partial H}{\partial u} \dot{u} \]

- one has \( (\frac{\partial H}{\partial \lambda})^T = f = \dot{x}, \dot{\lambda} = - (\frac{\partial H}{\partial x})^T \)

- therefore \( \frac{dH}{dt} = \frac{\partial H}{\partial t} + \frac{\partial H}{\partial u} \dot{u} \)

- when \( H \) does not depend explicitly on time \( (\frac{\partial H}{\partial t} = 0) \) and the optimal control law is adopted \( (\frac{\partial H}{\partial u} = 0) \), the Hamiltonian is constant
Simple Boundary Conditions on State Variables

- variable assigned at $t_0$, $(x_i)_0 - a_i = 0 \Rightarrow (\lambda_i)_0 + \mu = 0$, that is, the corresponding adjoint variable is free
- variable assigned at $t_f$, $(x_i)_f - b_i = 0 \Rightarrow - (\lambda_i)_f + \mu = 0$, that is, the corresponding adjoint variable is free
- one can consider $d(x_i)_0 = 0$ (or $d(x_i)_f = 0$) and drop the equation obtained by nullifying its coefficient
- variable free at $t_0$ and $\partial \phi / \partial (x_i)_0 = 0 \Rightarrow (\lambda_i)_0 = 0$, that is, the corresponding adjoint variable is null
- variable free at $t_f$ and $\partial \phi / \partial (x_i)_f = 0 \Rightarrow (\lambda_i)_f = 0$, that is, the corresponding adjoint variable is null
Simple Boundary Conditions on Time

- initial time assigned $t_0 - t_a = 0 \Rightarrow -H_0 + \mu = 0$, that is, the corresponding Hamiltonian is free
- final time assigned $t_f - t_b = 0 \Rightarrow H_f + \mu = 0$, that is, the corresponding Hamiltonian is free
- one can consider $dt_0 = 0$ (or $dt_f = 0$) and drop the equation obtained by nullifying its coefficient
- initial time free and $\partial\phi/\partial t_0 = 0 \Rightarrow H_0 = 0$, that is, the corresponding Hamiltonian is null
- final time free and $\partial\phi/\partial t_f = 0 \Rightarrow H_f = 0$, that is, the corresponding Hamiltonian is null
Initial State Specified - No Terminal Constraints

Fixed Terminal Time (1)

- \( x_0 = a, \ t_0 = t_a, \ t_f = t_b \)
- \( \phi \) can be regarded as a function of \( x_f \) only, i.e., \( \phi = \phi(x_f) \)
- \( H_0, \ H_f \) and \( \lambda_0 \) are free as \( \delta x_0 = 0, \ \delta t_0 = 0 \) and \( \delta t_f = 0 \)
- optimality conditions provide \( \lambda_f = \left( \frac{\partial \phi}{\partial x_f} \right)^T \) producing a two-point boundary value problem
- initial values \( \lambda_0 \) are free and must be determined to satisfy the optimality conditions at the final point
Indirect Optimization Methods
Continuous systems - CoV

Initial State Specified - No Terminal Constraints

Fixed Terminal Time (2)

- for Euler-Lagrange equations and optimality conditions satisfied (i.e., on an extremal path) non-admissible variations provide

\[ dJ = \lambda_0^T dx_0 - H dt_0 + \int_{t_0}^{tf} \frac{\partial H}{\partial u} \delta u dt \]

- \( \lambda_0 \) is the gradient of \( J \) with respect to changes in \( x_0 \) (influence functions)

\[ \lambda_0 = (\partial J/\partial x_0)^T \]

- \( -H_0 \) is the gradient of \( J \) with respect to changes in \( t_0 \) (time influence function)

\[ -H_0 = \partial J/\partial t_0 \]

- \( \frac{\partial H}{\partial u} \) represents the change in \( J \) for a unit impulse \( \delta u \) (impulse response functions, must be null on an extremal path)
State Variables Specified at Fixed Final Time

State Variables Unspecified at Fixed Initial Time

- \((x_i)_f = b_i\) replaces \((\lambda_i)_f = \frac{\partial \phi}{\partial (x_i)_f}\)
- similarly, if \((x_i)_0\) is free, \((\lambda_i)_0 = 0\) replaces \((x_i)_0 = a_i\) (\(\phi\) is assumed to be a function of \(x_f\) only)
- the change in \(J\) with respect to a unit change in \((x_i)_0\), i.e., \((\lambda_i)_0\), must be null at the optimal initial point if \((x_i)_0\) is free
- if \(\partial \phi / \partial (x_i)_0 \neq 0\) then \(dJ/d(x_i)_0 = \partial \phi / \partial (x_i)_0 + (\lambda_i)_0 = 0\) must be enforced
Unspecified Initial and Final Time

- $t_0 = t_a$ and $t_0 = t_b$ are dropped from $\psi$
- the transversality conditions provide the initial and final time

$$-H_0 + \frac{\partial \varphi}{\partial t_0} + \mu^T \frac{\partial \psi}{\partial t_0} = 0$$

$$H_f + \frac{\partial \varphi}{\partial t_f} + \mu^T \frac{\partial \psi}{\partial t_f} = 0$$
Equality Path Constraints

• \[ \int_{t_0}^{t_f} N(x, u, t) dt = k_b \]

• introduce a new variable \( x_{n+1} \) with differential equation
  \[ \dot{x}_{n+1} = N \]
  and boundary conditions \( (x_{n+1})_0 = 0, \)
  \( (x_{n+1})_f = k_b \)
Equality Constraints on Control Variables

• $C(u, t) = 0$ (only for $m \geq 2$)

• two options
  – eliminate one control variable $u_j$ from $C(u, t) = 0$ and consider the unconstrained problem with $m$ reduced by one
  – use augmented Hamiltonian $H = \Phi + \lambda^T f + \lambda' C$ and derive E-L equations, controls and boundary conditions consequently
Equality Constraints on State and Control Variables

- \( C(x, u, t) = 0 \)
- same approach as before, two options
  - eliminate one control variable \( u_j \) from \( C(x, u, t) = 0 \) and consider the unconstrained problem with \( m \) reduced by one
  - use augmented Hamiltonian \( H = \Phi + \lambda^T f + \lambda' C \) and derive E-L equations, controls and boundary conditions consequently
- one of the state variables can alternatively be eliminated
Equality Constraints on Function of State Variables

- \( S(x, t) = 0 \)
- since \( S \) must be constant \( \dot{S} = \frac{\partial S}{\partial x} \dot{x} + \frac{\partial S}{\partial t} = 0 \)
- if \( \dot{S} \) depends on \( u \) eliminate one control variable \( u_j \) from \( \dot{S} = 0 \) and consider the unconstrained problem with \( m \) reduced by one
- if \( \dot{S} \) is independent of \( u \) either
  - eliminate one state variable \( x_j \) from \( \dot{S} = 0 \) and consider the unconstrained problem with \( n \) reduced by one
  - compute subsequent time derivatives until \( u \) appears
Interior Point Constraints (1)

- functions of state variables specified or variables discontinuities at interior point (time may or may not be specified)
- split the integration into \( f \) intervals at the relevant \( f - 1 \) intermediate points
- the \( j \)-th interval spans from \( t_{(j-1)+} \) to \( t_{(j)-} \), the variable values at the extremities are \( x_{(j-1)+} \) and \( x_{(j)-} \), respectively, \( j = 1, f \)
Interior Point Constraints (2)

• define the performance index $J$

$$J = \phi(x_0, x_{1\pm}, ..., x_f, t_0, t_{1\pm}, ..., t_f) + \sum_{j=i}^{f} \int_{t_{(j-1)+}}^{t_{(j)-}} \Phi(x, \dot{x}, t) dt$$

• add conditions at intermediate points and build

$$\psi = \psi(x_{(j-1)+}, x_{(j)-}, t_{(j-1)+}, t_{(j)-}), j = 1, ..., f$$

• define the augmented performance index $J^*$

$$J^* = \phi + \mu^T \psi + \sum_{j=i}^{f} \int_{t_{(j-1)+}}^{t_{(j)-}} \Phi + \lambda^T (f - \dot{x}) dt$$

and derive Euler-Lagrange and control equations, and boundary conditions for optimality.
**Interior Point Constraints (3)**

- boundary conditions for optimality \((j = 1, f)\)

\[
\lambda_T^{(j-1)+} + \frac{\partial \phi}{\partial x_{(j-1)+}} + \mu_T \frac{\partial \psi}{\partial x_{(j-1)+}} = 0 \text{ (start of phase } j) \\
-\lambda_T^{(j)-} + \frac{\partial \phi}{\partial x_{(j)-}} + \mu_T \frac{\partial \psi}{\partial x_{(j)-}} = 0 \text{ (end of phase } j) 
\]

- transversality conditions \((j = 1, f)\)

\[
-H_{(j-1)+} + \frac{\partial \phi}{\partial t_{(j-1)+}} + \mu_T \frac{\partial \psi}{\partial t_{(j-1)+}} = 0 \text{ (start of phase } j) \\
H_{(j)-} + \frac{\partial \phi}{\partial t_{(j)-}} + \mu_T \frac{\partial \psi}{\partial t_{(j)-}} = 0 \text{ (end of phase } j) 
\]

- a multipoint boundary value problem must be solved
Interior Point Constraints (4)

- at a generic point $j$ (end of phase $j$, start of phase $j+1$)

$$
\lambda^T_{j+} + \frac{\partial \varphi}{\partial x_{j+}} + \mu^T \frac{\partial \psi}{\partial x_{j+}} = 0, \ j = 0, \ldots, f - 1
$$

$$
-\lambda^T_{j-} + \frac{\partial \varphi}{\partial x_{j-}} + \mu^T \frac{\partial \psi}{\partial x_{j-}} = 0, \ j = 1, \ldots, f
$$

$$
-H_{j+} + \frac{\partial \varphi}{\partial t_{j+}} + \mu^T \frac{\partial \psi}{\partial t_{j+}} = 0, \ j = 0, \ldots, f - 1
$$

$$
H_{j-} + \frac{\partial \varphi}{\partial t_{j-}} + \mu^T \frac{\partial \psi}{\partial t_{j-}} = 0, \ j = 1, \ldots, f
$$
Interior Point Constraints (5)

- variable continuous and specified at intermediate time $t_j$,
  
  \[
  (x_i)_{j+} - c_i = 0, \quad (x_i)_{j-} - c_i = 0, \quad \phi \text{ independent of } (x_i)_{j+} \text{ and } (x_i)_{j-} \Rightarrow (\lambda_i)_{j+} - (\lambda_i)_{j-} = -(\mu_1 + \mu_2), \]
  
  the corresponding adjoint variable has a free discontinuity

- variable continuous but unspecified at intermediate time $t_j$,

  \[
  (x_i)_{j+} - (x_i)_{j-} = 0, \quad \phi \text{ independent of } (x_i)_{j+} \text{ and } (x_i)_{j-} \Rightarrow (\lambda_i)_{j+} + \mu = 0 \text{ and } -(\lambda_i)_{j-} - \mu = 0, \]

  the corresponding adjoint variable is continuous $$(\lambda_i)_{j+} = (\lambda_i)_{j-}$$
Interior Point Constraints (6)

• intermediate time fixed (and continuous), \( t_{j^+} - t_c = 0, \ t_{j^-} - t_c = 0 \), \( \phi \) independent of \( t_{j^+} \) and \( t_{j^-} \) ⇒ \(-H_{j^+} + \mu_1 = 0, \ H_{j^-} + \mu_2 = 0\), the Hamiltonian has a free discontinuity

• intermediate time free and continuous, \( t_{j^+} - t_{j^-} = 0 \), \( \phi \) independent of \( t_{j^+} \) and \( t_{j^-} \) ⇒ \(-H_{j^+} + \mu = 0, \ H_{j^-} - \mu = 0\), the Hamiltonian is continuous \( H_{j^+} = H_{j^-} \)
Inequality Constraints

- $C(x, u, t) \geq 0$
- assume the constraint being active $C(x, u, t) = 0$
- define augmented Hamiltonian $H = \Phi + \lambda^T f + \lambda'C$ and solve the equality constrained problem
  - $\lambda' \geq 0$ is required when $C = 0$ (active constraint)
  - $\lambda' = 0$ is required when $C > 0$ (inactive constraint)
- constraint to be removed if $\lambda' < 0$
- main issue: determine the constrained interval(s)
Hamilton-Jacobi-Bellman Equations (1)

- for a given initial point \( x_0, t_0 \) and final constraints \( \psi(x_f, t_f) = 0 \) the optimal path provides the maximum index \( J = J^o \) with control variables \( u^o(t) \) (unicity of the optimal path is here supposed)

- Hamilton-Jacobi theory provides differential equations concerning \( J^o \) and the optimal control \( u^o \), extended by Bellman to discrete systems (dynamic programming)

- any point \( x, t \) on the optimal path may be considered as the initial point

- starting from \( x, t \) on the optimal path the maximization of \( J \) provides the same control law \( u^o(t) \)
**Hamilton-Jacobi-Bellman Equations (2)**

- for a generic starting point $x, t$, one has
  $$J^o(x, t) = \max \left\{ \phi(x_f, t_f) + \int_t^{t_f} \Phi(x, u, \tau) d\tau \right\}$$
  with boundary condition $J^o(x_f, t_f) = \phi(x_f, t_f)$ on the terminal hypersurface $\psi(x_f, t_f) = 0$
- maximization with respect to $u$
- assume $J^o$ continuous with continuous first and second derivatives
- if a generic (possibly nonoptimal) control is used from $x, t$ to $x + f(x, u, t) \Delta t, t + \Delta t$, and then the optimal control is used up to the final point, the performance index will be
  $$J'(x, t) = \Phi(x, u, t) \Delta t + J^o[x + f(x, u, t) \Delta t, t + \Delta t]$$
Hamilton-Jacobi-Bellman Equations (3)
H **Hamilton-Jacobi-Bellman Equations** (4)

- \( J' = J^o \) only if the optimal control is used from \( t \) to \( t + \Delta t \), that is,
  \[
  J^o(x, t) = \max \{ J^o[x + f(x, u, t)\Delta t, t + \Delta t] + \Phi(x, u, t)\Delta t \} 
  \]

- first order expansion provides \( J^o(x, t) = \)
  \[
  \max \left\{ J^o(x, t) + \frac{\partial J^o}{\partial x} f(x, u, t)\Delta t + \frac{\partial J^o}{\partial t} \Delta t + \Phi(x, u, t)\Delta t \right\} 
  \]

- for \( \Delta t \to 0 \) and considering that \( J^o \) and \( \partial J^o / \partial t \) do not depend on \( u \) one has
  \[
  -\frac{\partial J^o}{\partial t} = \max \left\{ \frac{\partial J^o}{\partial x} f(x, u, t) + \Phi(x, u, t) \right\} 
  \]
Hamilton-Jacobi-Bellman Equations (5)

- according to the definition of influence functions
  \[ dJ^o = \lambda^T \delta x - H \delta t \]
- therefore
  \[ \frac{\partial J^o}{\partial x} = \lambda^T \quad \frac{\partial J^o}{\partial t} = -H \]
  and
  \[ \phi + \frac{\partial J^o}{\partial x} f = \phi + \lambda^T f = H \]
- comparison provides the Hamilton-Jacobi-Bellman equation
  \[ -\frac{\partial J^o}{\partial t} = H^o \left( x, \frac{\partial J^o}{\partial x}, t \right) \]
  \[ H^o \left( x, \frac{\partial J^o}{\partial x}, t \right) = \max \left\{ H \left( x, \frac{\partial J^o}{\partial x}, u, t \right) \right\} = \max \left\{ H \left( x, \lambda, u, t \right) \right\} \]
- maximization with respect to \( u \) given \( x, \lambda, \) and \( t \)
Hamilton-Jacobi-Bellman Equations (6)

• HJB equation state that the optimal controls must maximize the Hamiltonian over the whole set of admissible controls (global condition)

• the same results was obtained independently by Pontryagin and is known as Pontryagin’s maximum principle (PMP)

• in the absence of bounds on state and control variables, HJB implies $H_u = 0$ and $H_{uu}$ negative definite (local conditions)

• HJB also implies that Euler-Lagrange equations and the boundary conditions for optimality must be verified
Hamilton-Jacobi-Bellman Equations (7)

- for any final point $x_f, t_f$ satisfying $\psi(x_f, t_f) = 0$ ($n + 1$ variables, $q \leq n + 1$ conditions), backward integration of HJB equation for any choice of $\lambda_f, \mu$ satisfying the boundary conditions for optimality and transversality condition at the final point ($n + q$ variables, $n + 1$ conditions) provides all the possible optimal paths and the corresponding control laws (*field of extremals*).

- in general, only one optimal path will pass through a given point $x, t$ and a unique optimal control $u^0(x, t)$ will be associated with each point (*optimal feedback control law*).

- main issue: *curse of dimensionality* - there are $n$ free parameters and it is very difficult to find and store all the possible solutions.
Bang-Bang Controls (1)

- if $H$ is linear with respect to a control variable $u_j$ then $\frac{\partial H}{\partial u_j} = 0$ does not contain $u_j$ (indeterminate)
- HJB equation or PMP show that the control value which maximizes $H$ must be adopted
- the problem has sense only if $u_j$ is bounded
  \[(u_j)_{\text{min}} \leq u_j \leq (u_j)_{\text{max}}\]
- the switching function $S_F = \frac{\partial H}{\partial u_j}$ is introduced
  - $u_j = (u_j)_{\text{max}}$ when $S_F > 0$
  - $u_j = (u_j)_{\text{min}}$ when $S_F < 0$
  - singular arc when $S_F = 0$ over a finite interval ($u_j$ obtained by nullifying the subsequent time derivatives of $S_F$)
Bang-Bang Controls (2)

\[ S_F < 0 \quad \text{for} \quad u = u_{\text{min}} \]

\[ S_F > 0 \quad \text{for} \quad u = u_{\text{max}} \]
**Necessary Conditions for a Maximum (1)**

- Stationary feasible point: state and Euler-Lagrange equations verified for $t_0 \leq t \leq t_f$, imposed conditions and necessary conditions for optimality verified at initial, final, and intermediate points.

- Weierstrass condition: $H$ maximized by the optimal controls (global maximum) for $t_0 \leq t \leq t_f$.

- For a local maximum Weierstrass condition is replaced by $\partial H/\partial u = 0$ and Legendre-Clebsch condition $H_{uu}$ negative semidefinite (usually, a minimization problem is considered and $H_{uu}$ positive semidefinite is imposed, convexity condition).
Necessary Conditions for a Maximum (2)

- normality condition: $\hat{Q}$ positive semidefinite, $\alpha \leq 0$ for $t_0 \leq t < t_f$
- Jacobi condition: $\hat{S} - \hat{R}\hat{Q}^{-1}\hat{R}^T$ finite for $t_0 < t < t_f$, that is, no conjugate points exist on the path (if the matrix becomes infinity some $\delta x$ must be restricted and no maximization is possible as infinite equivalent solutions exist)
- Riccati equation $\dot{S} = -C - A^TS - SA + SBS = -Sf_x - f_x^TS - H_{xx} + (Sf_u + H_{ux}^T)H_{uu}^{-1}(H_{ux} + f_u^TS)$
  $\dot{R} = -(A^T - SB)R$ and $\dot{Q} = R^TBR$
  $\dot{m} = -(A^T - SB)m$  $\dot{n} = R^T B n$  $\dot{\alpha} = m^T B m$
- with $A = f_x - f_u H_{uu}^{-1} H_{ux}$, $B = f_u H_{uu}^{-1} f_u^T$, $C = H_{xx} - H_{xu} H_{uu}^{-1} H_{ux}$
Sufficient Conditions for a Maximum (1)

- Stationary feasible point: state and Euler-Lagrange equations verified for \( t_0 \leq t \leq t_f \), imposed and necessary conditions for optimality verified at initial final and intermediate points.

- Weierstrass condition: \( H \) maximized by the optimal controls (global maximum) for \( t_0 \leq t \leq t_f \).

- For a local maximum, Weierstrass condition is replaced by \( \partial H/\partial u = 0 \) and strengthened Legendre-Clebsch condition \( H_{uu} \) negative definite.
Sufficient Conditions for a Maximum (2)

- normality condition: $\hat{\mathcal{Q}}$ positive definite, $\alpha < 0$ for $t_0 \leq t < t_f$
- Jacobi condition: $\hat{\mathcal{S}} - \hat{\mathcal{R}}\hat{\mathcal{Q}}^{-1}\hat{\mathcal{R}}^T$ finite for $t_0 \leq t < t_f$, that is, no conjugate points exist on the path (if the matrix becomes infinity some $\delta x$ must be restricted and no maximization is possible as infinite equivalent solutions exist)
Practical Approach

• rely on intuition
• check results with perturbed conditions for optimality
Optimization Problem Formulation

- trajectory split into arcs
- homogeneous control law in each arc
- constraints and discontinuities at the arcs boundaries
Boundary Value Problem

• application of the theory of optimal control produces a multipoint boundary value problem (BVP)

• some constants (e.g., relevant times) and the initial values of some of the state and adjoint variables are unknown

• boundary conditions at the relevant points must be satisfied

• an iterative procedure based on Newton’s method can be adopted
BVP Formulation (1)

- $t$ independent variable
- $x$ state variables $\frac{dx}{dt} = f$
- $\lambda$ adjoint variables $\frac{d\lambda}{dt} = - (\frac{\partial H}{\partial x})^T$
- $u$ control variables obtained as functions of $x$ and $\lambda$ by maximizing the Hamiltonian
- boundary conditions concerning $x$, $\lambda$ and $t$ at initial, final and intermediate points
BVP Formulation (2)

- independent variable transformation: in the $j$-th arc
  \[ \varepsilon = j - 1 + \frac{t - t_{j-1}}{t_j - t_{j-1}} = j - 1 + \frac{t - t_{j-1}}{\tau_j} \]
- in the $j$-th arc $\varepsilon$ varies between $j - 1$ and $j$
- $\frac{dx}{d\varepsilon} = \tau_j \frac{dx}{dt}$
- $\frac{d\lambda}{d\varepsilon} = \tau_j \frac{d\lambda}{dt}$
- the relevant times or arc time-lengths are additional constant parameters
- constant parameters $y$ with equation $\frac{dy}{d\varepsilon} = 0$
BVP Formulation (3)

- define vector of unknowns
  \[ z = \begin{pmatrix} x \\ \lambda \\ y \end{pmatrix} \]
- differential equations \( \partial z / \partial \epsilon = g(z, \epsilon) \)
- collect values at relevant boundaries
  \[ s = (z_{0+}, z_{1\pm}, \ldots, z_{(f-1)\pm}, z_{f-}) \]
- boundary conditions in the form \( \Psi(s) = 0 \)
- assume initial tentative values \( z_0 = p \)
- Newton’s method to reduce the error on boundary conditions
BVP Solution (1)

- at the $r$-th iteration correct the initial values
  
  \[ p^{r+1} = p^r + \Delta p \quad \text{with} \quad \Delta p = p^{r+1} - p^r = -[\partial \Psi / \partial p]^{-1} \Psi^r \]

- \([\partial \Psi / \partial p] = [\partial \Psi / \partial s] [\partial s / \partial p] \]
- \([\partial \Psi / \partial s]\) by derivation
- \([\partial s / \partial p]\) collects values at the boundaries of \([\partial z / \partial p]\)
- matrix \([\partial z / \partial p]\) by integration of the homogeneous differential system
  
  \[
  \begin{bmatrix}
  \frac{\partial z}{\partial p} \\
  \frac{\partial^2 z}{\partial p^2}
  \end{bmatrix} = \begin{bmatrix}
  \partial g \\
  \partial z
  \end{bmatrix} \begin{bmatrix}
  \partial z \\
  \partial p
  \end{bmatrix}
  \]
BVP Solution (2)

• the matrix $[\partial \Psi / \partial p]$ can also be obtained numerically
• vary $p_i$ by small quantity $\delta p_i$
• integrate equations and compute change of variable values at relevant points $\delta s$
• compute change in error on boundary conditions $\delta \Psi$
• approximate by linearization the $i$-th column of $[\partial \Psi / \partial p]$
  $[\partial \Psi / \partial p_i] = \delta \Psi / \delta p_i$
BVP Solution (3)

- Linearization may introduce errors that can prevent convergence
  - Perform check on error variation and reduce the parameter correction by posing $p^{r+1} = p^r + \Delta p/2$ when $\Psi^{r+1} > K_1 \Psi^r$
  - Use a reduced parameter correction $\Delta p = -K_2 [\partial \Psi / \partial p]^{-1} \Psi^r$ when the errors are large
  - $K_1 = 2$ usually provides good results
  - $0.01 \leq K_2 \leq 1$ usually provides good results (the larger values when the solution is close to the optimal one)
State Equations

- two-body problem with propulsion - vectorial formulation

\[
\frac{dr}{dt} = V
\]
\[
\frac{dV}{dt} = g + \frac{T}{m}
\]
\[
\frac{dm}{dt} = -\frac{T}{c}
\]

- Hamiltonian

\[
H = \lambda_r^T V + \lambda_V^T (g + T/m) - \lambda_m T/c
\]
Euler-Lagrange Equations

- for thrust independent of state variables

\[
\frac{d\lambda_r}{dt} = - \left[ \frac{\partial g}{\partial r} \right]^T \lambda_V
\]

\[
\frac{d\lambda_V}{dt} = -\lambda_r
\]

\[
\frac{d\lambda_m}{dt} = \frac{\lambda_V T}{m^2}
\]
Indirect Optimization Methods

Indirect trajectory optimization

Optimal Thrust

• PMP states that the optimal controls must maximize $H$ (given state and adjoint variables)
• thrust direction parallel to the velocity adjoint vector $\lambda_V$ (also named primer vector), i.e., $T = T\lambda_V/\lambda_V$
• the Hamiltonian becomes
  $H = \lambda_r^T V + \lambda_V^T g + T(\lambda_V/m - \lambda_m/c)$
• maximum thrust $T_{max}$ when the switching function $S_F = \lambda_V/m - \lambda_m/c$ is positive
• minimum thrust (typically, 0) when the switching function $S_F = \lambda_V/m - \lambda_m/c$ is negative
• $S_F = 0$ at the thrust switching points
Spherical Reference Frame

- $k$: north
- $i$: zenith
- $j$: east
- $r$: ecliptic or equatorial plane
Scalar State Equations

- two-body problem with propulsion

\[
\begin{align*}
\dot{r} &= u \\
\dot{\theta} &= v/(r \cos \phi) \\
\dot{\phi} &= w/r \\
\dot{u} &= -\mu/r^2 + (v^2 + w^2)/r + T \sin \gamma_T/m \\
\dot{v} &= (-uv + vw \tan \phi)/r + T \cos \gamma_T \cos \psi_T/m \\
\dot{w} &= (-uw - v^2 \tan \phi)/r + T \cos \gamma_T \sin \psi_T/m \\
\dot{m} &= -T/c
\end{align*}
\]

- \( H = \lambda_r u + \lambda_\theta v/(r \cos \phi) + \lambda_\phi w/r \\
  + \lambda_u[-\mu/r^2 + (v^2 + w^2)/r + T \sin \gamma_T/m] \\
  + \lambda_v[(-uv + vw \tan \phi)/r + T \cos \gamma_T \cos \psi_T/m] \\
  + \lambda_w[(-uw - v^2 \tan \phi)/r + T \cos \gamma_T \sin \psi_T/m] - \lambda_m T/c \)
Scalar Euler-Lagrange Equations

- thrust independent of state variables (optimal thrust direction is assumed)

\[
\begin{align*}
\dot{\lambda}_r &= \left[ \lambda_{\theta} \frac{v}{\cos \phi} + \lambda_{\phi} \right.
\left. w + \lambda_u \left( -\frac{2}{r} + v^2 + w^2 \right) + 
\left. \lambda_v \left( -uv + vw \tan \phi \right) + \lambda_w \left( -uw - v^2 \tan \phi \right) \right] / r^2 \\
\dot{\lambda}_{\theta} &= 0 \\
\dot{\lambda}_\phi &= \left( -\lambda_{\theta} \sin \phi - \lambda_v \right.
\left. vw + \lambda_w v^2 \right) / \left( r \cos^2 \phi \right) \\
\dot{\lambda}_u &= \left( -\lambda_r r + \lambda_v u + \lambda_w w \right) / r \\
\dot{\lambda}_v &= \left[ -\lambda_{\theta} \frac{1}{\cos \phi} - 2\lambda_u v + \lambda_v \left( u - w \tan \phi \right) + 2\lambda_w v \tan \phi \right] / r \\
\dot{\lambda}_w &= \left( -\lambda_{\phi} - 2\lambda_u w - \lambda_v \tan \phi + \lambda_w u \right) / r \\
\dot{\lambda}_m &= T \lambda_V / m^2
\end{align*}
\]
Thrust Direction

\[ i \quad (\text{zenith}) \]

\[ j \quad (\text{east}) \]

\[ T \]

\[ k \quad (\text{north}) \]

\[ \gamma_T \]

\[ \psi_T \]
Optimal Thrust Angles

- $\gamma_T$ thrust elevation angle, $\psi_T$ thrust heading angle
- optimal values by posing $\partial H/\partial \gamma_T = 0$ and $\partial H/\partial \psi_T = 0$
- one obtains
  \[
  \lambda_u \cos \gamma_T - (\lambda_v \cos \psi_T + \lambda_w \sin \psi_T) \sin \gamma_T = 0 \\
  -\lambda_v \sin \psi_T + \lambda_w \cos \psi_T = 0
  \]
- these equations provide
  \[
  \sin \gamma_T = \frac{\lambda_u}{\lambda_V} \\
  \cos \gamma_T \cos \psi_T = \frac{\lambda_v}{\lambda_V} \\
  \cos \gamma_T \sin \psi_T = \frac{\lambda_w}{\lambda_V}
  \]
- primer vector magnitude $\lambda_V = \sqrt{\lambda_u^2 + \lambda_v^2 + \lambda_w^2}$
Thrust Management (1)

- the thrust magnitude is discontinuous
- the thrust level can be decided during integration according to the sign of $S_F$
- the thrust magnitude at a given instant and the result of the integration may depend on the integration step
- convergence difficulties may be experienced because of scarce accuracy in the evaluation of the error gradients
- variable-step integration strategies usually provide better results
Indirect Optimization Methods

Thrust Management (2)

- an alternative approach consists in assuming “a priori” the switching structure, i.e., the succession of thrusting and coasting arcs, by splitting the trajectory into phases
- the thrust level during each phase is given and an accurate integration is allowed
- the switching times are additional unknown parameters of the BVP problem
- for free switching time, boundary conditions for optimality prescribe the Hamiltonian continuity $H_+ = H_-$
- for continuous state and adjoints variables (no constraints on the variables at the impulse) $S_F = 0$ is required

\[
\lambda^T_{r+} V_+ + \lambda^T_{V+} g_+ + T_+ S_{F+} = \lambda^T_{r-} V_- + \lambda^T_{V-} g_- + T_- S_{F-}
\]
Thrust Management (3)

- the switching structure must be assigned a priori, based on experience and/or intuition
- the approach that decides the thrust magnitude during integration should be tried first
- if convergence difficulties are experienced, the switching function history may suggest the switching structure (thrusting arcs when $S_F > 0$, coasting arcs when $S_F < 0$)
- for a given switching structure, the solution must be checked in the light of PMP (verify that $S_F > 0$ during a thrusting phase and $S_F < 0$ during a coasting phase)
- thrusting or coasting arcs must be added or removed when PMP is violated
Switching Structure - Examples
Terminal Conditions (1)

- for maximum final mass $\lambda_{mf} = 1$
- for prescribed final position and velocity the final values of the corresponding adjoint variables are free
- for prescribed final position and free velocity the final values of the velocity adjoint variables are null and $\lambda_{Vf} = 0$ and, therefore, $S_{Ff} < 0$ (the thruster is off, since there is no use in thrusting if the velocity is free)
Terminal Conditions (2)

- minimum time, free final mass $\lambda_{mf} = 0$
- $\lambda_m \leq 0$ (in fact, its derivative is always positive) and $S_F > 0$ at any time (the thruster is always on)
Terminal Conditions (3)

- rendezvous requires $r_f = r_p(t_f)$, $V_f = V_p(t_f)$
- one has $\frac{\partial \psi}{\partial t_f} = \frac{\partial \psi}{\partial (r_p)} (\dot{r}_p) + \frac{\partial \psi}{\partial (V_p)} (\dot{V}_p) = -\frac{\partial \psi}{\partial r_f} (V_p) + \frac{\partial \psi}{\partial V_f} g_f$ (with $\dot{r}_p = V_p$ and $\dot{V}_p = g$)
- boundary conditions for optimality and transversality condition combined provide $H - \lambda^T_v V_p - \lambda^T_r g = 0$ that is $T S_F = 0$, since $x = x_p$ at the rendezvous
- since the rendezvous is actually obtained when the thruster is switched on ($T > 0$), then $S_F = 0$ is usually imposed at the final point
Indirect Optimization Methods

Indirect trajectory optimization

Optimal Phasing

- Assume optimal phasing between the relevant planets to define unconstrained optimum and find tentative solutions for the rendezvous problem.
- "Move" the target planet to the required position by evaluating position and velocity at the relevant time $t$ as the values assumed at $t + t^*$, with $t^*$ optimization parameter.
- Impose $r_f = r_p(t_f + t^*)$, $V_f = V_p(t_f + t^*)$.
- Boundary conditions for optimality require $\mu^T(\partial \psi/\partial t^*) = 0$.
- Since $\partial \psi/\partial t^* = \partial \psi/\partial t_f$, the transversality condition becomes $H_f = 0$, as in a time-free problem.
- Similar conditions are obtained if optimal phasing is assumed for an intermediate planet (flyby).
**Impulsive Thrust (1)**

- an impulse may me considered as a discontinuity in velocity and mass

\[
\Delta V = V_+ - V_-, \quad m_+ = m_- \exp(-\Delta V/c)
\]

- boundary conditions for optimality require

\[
\begin{align*}
-\lambda_{V_-} + (\mu m_-/c) \exp(-\Delta V/c) \Delta V/\Delta V &= 0 \\
\lambda_{V_+} - (\mu m_-/c) \exp(-\Delta V/c) \Delta V/\Delta V &= 0 \\
-\lambda_{m_-} - \mu \exp(-\Delta V/c) &= 0 \\
\lambda_{m_+} + \mu &= 0
\end{align*}
\]

- this implies

\[
\begin{align*}
\lambda_{V_-} &= \lambda_{V_+} = \lambda_V \Delta V/\Delta V \text{ (the impulse is parallel to the primer vector, which is continuous)} \\
\lambda_{m_+} &= \lambda_{m_-} \exp(+\Delta V/c) \text{ (the product } m\lambda_m \text{ is continuous)} \\
\lambda_V &= m\lambda_m/c \text{ (the switching function is null)}
\end{align*}
\]
Indirect Optimization Methods

Indirect trajectory optimization

Impulsive Thrust (2)

- if the impulse time is free, transversality conditions require the continuity of the Hamiltonian through the impulse
  \( H_+ = H_- \) (note that the mass and the corresponding adjoint variable do not appear in \( H \), as no thrust is applied except at the impulse)
  \[
  \lambda^T_r V_+ + \lambda^T_v g_+ = \lambda^T_r V_- + \lambda^T_v g_-
  \]
- since position, gravity, and adjoint variable are continuous, one obtains
  \[
  \lambda^T_r (V_+ - V_-) = \lambda^T_r \Delta V = 0
  \]
- this implies \( \lambda^T_r \lambda_v = 0 \), that is \( \dot{\lambda}_v \lambda_v = \lambda_v \dot{\lambda}_v = 0 \)
- the derivative of the primer magnitude must be null, and, therefore, also \( \dot{S}_F = 0 \)
Impulsive Thrust (3)

- the switching structure (i.e., number and sequence of impulses) must be assumed a priori
- impulses must be added when the switching function becomes positive, i.e., $\lambda_V > m\lambda_m/c$ (note that the larger is $c$, the smaller is the value that $\lambda_V$ must assume at the impulse)
- impulse magnitude tends to vanish when the impulse is not required (signaling that the impulse should be removed from the switching structure)
Indirect Optimization Methods

Interplanetary Trajectories - Departure (1)

- the dimension of and the time spent inside the sphere of influence is neglected
- impulsive departure from parking orbit
- hyperbolic excess velocity $V_\infty = V_0 - V_p$ with $V_p$ planet heliocentric velocity and $V_0$ spacecraft heliocentric velocity at escape
- impulse magnitude $\Delta V = \sqrt{V_\infty^2 + V_{esc}^2} - V_{orb}$, with $V_{esc}$ and $V_{orb}$ escape velocity and orbital velocity at departure, respectively (the departure point should be the periapsis of the parking orbit)
- escape mass $m_0 = m_{orb} \exp(-\Delta V/c)$
- position $r_0 = r_p$
Interplanetary Trajectories - Departure (2)

- boundary conditions for optimality require
  \[ \lambda V_0 - (\mu m_{orb}/c) \exp(-\Delta V/c) V_\infty/\sqrt{V_\infty^2 + V_{esc}^2} = 0 \]
  \[ \lambda m_0 + \mu = 0 \]

- this implies
  \[ \lambda V_0 = \lambda V_0 V_\infty/\sqrt{V_\infty^2 + V_{esc}^2} \quad \text{(the hyperbolic excess velocity is parallel to the primer vector)} \]
  \[ \lambda V_0 = (m_0 \lambda m_0/c)V_\infty/\sqrt{V_\infty^2 + V_{esc}^2} \]

- in comparison with a deep-space impulse, where
  \[ \lambda V = m\lambda m/c, \text{ an impulse inside a sphere of influence occurs with a negative switching function, i.e., it is more convenient to use thrust close to a planet than outside its sphere of influence} \]
Interplanetary Trajectories - Arrival

- specular maneuver compared to departure
- under the same assumptions $r_f = r_p$, $V_\infty = V_f - V_p$,
  \[
  \Delta V = \sqrt{V_\infty^2 + V_{esc}^2} - V_{orb}
  \]
- mass delivered into the parking orbit $m_{orb} = m_f \exp(-\Delta V/c)$ is maximized
- boundary conditions are $\lambda_{Vf} = -\lambda_{Vf} V_\infty / \sqrt{V_\infty^2 + V_{esc}^2}$ (the hyperbolic excess velocity is parallel and opposite to the primer vector)
  \[
  \lambda_{Vf} = (m_f \lambda_m/c) V_\infty / \sqrt{V_\infty^2 + V_{esc}^2}
  \]
Interplanetary Trajectories - Flyby (1)

- the dimension of and the time spent inside the sphere of influence is neglected
- planet intercept requires $r_+ = r_- = r_p$
- planetocentric energy conservation requires conservation of hyperbolic excess velocity magnitude
  \[ V_{\infty+}^2 = (V_+ - V_p)^2 = V_{\infty-}^2 = (V_- - V_p)^2 \]
- the $V_\infty$ turn depends on the height of the hyperbola periapsis $R_p$
Interplanetary Trajectories - Flyby (2)

- rotation angle
  \[ \delta = \pi - 2\Phi \]
- \( \cos \Phi = \sin(\delta/2) = 1/e \)
- \( \sin(\delta/2) = \frac{\mu_p/R_p}{V_\infty^2 + \mu_p/R_p} \)
- \( b = r_p \sqrt{V_\infty^2 + 2\mu_p/R_p} V_\infty \)
Indirect Optimization Methods

Indirect trajectory optimization

Free-Height Flyby

- imposed boundary conditions $r_+ = r_- = r_p$, $V_{\infty+}^2 = V_{\infty-}^2$, and $m_+ = m_-$

- boundary conditions for optimality
  $\lambda_{r-} - \lambda_{r+} = \mu_1$ (the position adjoint vector presents a free discontinuity)
  $\lambda_{ms+} = \lambda_{ms-}$ (the mass adjoint variable is continuous)
  $\lambda_{V-} = -2\mu_2 V_{\infty-}$ \quad $\lambda_{V+} = -2\mu_2 V_{\infty+}$

- this implies that the primer must be parallel to the hyperbolic excess velocity before and after the flyby and that its magnitude is continuous
Minimum-Height Flyby (1)

• additional boundary condition on the velocity turn
  \[ V_{\infty}^T + V_{\infty}^- = - \cos 2\phi \ V_{\infty}^2^- \] with \( \cos \phi = V_p^2 / (V_{\infty}^2^- + V_p^2) \)
  and \( V_p = \sqrt{\mu_p / R_p} \) circular velocity at the minimum allowable distance from the planet

• boundary conditions for optimality state that the position adjoint vector presents a free discontinuity and the mass adjoint variable is continuous

• \( \lambda_{V_-} = -2\mu_2 V_{\infty}^- + \mu_4 V_{\infty}^+ + 2\mu_4 B \ V_{\infty}^- \)
  \( \lambda_{V^+} = -2\mu_2 V_{\infty}^+ - \mu_4 V_{\infty}^- \) with
  \( B = \cos 2\phi - A \sin 2\phi \) and \( A = \frac{d\phi}{dV_{\infty}^-} V_{\infty}^- = \frac{2}{\tan \phi} \frac{V_{\infty}^2^-}{V_{\infty}^2^- + V_p^2} \)
Minimum-Height Flyby (2)

- one has \( \lambda_{V^-} \times V_{\infty^-} = -\mu_4 V_{\infty^+} \times V_{\infty^-} \) and
  \( \lambda_{V^+} \times V_{\infty^+} = \mu_4 V_{\infty^-} \times V_{\infty^+} = \lambda_{V^-} \times V_{\infty^-} \) the primer component perpendicular to \( V_\infty \) is the same before and after the flyby (\( \lambda_{V^+}^{perp} = \lambda_{V^-}^{perp} \))

- also \( (\lambda_{V^-} \times V_{\infty^-}) \cdot V_{\infty^+} = 0 \) and \( (\lambda_{V^+} \times V_{\infty^+}) \cdot V_{\infty^-} = 0 \)
  that is \( \lambda_V \cdot (V_{\infty^+} \times V_{\infty^-}) = 0 \) the primer lies on the flyby plane before and after the flyby

- finally \( \lambda_{V^-} \cdot V_{\infty^-} - 2\mu_2 V_{\infty^-}^2 + \mu_4 V_{\infty^+} \cdot V_{\infty^-} + 2\mu_4 B V_{\infty^-}^2 = 0 \)
  and \( -\lambda_{V^+} \cdot V_{\infty^+} + 2\mu_2 V_{\infty^+}^2 + \mu_4 V_{\infty^-} \cdot V_{\infty^+} = 0 \)
  which provide the change in the primer component parallel to \( V_\infty \),
  that is \( \lambda_{V^+}^{par} = \lambda_{V^-}^{par} + 2A \lambda_{V^-}^{perp} \)
Minimum-Height Flyby (3)

- the direction of the primer component perpendicular to the hyperbolic excess velocity must be checked to ensure the height constraint requirement.
- the constraint is required when the $\lambda_V$ perpendicular component is directed toward the planet, i.e., is in the direction of the velocity rotation.
- the constraint must be removed when it is directed opposite to the velocity turn.
- it is sufficient to check the sign of $(\lambda_V \times V_{\infty}) \cdot (V_{\infty+} \times V_{\infty-})$, which must be positive.
Indirect Optimization Methods

Indirect trajectory optimization

Minimum-Height Flyby (4)
Flyby - Transversality Conditions (1)

- one has $t_+ = t_-$ and, as far as the condition at flyby on position and velocity are concerned (only the terms $r - r_p$ and $V - V_p$ appear), one has

$$\frac{\partial \psi}{\partial t} = -\left(\frac{\partial \psi}{\partial r}\right)\dot{r}_p - \left(\frac{\partial \psi}{\partial V}\right)\dot{V}_p$$

- the transversality condition becomes

$$H_- - \mu^T(\frac{\partial \psi}{\partial r_-} V_p + \frac{\partial \psi}{\partial V_-} g) - \mu_T = 0$$

$$-H_+ - \mu^T(\frac{\partial \psi}{\partial r_+} V_p + \frac{\partial \psi}{\partial V_+} g) + \mu_T = 0$$

- using the boundary conditions on the adjoint variables these equations become

$$H_- - \lambda^T_{r-} V_p - \lambda^T_{V-} g = H_+ - \lambda^T_{r+} V_p - \lambda^T_{V+} g$$

- using $H = \lambda^T_r V + \lambda^T_V g + TS_F$ one has

$$\lambda^T_{r-} V \infty_- + T_- S_F_- = \lambda^T_{r+} V \infty_+ + T_+ S_F_+$$
Flyby - Transversality Conditions (2)

- for a free-height flyby $\lambda_V$ is parallel to $V_\infty$ and $\lambda_{V+} = \lambda_{V-}$
- therefore, $S_{F+} = S_{F-}$ and $T_+ = T_-$ (in fact, the thrust magnitude is only determined by $S_F$)
- the transversality condition becomes $\dot{\lambda}_{V+} = \dot{\lambda}_{V-} = 0$
- on the contrary, for a minimum-height flyby $S_F$ is discontinuous and the thrust may also be discontinuous
Propulsion System Optimization (1)

• when electric propulsion is employed, the available power and the values of thrust and specific impulse may be subject to optimization

• the propulsion system mass can be considered in the performance index by posing $\varphi = m_f - m_{ps}$

• assume $m_{ps}$ proportional to the thrust power $m_{ps} = \beta Tc/2$

• additional (constant) variables $T$ and $c$ with $\dot{T} = 0$ and $\dot{c} = 0$

• introduce adjoint variables $\lambda_T$ and $\lambda_c$ and derive Euler-Lagrange equations $\dot{\lambda}_T = -\partial H/\partial T$ and $\dot{\lambda}_c = -\partial H/\partial c$
Propulsion System Optimization (2)

- derive boundary conditions for optimality at the initial point $\lambda_{T0} = 0$, $\lambda_{c0} = 0$ and final point $\lambda_{mf} = 1$, $\lambda_{Tf} = -\beta c/2$, $\lambda_{cf} = -\beta T/2$
- solution provides the optimal values for $T$ and $c$
- a constraint on the trip time is required to avoid $T \to 0$ and $t \to \infty$
Variable Specific Impulse (1)

- electric propulsion system with variable specific impulse and thrust at constant power (and efficiency) \( P = Tc/(2\eta) \) and \( T = 2\eta P/c \)
- consider \( P \) and \( c \) as control variables with \( 0 \leq P \leq P_{max} \) with unbounded specific impulse (and thrust)
- the Hamiltonian for optimal thrust direction (parallel to the primer vector) is \( H = \lambda^T_r V + \lambda^T_V g + \left( \frac{\lambda V}{m} - \frac{\lambda m}{c} \right) 2\eta \frac{P}{c} \)
- optimal exhaust velocity from \( \partial H/\partial c = 0 \), that is \( c = 2m\lambda m/\lambda V \), which gives \( T = \eta P\lambda V/(m\lambda m) \)
- the Hamiltonian becomes \( H = \lambda^T_r V + \lambda^T_V g + \eta \frac{\lambda V P}{m c} \)
- \( \partial H/\partial P > 0 \) and \( P = P_{max} \) is always required
Variable Specific Impulse (2)

- a constraint on the total trip time is required to avoid solutions with $T \to 0$ and $c \to \infty$ for $t \to \infty$
- the quantity $m^2 \lambda_m = b$ is constant, as
  \[
  \frac{d(m^2 \lambda_m)}{dt} = \lambda V T - 2m \lambda_m \frac{T}{c} = 0
  \]
  for the optimal exhaust velocity
- one has $c = 2b/(m \lambda_V)$ and the acceleration $T/m = \eta P \lambda_V / b$
  only depends on $\lambda_V$
- for a given trip time the acceleration history is fixed and does not depend on the mission parameters (power, efficiency)
- exhaust velocity history and payload, instead, depend on $P$ and $\eta$
Variable Specific Impulse (3)

- for bounded specific impulse $c_{\text{min}} \leq c \leq c_{\text{max}}$ and thrust $T_{\text{min}} = 2\eta P/c_{\text{max}} \leq T \leq T_{\text{max}} = 2\eta P/c_{\text{min}}$, define the power switching function $S_P = \frac{\lambda_v}{m} - \frac{\lambda_m}{c_{\text{max}}} \geq \frac{\lambda_v}{m} - \frac{\lambda_m}{c}$

- when $S_P < 0$ the quantity $\partial H/\partial P$ is surely negative and $P = 0$, otherwise $P = P_{\text{max}}$ must be adopted

- $c = 2m\lambda_m/\lambda_V$ when $c_{\text{min}} \leq 2m\lambda_m/\lambda_V \leq c_{\text{max}}$, otherwise the closer extreme must be adopted ($c_{\text{min}}$ when $c_{\text{opt}} \leq c_{\text{min}}$, $c_{\text{max}}$ when $c_{\text{opt}} \geq c_{\text{max}}$)

- discrete values of the specific impulse could also be considered
Variable Specific Impulse (4)

![Diagram showing variable specific impulse with exhaust velocity c]

- $P = P_a; c = c_{\text{min}}$
- $P = P_a; c = c_{\text{opt}}$
- $P = P_a; c = c_{\text{max}}$
- $P = 0$

Exhaust velocity $c$
Aerodynamic Forces

- equations of motion

\[
\frac{d\mathbf{r}}{dt} = \mathbf{V} \\
\frac{d\mathbf{V}}{dt} = g + \frac{T}{m} + \frac{D}{m} + \frac{L}{m} \\
\frac{dm}{dt} = -\frac{T}{c}
\]

- lift \( L = qSC_L L/L \), drag \( D = qSC_D V_r/V_r \), with dynamic pressure \( q = \rho V_r^2/2 \) and relative velocity \( V_r \)

- assume parabolic polar \( C_D = C_{D_0} + KC_L^2 \)
Optimal Aerodynamic Control

- Hamiltonian
  \[ H = \lambda_r^T V + \lambda_V^T g + TS_F + \frac{qS}{m} A_F \]

- aerodynamic acceleration coefficient
  \[ A_F = -C_D \lambda_V^T V_r / V_r + C_L \lambda_V^T L / L \]

- optimal controls maximize the scalar projection of the aerodynamic force on \( \lambda_V \)

- lift in the plane defined by relative velocity \( V_r \) and primer vector \( \lambda_V \)

- lift coefficient \( C_L = C_L = \tan \delta / 2K \) (or \( C_L = C_{L_{max}} \) if \( \tan \delta / 2K > C_{L_{max}} \)) where \( \delta \) is the angle between \( V_r \) and \( \lambda_V \)
Optimal Aerodynamic Force
Indirect Optimization Methods

Indirect trajectory optimization

Geopotential Perturbation

• potential given by

\[
\frac{V}{\mu/r} = - \left[ 1 + \sum_{n=2}^{N} \left( \frac{r E}{r} \right)^n \sum_{m=0}^{n} \left( C_{nm} \cos m \lambda + S_{nm} \sin m \lambda \right) P_{nm}(\sin \varphi) \right]
\]

• perturbing acceleration

\[(a_J)_u = \frac{\partial \Phi}{\partial r} \]
\[(a_J)_v = \frac{\partial \Phi}{\partial \vartheta} \left( \frac{1}{r \cos \varphi} \right) \]
\[(a_J)_w = \frac{\partial \Phi}{\partial \varphi} \left( \frac{1}{r} \right) \]

• derivatives of the adjoint variables are modified: simple calculations through a recursive scheme
Third-Body Gravitational Perturbation

• perturbing acceleration ($\mathbf{R} = \mathbf{r} - \mathbf{r}_p$ is the spacecraft position vector wrt third body)

$$a_p = -\left(\frac{\mu_p}{R^3}\right)\mathbf{R} - \left(\frac{\mu_p}{r_p^3}\right)\mathbf{r}_p$$

• projection in the topocentric frame gives components of the perturbing accelerations as functions of position vector

• derivatives of the adjoint variables to position are modified: simple calculations (even though tedious)
Ascent: Maximum Dynamic Pressure

- an arc with constant dynamic pressure $q$ at the maximum allowable value may be required
- thrust modulation to assure constant $q$
- $q$ is a function of state variables through relative velocity and density $q = q[(\rho(r), V_{rel})] = q(r, u, v, w)$
- the equation
  \[
  \frac{dq}{dt} = \frac{\partial q}{\partial \dot{x}} \dot{x}
  \]
  provides the thrust magnitude
- $q = q_{max}$ at the start of the arc and $T = T_{max}$ at the arc end are the required boundary conditions
- Euler-Lagrange equations and boundary conditions for optimality are modified
Ascent: Maximum Heat Flux

- a constraint on the heat flux $h$ may be required
- $h = h_{max}$ and possibly $\dot{h} = 0$ are the required boundary conditions
- $h$ is a function of state variables through relative velocity and density $h = h[(\rho(r), V_{rel}) = h(r, u, v, w)]$
- discontinuities in the adjoint variables arise
- good point for multiple shooting
Singular Arcs: When

- bang-bang control can be explained by the requirement of thrusting in the most favorable positions
  - larger thrust reduces velocity losses: use $T_{\text{max}}$
  - gravitational losses are smaller for larger velocity
- singular arcs arise when a reduced thrust is useful
  - aerodynamic drag: grows with velocity
  - three-body ?
- signal of requirement
  - chattering
  - irregular behavior of the switching function (it is not possible to satisfy PMP)
Singular Arcs in Atmospheric Flight

- Hamiltonian

\[ H = \lambda_r^T v + \lambda_v^T g + \frac{T}{m} S_F + \frac{q}{m} A_F \]

- thrust switching function and aerodynamic coefficient (\(\delta\): angle between velocity and primer vector)

\[ S_F = \lambda_v - \lambda_m m/c \]

\[ A_F = -C_D \lambda_v \cos \delta + C_L \lambda_v \sin \delta \]

\[ C_L = \frac{\tan \delta}{2K}, \quad C_D = C_{D0} + KC_L^2 \]
Singular Arcs in Atmosphere: Derivation

- $S_F$ and $\dot{S}_F$ vanish at arc extremities: boundary conditions to determine initial and final time, with

$$\dot{S}_F = \frac{\partial S_F}{\partial x} \dot{x} + \frac{\partial S_F}{\partial \lambda} \dot{\lambda}$$

- $\dot{S}_F$ does not contain $T$

$$\dot{S}_F = -\frac{\lambda_T^T \lambda_v}{\lambda_v} - \frac{S}{m} \left[ \frac{q A_F}{c} + (A_F Q_v^T + q A_v^T) \frac{\lambda_v}{\lambda_v} \right]$$

- coefficients

$$Q_v^T = \partial q / \partial v \quad A_v^T = \partial A_F / \partial v$$
Indirect Optimization Methods

Singular Arcs in Atmosphere: Thrust

- $\ddot{S}_F = 0$ provides thrust magnitude (first-order singular arc)
  \[
  \ddot{S}_F = \frac{\partial}{\partial x}\dot{x} + \frac{\partial}{\partial \lambda}\dot{\lambda}
  \]

- $T$ appears in derivatives of velocity components and mass
  ($a_0$: inertial + aerodynamic acceleration)
  \[
  \ddot{S}_F = \Lambda - \frac{S}{m} \left( S^T_r \dot{r} + S^T_v a_0 + S^T_\lambda \dot{\lambda}_v + S_T T \right)
  \]

- setting $\ddot{S}_F = 0$ provides
  \[
  T^* = \frac{m\Lambda/S - S^T_r v - S^T_v a_0 - S^T_\lambda \dot{\lambda}_v}{S_T}
  \]

- $S_T = 0$ vanishes in the absence of aerodynamic terms
Requirements for Convergence

- BVP solution is based on assumption of linear behavior
  - accurate evaluation of the error gradients
  - small corrections, i.e., tentative solution close to optimal solution
Error Gradient Evaluation

• accurate integration of the equations of motion
  – variable step (e.g., Adams Moulton) can handle control discontinuities
  – separation into phases can handle control discontinuities
• multiple shooting to reduce influence of single parameters
• normalization is usually mandatory (same order of magnitude for all the variables)
Tentative Solution

• must be sufficiently close to the optimal solution
  – switching structure/constraints assessment
  – avoid divergence
  – small values for the adjoint variables
• correction relaxation (i.e., use a fraction of the computed correction)
Indirect Optimization Methods

Hints and tips

Optimal Phasing

• define mission opportunities in a large launch window (in particular when gravity assist is exploited)
  – evaluate optimal phasing transfer and its periodicity
  – check for target position on possible arrival dates

• find most favorable positions for gravity assists (plane/eccentricity changes)

• favorable opportunities when actual position close to desired one
Indirect Optimization Methods

Hints and tips

Optimal Phasing - Asteroid Rendezvous

• optimal phasing transfers from earth repeat every year
• check target position on arrival day/month in available years
• one/two good opportunities every synodic cycle
Multiple Shooting

• some cases are very sensitive to initial values
  – ascent: strong nonlinearity
  – multiple flybys: sudden discontinuities
Multiple Shooting - Ascent

- uncorrect values: no integration possible
- split (e.g., at maximum heat flux)
- re-start with values that assure continuous ascent (e.g. positive radial velocity)
Multiple Shooting - Flybys

- use point immediately after flyby as split point
- accommodate for errors due to inaccurate estimations of the initial values
- position and velocity components suggested by physics of the problem
- try nonpropelled tentative solutions
Trajectory Patching

- trajectories with multiple arcs
  - $\Delta V$ gravity-assist orbits
  - multiple-flyby trajectories
- solve simple one-arc problem with proper boundary/optimality conditions
- patch simple trajectories with a multiple-shooting approach
Trajectory Patching - $\Delta V$-GA Trajectories

- maximize $V_\infty$ gain for assigned propellant consumption (or dual problem)
- use nonpropelled resonant orbit as tentative solution
**Trajectory Patching - Multiple Flybys**

- connecting planets: Hohmann transfer
- resonant orbits: assign arrival $V_\infty$ for following leg
- orientation suggested by geometry of orbits
Trajectory Patching - GTOC1

- EVVJ-GA transfer to deflect asteroid 2001 TW229
- first, each leg optimized separately
- final optimization of complete trajectory
**Trajectory Patching - GTOC4**

- flyby of 36 asteroids
- legs encompassing up to 15 flybys are optimized
- start following leg with 3 or 4-asteroid overlap
Trajectory Patching - GTOC5

- rendezvous and flyby of 17 asteroids
- first, simple R(i)-F(i)-R(i+1) legs optimized separately, mixed time-mass performance index
- optimization of up to 5 joined legs follows

![Graph showing Earth, Rendezvous, and Flyby over time](image)
Homotopy

- complex trajectories/dynamics
  - three-body problem
- solve simple constrained trajectory
- solve transition problems towards optimal solution with blended boundary conditions between constrained/optimal problem
Homotopy - Moon Return

- fix proper times and impulses for starting solution
- linear combination of the boundary conditions with multipliers $c$ and $1 - c$
- quite easy convergence
Indirect Optimization Methods

Hints and tips

Continuation

- unknown switching structure:
  - lunisolar effect on Highly Elliptic Orbit deployment
- gradual introduction of disturbing parameter
- (automatic) check evolution of switching structure
- extremely more efficient than smoothing techniques
Indirect Optimization Methods  
Hints and tips

Continuation - HEO Deployment

- multiple revolutions
- energy modulated to have favorable Moon phasing at apogee passages
- (almost) equal burns at each apogee passage without the Moon
- apogee burns may vanish as Moon’s gravity is introduced

L. Casalino 2015
Independent Variable Selection

- thrusting is usually preferred at specific positions, fixed in space
- quasi-periodicity with respect to longitude
- passage times may be unknown (depend on trajectory)
  - multiple revolution geocentric transfers
  - lunisolar effect on Highly Elliptic Orbit deployment
- use longitude as the independent variable
Indirect Optimization Methods

Independent Variable Selection - HEO

Deployment

- burns at apsides
- orbit periods dictated by better use of lunisolar perturbation
- accurate integration requires splitting the trajectory into arcs
- switching times cannot be estimated, but switching angles can

<table>
<thead>
<tr>
<th>Time, hr</th>
<th>Long., rad</th>
</tr>
</thead>
<tbody>
<tr>
<td>84.4-84.3</td>
<td>10.73-10.75</td>
</tr>
<tr>
<td>84.5-84.4</td>
<td>11.27-11.25</td>
</tr>
<tr>
<td>125.7-126.5</td>
<td>14.13-14.14</td>
</tr>
<tr>
<td>129.2-126.5</td>
<td>14.16-14.14</td>
</tr>
<tr>
<td>212.4-211.7</td>
<td>20.41-20.42</td>
</tr>
<tr>
<td>219.2-214.9</td>
<td>20.46-20.44</td>
</tr>
<tr>
<td>311.8-299.9</td>
<td>26.70-26.69</td>
</tr>
<tr>
<td>311.8-304.9</td>
<td>26.70-26.74</td>
</tr>
<tr>
<td>408.2-397.8</td>
<td>33.00-33.00</td>
</tr>
</tbody>
</table>

Comparison 1/1/15-15/1/15

L. Casalino 2015
Indirect Optimization Methods

Hints and tips

Independent Variable Selection - Multiple Revs.

• typical case: LeoGeo transfer
• continuation approach: gradually increase number of revs., final radius, inclination change
• small changes are accepted easily
• very difficult strict fulfillment of boundary conditions

<table>
<thead>
<tr>
<th>Type</th>
<th>$I_{sp}, \text{s}$</th>
<th>t, d</th>
<th>Revs.</th>
<th>mass</th>
</tr>
</thead>
<tbody>
<tr>
<td>Edelbaum</td>
<td>$\infty$</td>
<td>191</td>
<td>1048</td>
<td>1</td>
</tr>
<tr>
<td>Ion thruster</td>
<td>3000</td>
<td>174</td>
<td>989</td>
<td>0.822</td>
</tr>
<tr>
<td>Hall thruster</td>
<td>1500</td>
<td>158</td>
<td>936</td>
<td>0.675</td>
</tr>
<tr>
<td>Arcjet</td>
<td>600</td>
<td>122</td>
<td>802</td>
<td>0.374</td>
</tr>
</tbody>
</table>

L. Casalino 2015
Smart Constraint Position

• proper constraint definition may simplify derivation and solution
• use alternative (equivalent or almost equivalent) simpler form of constraint
  – separation constraint at apogee in HEO: $|r_1 - r_2| = k$
  – can be turned into time constraint $t_{a2} - t_{a1} = k/v_a$
Nested Indirect Methods

- stepped approach to solve very complex problem
- in multiple-revolution transfers solve single revolution problem first
- patch single revolution solutions together
  - LeogGeo transfer with electric propulsion
Nested Indirect Methods: Single Revolution

- one revolution: almost circular orbit, small inclination change, constant mass (Edelbaum’s approximation)
- apply OCT to define optimal controls (thrust angles, specific impulse if variable) and evaluate changes of $a$ and $m$ in one revolution
- changes for unit radius orbit with unit mass
  - constant specific impulse $\Delta r = 8\pi P \cos \beta / c_N$, $\Delta i = 8P \sin \beta / c_N$, $\Delta m = -4\pi P / c_N^2$, control is out-of-plane angle $\beta$
  - variable specific impulse $\Delta r = 16\pi PK_1$, $\Delta i = 2\pi PK_3$, $\Delta m = -2\pi P(8K_1^2 + K_3^2)$, controls are $K_1$ and $K_3$
Nested Indirect Methods: Multiple Revolutions

- multiple revolutions: radius and mass changes are considered and elements changes are modified accordingly
- differential equations with $r$ as independent variable
  - constant specific impulse
    \[
    \frac{di}{dr} = \frac{\tan \beta}{(\pi r)} \quad \frac{dt}{dr} = \frac{m}{(2\sqrt{r^3T \cos \beta})}
    \]
    \[
    \frac{dm}{dr} = -\frac{m}{(2c\sqrt{r^3 \cos \beta})}
    \]
  - variable specific impulse
    \[
    \frac{di}{dr} = \frac{K_3}{(8rK_1)} \quad \frac{dt}{dr} = \frac{m}{(8P\sqrt{r^3 K_1})}
    \]
    \[
    \frac{dm}{dr} = -\frac{m(8K_1^2 + K_2^2)}{(8\sqrt{r^3K_1})\sqrt{r^3T \cos \beta})}
    \]
Nested Indirect Methods: Optimal Problem

• apply OCT to determine optimal values of the controls (either $\beta$ or $K_1$ and $K_2$)
  
  - constant specific impulse

  \[
  H = \lambda_i \tan \beta/(\pi r) + \lambda t m/(2\sqrt{r^3} T \cos \beta) - \lambda m m/(2c\sqrt{r^3} \cos \beta)
  \]

  \[
  \sin \beta = \frac{2\lambda_i}{\pi m(\lambda_m/c - \lambda_t/T) \sqrt{r}}
  \]

  - variable specific impulse

  \[
  H = \lambda_i K_3/(8r K_1) + \lambda t m/(8P\sqrt{r^3} K_1) - \lambda m m(8K_1^2 + K_3^2)/(8\sqrt{r^3} K_1)
  \]

  \[
  K_3 = \frac{\lambda_i \sqrt{r}}{2m \lambda_m} \quad K_1 = \sqrt{\frac{1}{8} \left( \frac{1}{\lambda_m P} - K_3^2 \right)}
  \]
Nested Indirect Methods: LeoGeo Results

- Inclination change $\Delta i/T_n$
- Radius change $\Delta r/(\pi T_n)$
- Thrust angle $\beta$ (deg)
- Effective exhaust velocity $c$

L. Casalino 2015
Indirect Methods and Evolutionary Algorithms

- use an EA to determine the unknowns of the MPBVP from the application of OCT
  - solution may be inaccurate
  - usually not effective in complex problems
- use an EA to estimate the unknowns of the MPBVP (that is, to get a tentative solution)
  - requires solution of MPBVP with Newton’s method
Future Developments

• applications suited for indirect methods are almost infinite

• possible interesting developments
  – nested optimization of geocentric transfers with eclipses and perturbations
  – developments of general (automatic) techniques to define and correct switching and constraint structure
Conclusions

• savvy application of indirect methods may allow for solution of very complex problems
• several techniques to permit convergence may be devised depending on the peculiarities of the problem
• fundamental relevance of problem formulation
• user’s intuition necessary to guide the solution process