



UNIVERSIDAD AUTÓNOMA DE
CHIAPAS



FACULTAD DE CIENCIAS EN FÍSICA Y
MATEMÁTICAS

Del cálculo variacional al control óptimo:
análisis estructural y condiciones de
optimalidad en sistemas multivariantes
deterministas

T E S I S

QUE PARA OBTENER EL TÍTULO DE:
LICENCIADO EN MATEMÁTICAS

PRESENTA:
MARCO ANTONIO MORENO MONTOYA

DIRECTOR:
Dr. Yofre Hernán García Gómez

Tuxtla Gutiérrez, Chiapas a 27 de octubre del 2025.

Tabla de contenidos

Resumen	3
Introducción	4
Objetivos	6
Objetivo General	6
Objetivos Especificos	6
1. Preliminares	7
1.1. Notación	7
1.2. Conjuntos y Espacios	8
1.3. Elementos de análisis funcional	10
1.4. Resultados importantes	12
I. Cálculo Variacional	15
2. Cálculo Variacional	18
2.1. ¿Qué es un Funcional?	18
2.2. Funcionales en Varias Variables	18
2.3. Espacios Funcionales	19
2.4. Continuidad de un Funcional	19
2.5. Linealidad de un Funcional	22
2.6. El Incremento de un Funcional	23
2.7. La Primera Variación	24
2.8. Máximo y Mínimo de un Funcional	27
2.9. Lema y Teorema Fundamental del Cálculo Variacional	27
2.10. Ecuaciones de Euler	29
2.11. Teorema de Noether	31
2.12. Transformación de Legendre y Hamiltoniano	38
2.13. Problema Primal y Dual en Cálculo Variacional	41
II. Control Óptimo	43
3. Del Cálculo Variacional a la Teoría del Control Óptimo	45
3.1. Trayectorias y Controles	45

3.2. Formulación del Problema de control óptimo	47
3.3. Incremento del Funcional y Primera Variación en Control Óptimo	48
3.4. Ecuaciones de Euler-Lagrange Generalizadas	50
3.5. Hamiltoniano en Control Óptimo	55
3.6. Principio del Mínimo de Pontryagin (PMP)	55
4. Existencia de Soluciones en Problemas de Control Óptimo	59
4.1. Teorema de Existencia de Filippov	59
III. Caso de Estudio	64
Problema de Goddard	65
5. Problema Original	66
5.1. Formulación	66
5.2. Resultados Numéricos	67
5.2.1. Método Directo	67
5.2.2. Método Indirecto	70
5.2.3. Análisis de Sensibilidad	73
5.2.4. Interpretación de los resultados	77
6. Problema Alternativo	79
6.1. Método Directo	79
6.2. Análisis de Sensibilidad	83
7. Conclusiones	86
Referencias	88

Resumen

Este trabajo establece la conexión entre el cálculo variacional y la teoría del control óptimo. A partir de la formulación clásica de los funcionales y de las ecuaciones de Euler–Lagrange, se desarrolla un enfoque que permite generalizar dichas condiciones a sistemas con restricciones dinámicas. El estudio se centra en el problema de Bolza, por su carácter general y su capacidad para englobar las formulaciones de Lagrange y Mayer.

Mediante la introducción del Hamiltoniano y del funcional aumentado, se establece el vínculo entre el problema primal y su representación dual, lo que conduce de manera natural a las condiciones de optimalidad de Pontryagin. El trabajo adopta una perspectiva multivariable y determinista, que permite analizar las ecuaciones de Euler–Lagrange y de Hamilton, junto con sus implicaciones en la estructura de los sistemas controlados.

De este modo, el trabajo ofrece una exposición rigurosa que integra los fundamentos teóricos del cálculo variacional con los principios del control óptimo, destacando la perspectiva multivariable y su relevancia dentro del análisis moderno de sistemas dinámicos.

Introducción

El *Cálculo Variacional* ha desempeñado un papel fundamental en las formulaciones matemáticas de problemas de optimización desde el siglo XVIII. Su desarrollo, principalmente impulsado por la física y la mecánica, permitió establecer un marco teórico para describir leyes de la naturaleza mediante principios de mínima acción (Ferguson (2004)). En este sentido, los problemas de determinar funciones que optimizan funcionales dio origen a una estructura matemática en la que las ecuaciones de **Euler-Lagrange** emergen como condiciones necesarias de optimalidad.

Sin embargo, conforme los sistemas físicos, de ingeniería, de economía, etc. se tornaron más complejos, el enfoque clásico comenzó a dejar de ser suficiente. Muchos fenómenos reales requieren de la incorporación de *variables de control* que actúan sobre la dinámica de un sistema, y sobre los cuales la trayectoria del estado ya no se encuentra libre, sino que debe satisfacer ciertas restricciones que se imponen en el sistema mediante una ecuación diferencial. Este cambio de paradigma condujo a lo que hoy se conoce como **Teoría de Control Óptimo**, donde los métodos de cálculo variacional se amplían y reinterpretan para incluir restricciones dinámicas, dominios de control y condiciones mucho más generales.

Dentro de este marco, los problemas de control óptimo pueden formularse en diversas formas equivalentes. Las más comunes son las formulaciones de **Lagrange, Mayer y Bolza**, cuya diferencia radica en la manera en que expresan la función objetivo. El problema de Lagrange considera la optimización de una integral dependiente de la trayectoria y el control; el de Mayer, en cambio, se centra en minimizar una función terminal que depende del estado final; mientras que el problema de Bolza combina ambos términos en una formulación unificada. Dado que el problema de Bolza contiene como casos particulares a los otros dos, este trabajo adopta dicha forma como modelo general, sobre el cual se establecen las condiciones de optimalidad y las relaciones estructurales que se desarrollan a lo largo del texto.

El presente trabajo aborda el estudio de esta transición conceptual y formal **del cálculo variacional clásico al control óptimo**, haciendo énfasis en las estructuras matemáticas que da vida a ambos enfoques. En particular, se analiza cómo las condiciones de Euler-Lagrange pueden reinterpretarse dentro de un marco más general al introducir el **Hamiltoniano**, el cual actúa como vínculo entre la formulación primal de un problema y su formulación dual. Este enfoque permite establecer una conexión natural con el **Principio del Mínimo de Pontryagin**. Además, se incorpora una perspectiva

multivariable que permite tratar sistemas más complejos con varias funciones de estado y control, destacando la importancia de los estados funcionales adecuados y de la formulación vectorial de las ecuaciones de Euler-Lagrange y del Hamiltoniano.

La estructura de la tesis se organiza de la siguiente manera: en el Capítulo 1 (Capítulo 1) se introducen los preliminares necesarios, incluyendo nociones básicas de análisis, Conjuntos y la notación que se utiliza a lo largo del trabajo. En el Capítulo 2 (Capítulo 2) se desarrolla la teoría del cálculo variacional en varias variables, presentando el concepto de incremento, la primera variación y la derivación formal de las ecuaciones de Euler-Lagrange, así como la transformación de Legendre y el surgimiento del Hamiltoniano como objeto fundamental. Posteriormente, los Capítulos 3 y 4 (Capítulo 3) extiende estas ideas a problemas de control óptimo, incorporando restricciones dinámicas, explorando la formulación aumentada del funcional y el principio de Pontryagin. Finalmente en los Capítulos 5 y 6 (Capítulo 5 y Capítulo 6) abordamos el caso de estudio, utilizando el problema clásico de Goddard (1919).

Objetivos

Objetivo General

Analizar los fundamentos del cálculo variacional y su extensión al control óptimo en varias variables, aplicando el Principio del Mínimo de Pontryagin a un problema clásico (Goddard) y evaluando soluciones numéricas mediante métodos directos e indirectos.

Objetivos Especificos

- Revisar los conceptos básicos necesarios para definir problemas de calculo variacional y control óptimo en varias variables con restricciones.
- Exponer los fundamentos del cálculo variacional, deducir las ecuaciones de Euler–Lagrange y el teorema de Noether.
- Formular el problema de control óptimo y demostrar condiciones de existencia mediante el teorema de Filippov.
- Implementar aproximaciones numéricas de la solución en un problema aplicado (Goddard) utilizando métodos directos e indirectos y ver la sensibilidad de dichas aproximaciones ante variaciones en parámetros o restricciones.

1. Preliminares

El propósito de este capítulo es reunir conceptos, resultados y convenciones que serán utilizados a lo largo del trabajo. Se busca establecer las bases necesarias para el desarrollo de los temas posteriores, facilitando la comprensión y evitando repeticiones innecesarias.

1.1. Notación

En esta sección fijamos la notación que utilizaremos durante todo el trabajo. Adoptamos las convenciones usuales en análisis real, cálculo variacional y teoría de control óptimo, tomando como referencia principalmente a Rudin (1976), Cesari (1983), Rockafellar (197d. C.).

Denotamos por $\mathbb{R}^n$ el **espacio euclidiano n -dimensional**, provisto del producto interno usual

$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i,$$

y de la norma inducida $\|x\| = \sqrt{\langle x, x \rangle}$. Cuando el contexto lo permita, escribiremos simplemente $\langle \cdot, \cdot \rangle$ y $\|\cdot\|$ para denotar este producto interno y la norma asociada. Además, utilizaremos $\mathbb{N}$ para el conjunto de los números naturales, $\mathbb{Z}$ para los enteros, y $\mathbb{Q}$ para los racionales. El conjunto de los números reales se denota por $\mathbb{R}$, y el de los complejos por $\mathbb{C}$.

La **transpuesta** de un vector o matriz A se denotará por $A^\top$. El producto matricial de dos matrices A y B se denota por AB , y se siguen las reglas usuales. Para el producto escalar entre dos vectores $x, y \in \mathbb{R}^n$ escribiremos $x^\top y = \langle x, y \rangle$.

Si $x = (x_1, \dots, x_n)$, el símbolo x_i designa su i -ésima componente.

En el caso de funciones vectoriales $y : \mathbb{R}^n \rightarrow \mathbb{R}^m$, denotamos por $Dy(x)$ a su **matriz jacobiana** sobre x , y por $D_i y$ la derivada parcial respecto a la i -ésima variable.

Para un campo $F = (F_1, \dots, F_n) : \mathbb{R}^n \rightarrow \mathbb{R}^n$ y $x = (x_1, \dots, x_n)$, definimos su **divergencia** como

$$\operatorname{div} F = \sum_{i=1}^n \frac{\partial F_i}{\partial x_i}.$$

Este operador mide la tasa neta o el flujo del campo en un punto dado.

Cuando una propiedad se cumpla ***casi en todas partes*** sobre un conjunto medible, lo indicaremos mediante la abreviatura **c.t.p.**

Estas convenciones se mantendrán uniformes en todo el trabajo, y cualquier notación adicional se introducirá en las secciones donde sea necesaria.

1.2. Conjuntos y Espacios

En esta sección introducimos las definiciones y propiedades básicas de conjuntos y espacios, que serán esenciales para el desarrollo posterior. Nos enfocaremos en las operaciones fundamentales entre conjuntos, así como en la estructura de los espacios donde se desarrollan los problemas de interés. Las definiciones y propiedades presentadas siguen la exposición clásica de Munkres (2000) Rudin (1976), Rockafellar (197d. C.), proporcionando el marco necesario para el análisis matemático riguroso.

Definición 1.1 (Conjunto). Un **conjunto** se considera una colección de objetos en la que cada uno se denomina como punto (o elemento).

Si x es un *elemento* que pertenece a un conjunto A , lo expresamos con la notación $x \in A$, análogamente, si x no es un elemento que pertenece al conjunto A , escribimos $x \notin A$. El **conjunto vacío** $\emptyset$ es el conjunto que no tiene elementos.

Definición 1.2 (Subconjunto). Decimos que A es un **subconjunto** de B , si cada elemento de A también es un elemento de B , y expresamos esto escribiendo $A \subset B$.

En general, todos los conjuntos se consideran subconjuntos de un **conjunto universal** (o **espacio** o **universo**), que contiene a todos los elementos bajo consideración en un contexto dado. Este conjunto universal se denota comúnmente por U . En muchos casos, se toma $U = \mathbb{R}^n$ para algún $n \in \mathbb{N}$, dependiendo del problema o área de estudio. Esta noción permite definir las operaciones entre conjuntos.

Definición 1.3 (Conjunto Abierto). Un conjunto A se dice **abierto** si para todo $x \in A$ existe $\varepsilon > 0$ tal que la **bola abierta** centrada en x y de radio ε , definida por

$$B_\varepsilon(x) = \{y \in \mathbb{R}^n : \|y - x\| < \varepsilon\},$$

está contenida en A , es decir, $B_\varepsilon(x) \subset A$.

Definición 1.4 (Conjunto Cerrado). Un conjunto C se dice **cerrado** si su complemento $U \setminus C$ es un conjunto abierto, es decir, si todo punto que no pertenece a C admite una vecindad completamente contenida fuera de C . Equivalentemente, C es cerrado si contiene todos sus puntos de acumulación.

Definición 1.5 (Clausura de un Conjunto). La **clausura** (o **cerradura**) de un conjunto A , denotada por $\overline{A}$, es el conjunto formado por todos los puntos de A junto con todos sus puntos de acumulación. Es el conjunto cerrado más pequeño que contiene a A , es decir,

$$\overline{A} = \bigcap \{C \subset U : C \text{ es cerrado y } A \subset C\}.$$

Definición 1.6 (Compacidad). Un conjunto K es **compacto** si para toda familia de conjuntos abiertos $\{U_\alpha\}_{\alpha \in A}$ tal que $K \subset \bigcup_{\alpha \in A} U_\alpha$, existe una subfamilia finita $\{U_{\alpha_1}, \dots, U_{\alpha_m}\}$ con $K \subset \bigcup_{j=1}^m U_{\alpha_j}$. Es decir, toda cubierta abierta de K admite una subcubierta finita.

Por el **teorema de Heine-Borel**, en $\mathbb{R}^n$, un conjunto es compacto si y solo si es cerrado y acotado (Rudin (1976)).

Definición 1.7 (Conjunto Convexo). Un conjunto C se dice **convexo** si para cualesquiera dos puntos $x, y \in C$ y para todo $\lambda \in [0, 1]$, el punto

$$\lambda x + (1 - \lambda)y \in C.$$

Definición 1.8 (Proyección). Sea $A \subset X \times Y$, donde X y Y son espacios euclidianos. La proyección de A sobre X se define como

$$\pi_X(A) = \{x \in X : \exists y \in Y \text{ tal que } (x, y) \in A\}.$$

De manera análoga, la proyección de A sobre Y es

$$\pi_Y(A) = \{y \in Y : \exists x \in X \text{ tal que } (x, y) \in A\}.$$

Proposición 1.1 (Proyección de un conjunto compacto). *Si $K \subset X \times Y$ es un conjunto compacto, entonces su proyección sobre X (respectivamente sobre Y), $\pi_X(K)$, también es un conjunto compacto.*

Prueba. Sea $\{U_\alpha\}_{\alpha \in A}$ una familia de abiertos en X tal que $\pi_X(K) \subset \bigcup_{\alpha \in A} U_\alpha$. Entonces, $\{U_\alpha \times Y\}_{\alpha \in A}$ es una familia de abiertos en $X \times Y$ que cubre K , ya que si $(x, y) \in K$, entonces $x \in \pi_X(K)$ y existe α tal que $x \in U_\alpha$, por lo que $(x, y) \in U_\alpha \times Y$. Como K es compacto, existe una subfamilia finita $\{U_{\alpha_1} \times Y, \dots, U_{\alpha_n} \times Y\}$ que cubre K . Por lo tanto, $\{U_{\alpha_1}, \dots, U_{\alpha_n}\}$ cubre $\pi_X(K)$, mostrando que $\pi_X(K)$ es compacto. $\square$

Para elementos $x(t), y(t)$, $t_1 \leq t \leq t_2$, definidos sobre el mismo intervalo, la función de distancia usual

$$d(x, y) = \sup_{t_1 \leq t \leq t_2} \|x(t) - y(t)\|,$$

es una manera natural de medir su separación. (Cesari (1983))

Definición 1.9 (Métrica rho (ρ)). Si $x(t)$, $a \leq t \leq b$, y $y(t)$, $c \leq t \leq d$, están definidos sobre intervalos que pueden ser distintos, entonces definimos

$$\rho(x, y) = |a - c| + |b - d| + \sup_{t \in [a, b] \cap [c, d]} |x(t) - y(t)|$$

1.3. Elementos de análisis funcional

En esta sección introducimos las nociones fundamentales de funciones reales y vectoriales que serán empleadas en el desarrollo posterior. Presentamos definiciones, propiedades y ejemplos básicos que servirán como herramientas esenciales para el análisis y la formulación de resultados en los capítulos siguientes. El enfoque estará en los aspectos más relevantes para el cálculo variacional y la teoría de control óptimo, siguiendo la notación y convenciones establecidas previamente, tomando como referencia a Rudin (1976) y Spivak (2008).

Definición 1.10 (Función continua). Una función f es **continua en** a si

$$\lim_{x \rightarrow a} f(x) = f(a).$$

Diremos que una función es **continua en un conjunto** X si es continua en todo punto de este.

Definición 1.11 (Función acotada). Una función f es **acotada** si existe una constante $M > 0$ tal que $\|f(x)\| \leq M$ para todo x en el dominio de f .

Definición 1.12 (Función absolutamente continua). Una función $f : [a, b] \rightarrow \mathbb{R}^n$ es **absolutamente continua** si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para cualquier colección finita de intervalos disjuntos $[(x_k, y_k)]$ en $[a, b]$ con $\sum_k |y_k - x_k| < \delta$, se cumple

$$\sum_k \|f(y_k) - f(x_k)\| < \varepsilon.$$

Toda función absolutamente continua es continua y, además, derivable casi en todas partes, con

$$f(x) = f(a) + \int_a^x f'(t) dt$$

para todo $x \in [a, b]$.

Definición 1.13 (Función L-integrable). Una función f se dice ser **L^p -integrable** en un conjunto medible $X \subset \mathbb{R}^n$ si

$$\int_X \|f(x)\|^p dx < \infty$$

para algún $1 \leq p < \infty$. El conjunto de todas las funciones L^p -integrables en X se denota por $L^p(X)$.

Definición 1.14 (Funciones Equi-Lipschitzianas). Decimos que una función f es **Lipschitziana** en un conjunto E si existe una constante $L > 0$ tal que

$$\|f(x) - f(y)\| \leq L\|x - y\|, \quad \forall x, y \in E$$

Por otra parte, una familia de funciones $\mathcal{F}$ definidas en E es **equi-Lipschitziana** si existe una constante $L > 0$ tal que toda $f \in \mathcal{F}$ es Lipschitziana.

Definición 1.15 (Funciones equicontinuas). Una familia de funciones $\mathcal{F}$ definidas en un conjunto E es **equicontinua** si para todo $\varepsilon > 0$ existe $\delta > 0$ tal que para todo $x, y \in E$ con $\|x - y\| < \delta$ y para toda $f \in \mathcal{F}$ se cumple

$$\|f(x) - f(y)\| < \varepsilon.$$

Definición 1.16 (Convergencia puntual). Una sucesión de funciones $\{f_n\}$ **converge puntualmente** a una función f si

$$\lim_{n \rightarrow \infty} f_n(x) = f(x),$$

para todo x en el dominio

Definición 1.17 (Convergencia uniforme). Una sucesión de funciones $\{f_n\}$ **converge uniformemente** a una función f si

$$\forall \varepsilon > 0, \exists N \text{ tal que } |f_k(x) - f(x)| < \varepsilon,$$

para todo x en el dominio y $k \geq N$.

Definición 1.18 (Mínimo y máximo de una función). Sea f una función definida sobre un conjunto A . Decimos que f alcanza su **mínimo** en $x_0 \in A$ si

$$f(x_0) \leq f(x), \quad \forall x \in A.$$

El valor $f(x_0)$ se llama el **mínimo absoluto** de f en A .

Análogamente, f alcanza su **máximo** en $x_1 \in A$ si

$$f(x_1) \geq f(x), \quad \forall x \in A,$$

y el valor $f(x_1)$ se llama el **máximo absoluto** de f en A .

Definición 1.19 (Arg min y Arg max). Sea f una función definida sobre un conjunto A . El **conjunto de puntos donde f alcanza su mínimo** se denota por

$$\operatorname{argmin}_{x \in A} f(x) = \{x^* \in A : f(x^*) \leq f(x), \forall x \in A\}.$$

Análogamente, el **conjunto de puntos donde f alcanza su máximo** se denota por

$$\operatorname{argmax}_{x \in A} f(x) = \{x^* \in A : f(x^*) \geq f(x), \forall x \in A\}.$$

1.4. Resultados importantes

Definición 1.20 (Propiedad (K) y Propiedad (Q)). Sea $x \mapsto Q(x)$, con $x \in X$, $Q(x) \subset Y$, una aplicación multivaluada de un espacio métrico X hacia un espacio topológico Y .

Decimos que se tiene la **propiedad (K)** en $x_0 \in X$ si

$$Q(x_0) = \bigcap_{\delta > 0} \overline{\bigcup_{x \in N_\delta(x_0)} Q(x)}. \quad (1.1)$$

Por otro lado, se tiene la **propiedad (Q)** en x_0 si

$$Q(x_0) = \bigcap_{\delta > 0} \overline{\operatorname{co} \bigcup_{x \in N_\delta(x_0)} Q(x)}.$$

Teorema 1.1 (Arzelá-Ascoli). Sean K un espacio métrico compacto. Si una sucesión $\{f_n\} \subseteq C(K)$ es puntualmente acotada y equicontinua, entonces tiene una subsucesión que converge uniformemente.

Prueba. Vease (67.2. p.336 de McShane (2015)) □

Teorema 1.2 (Función implícita). Sea A un subconjunto cerrado de $\mathbb{R}^v + n$. Para cada $(t, x) \in A$ sea $U(t, x)$ un subconjunto de $\mathbb{R}^m$. Supongamos que el conjunto

$$M = \{(t, x, u) \mid (t, x) \in A, u \in U(t, x)\} \subset \mathbb{R}^{v+n+m},$$

es cerrado. Sean $f = (f_1, \dots, f_r)$ y $f_0(t, x, u)$ funciones continuas en M . Y para cada $(t, x) \in A$ sea

$$\tilde{Q}(t, x) = \{(z_0, z) \mid z_0 \geq f_0(t, x, u), z = f(t, x, u), u \in U(t, x)\} \subset \mathbb{R}^{r+1}.$$

Sea G un subconjunto medible de $\mathbb{R}^v$, y sean $\eta(t), \xi(t) = (\xi_1, \dots, \xi_r), x(t) = (x_1, \dots, x_n), t \in G$, funciones medibles tales que $(t, x(t)) \in A, (\eta(t), \xi(t)) \in \tilde{Q}(t, x(t)), t \in G$ (c.t.p). Entonces existe una función medible $u(t) = (u_1, \dots, u_m), t \in G$, tal que $u(t) \in U(t, x(t)), \eta(t) \geq f_0(t, x(t), u(t)), \xi(t) = f(t, x(t), u(t)), t \in G$ (c.t.p).

Prueba. Vease (8.2.iii. p.278 de Cesari (1983)) □

Proposición 1.2. Sean $A \subset \mathbb{R}^v$ y, para cada $x \in A$, sea $U(x) \subset \mathbb{R}^m$. Definamos

$$M = \{(x, u) \mid x \in A, u \in U(x)\} \subset \mathbb{R}^{v+m}.$$

Sea $f = (f_1, \dots, f_n) : M \rightarrow \mathbb{R}^n$ y sea $f_0 : M \rightarrow \mathbb{R}$ funciones continuas en M . Para cada $x \in A$ sean

$$Q(x) = f(x, U(x)) = \{z \mid z = f(x, u), u \in U(x)\} \subset \mathbb{R}^n,$$

y

$$\tilde{Q}(x) = \{(z_0, z) \mid z_0 \geq f_0(x, u), z = f(x, u), u \in U(x)\}.$$

También, sean

$$M_0 = \{(x, z) \mid x \in A, z \in Q(x)\} \subset \mathbb{R}^{v+n},$$

y

$$\tilde{M}_0 = \{(x, z_0, z) \mid x \in A, (z_0, z) \in \tilde{Q}(x)\} \subset \mathbb{R}^{v+1+n}.$$

Si M es compacto y f_0, f son continuas en M , entonces todos los conjuntos $\tilde{Q}(x)$ son cerrados y cuya proyección $Q(x)$ sobre $\mathbb{R}^n$ es compacta. Además son semicontinuos superiormente respecto al conjunto inclusión, por lo que tienen la propiedad (K) en A , si además son convexos, entonces también la propiedad (Q) en A . Además se tiene que $\tilde{M}_0$ es cerrado y su proyección M_0 sobre $\mathbb{R}^{v+n}$ es compacta.

Prueba. La compacidad de M implica la compacidad de A y de cada conjunto $U(x)$ (Proposición 1.1). Sea $x \in A$ arbitrario, como $U(x)$ es compacto y f es continua, entonces su imagen $f(x, U(x)) = Q(x)$ es compacta.

Veamos que $\tilde{Q}(x)$ es cerrado. Sea $((z_0^k, z^k))_{k \in \mathbb{N}} \subset \tilde{Q}(x)$ una sucesión convergente a $(z_0, z) \in \mathbb{R}^{1+n}$. Por definición, para cada k existe $u_k \in U(x)$ tal que

$$z^k = f(x, u_k), \quad z_0^k \geq f_0(x, u_k).$$

Como $U(x)$ es compacto, la sucesión (u_k) tiene una subsucesión convergente $u_{k_j} \rightarrow u^* \in U(x)$. Dado que f y f_0 son continuas en M , se cumple

$$f(x, u_{k_j}) \rightarrow f(x, u^*) \quad \text{y} \quad f_0(x, u_{k_j}) \rightarrow f_0(x, u^*).$$

Ya que $z^{k_j} \rightarrow z$, se sigue que $z = f(x, u^*)$, y tomando límites en $z_0^{k_j} \geq f_0(x, u_{k_j})$ se obtiene $z_0 \geq f_0(x, u^*)$. Por lo tanto, $(z_0, z) \in \tilde{Q}(x)$, y así, $\tilde{Q}(x)$ es cerrado.

Probemos ahora que los conjuntos $\tilde{Q}(x)$ son semicontinuos superiormente respecto al conjunto inclusión.

Sea $x_0 \in A$ y consideremos una sucesión $(x_k)_{k \in \mathbb{N}} \subset A$ con $x_k \rightarrow x_0$. Para cada k tomamos un punto $(z_0^k, z^k) \in \tilde{Q}(x_k)$, por definición existen $u_k \in U(x_k)$ tales que

$$z^k = f(x_k, u_k), \quad z_0^k \geq f_0(x_k, u_k).$$

La sucesión (x_k, u_k) se encuentra en M , que es compacto; por lo tanto existe una subsucesión (x_{k_j}, u_{k_j}) y un punto $(x^*, u^*) \in M$ tales que $(x_{k_j}, u_{k_j}) \rightarrow (x^*, u^*)$. Como la proyección de M sobre la coordenada x es A y $x_{k_j} \rightarrow x_0$, necesariamente $x^* = x_0$ y, por tanto, $u^* \in U(x_0)$.

Si la sucesión $((z_0^k, z^k))$ no tiene ninguna subsucesión convergente, la condición de semicontinuidad superior se satisface por vacuidad (no hay límite que comprobar). En caso contrario, consideramos una subsucesión convergente, de modo que

$$(z_0^{k_j}, z^{k_j}) \rightarrow (z_0, z).$$

Como f y f_0 son continuas, obtenemos

$$f(x_{k_j}, u_{k_j}) \rightarrow f(x_0, u^*), \quad f_0(x_{k_j}, u_{k_j}) \rightarrow f_0(x_0, u^*).$$

Dado que $z^{k_j} = f(x_{k_j}, u_{k_j})$ y $z^{k_j} \rightarrow z$, entonces $z = f(x_0, u^*)$. Además, de $z_0^{k_j} \geq f_0(x_{k_j}, u_{k_j})$, se sigue que $z_0 \geq f_0(x_0, u^*)$. Por tanto $(z_0, z) \in \tilde{Q}(x_0)$.

Esto muestra la caracterización por sucesiones de la semicontinuidad superior: si $x_k \rightarrow x_0$ y $(z_0^k, z^k) \in \tilde{Q}(x_k)$ converge a (z_0, z) , entonces $(z_0, z) \in \tilde{Q}(x_0)$. Luego $\tilde{Q}$ es semicontinuo superiormente respecto a la inclusión.

En consecuencia, la correspondencia $x \mapsto \tilde{Q}(x)$ es semicontinua superiormente y toma valores cerrados y acotados en $\mathbb{R}^{1+n}$; por lo que tiene valores compactos. De acuerdo con la definición (Definición 1.20), la semicontinuidad superior de $\tilde{Q}$ junto con la compacidad de sus valores implica la propiedad (K) en A ; y si además los conjuntos $\tilde{Q}$ son convexos, también la propiedad (Q). Además notemos que $\tilde{M}_0$ es un grafo de $\tilde{Q}$ y tomando en cuenta la definición equivalente de la semicontinuidad que dice que *una aplicación multivaluada Q con valores cerrados es semicontinua superiormente si y solo si su grafo es cerrado*, concluimos que $\tilde{M}$ es cerrado. Finalmente, notemos que M_0 es la imagen de M bajo la función continua $(x, u) \mapsto (x, f(x, u))$, y ya que M es compacto concluimos que M_0 es compacto. $\square$

Parte I.

Cálculo Variacional

El cálculo variacional es una área de las matemáticas que estudia la optimización de funcionales, es decir, funciones cuyo dominio es un conjunto de funciones. Aunque los primeros problemas en este campo se remontan a la antigua Grecia, su desarrollo significativo ocurrió entre los siglos XVIII y XIX, impulsado por la búsqueda de soluciones óptimas en trayectorias, superficies mínimas y principios físicos (Ferguson 2004) y dando lugar a problemas clásicos como la braquistócrona de Johann Bernoulli, el principio de mínima acción de Maupertuis, etc.

Con el desarrollo de la teoría, se ha dado paso a problemas más complejos y aplicaciones en diversas áreas, incluyendo la economía, biología, control óptimo, ingeniería aeroespacial, entre otras. Presentaremos algunos fundamentos y abordaremos la existencia de soluciones de problemas de cálculo variacional. Posteriormente, introduciremos el Principio de Pontryagin, el cual nos permite extender los resultados del cálculo variacional hacia el ámbito del control óptimo.

Los problemas en el cálculo variacional consisten en encontrar una función $y(x)$ tal que minimice o maximice un funcional de la forma

$$J[y] = \int_a^b F(x, y, y') dx$$

donde F es una función dada que depende de la variable independiente x , la función $y(x)$ y su derivada $y'(x)$. Este problema puede plantearse agregándole diversas condiciones de frontera, restricciones, etc.

El primer problema real fue estudiado por Isaac Newton (1643-1727) examinando el movimiento de cuerpos en un medio resistente.

En junio de 1696, el matemático Johann Bernoulli (1667-1748) presentó en *Acta Eruditorum* (Bernoulli 1996), el primer periódico científico alemán, el problema de *La Braquistócrona* (**The Brachistochrone**) dirigido a “los matemáticos más inteligentes del mundo”, este problema se convirtió en uno de los más famosos del cálculo variacional. Su formulación, originalmente escrita en Latín, puede traducirse de la siguiente manera:

Si en un plano vertical dos puntos A y B son dados, encuentre la trayectoria AMB del punto movable M , a lo largo de la cual, partiendo de A , y bajo la influencia de su propio peso, llega a B en el menor tiempo posible” (H. J. Sussmann y Willems 1997).

Este problema atrajo la atención de numerosos matemáticos destacados, entre ellos Newton y Leonhard Euler. En 1744, Euler publicó *Methodus inveniendi lineas curvas maximi minimive proprietate gaudentes, sive solutio problematis isoperimetrici latissimo sensu*, una obra que muchos consideran el punto de partida formal del cálculo variacional. Posteriormente Joseph-Louis Lagrange (1736-1813) le propuso a Euler un nuevo método

analítico que simplificaba el enfoque geométrico previamente utilizado. La idea de Lagrange de comparar funciones condujo directamente a las ecuaciones Euler-Lagrange. Después de considerar el método de Lagrange, Euler abandonó su enfoque geométrico, y nombró el campo con el nombre que usamos hoy en día, *Cálculo Variacional* o *Cálculo de Variaciones*, en reconocimiento al método variacional de Lagrange (Ferguson 2004).

2. Cálculo Variacional

2.1. ¿Qué es un Funcional?

Un **funcional** es una regla de correspondencia que asigna un único número real a cada función de cierto conjunto determinado. Formalmente si $\mathcal{A}$ es un conjunto de funciones definidas sobre un conjunto Ω , un funcional es una aplicación

$$J : \mathcal{A} \rightarrow \mathbb{R}, \quad y \mapsto J[y]$$

Ejemplo 2.1.

Ejemplo de funcional

Supongamos que $x(\cdot)$ es una función continua de t definida en el intervalo $[t_0, t_f]$ y

$$J(x) = \int_{t_0}^{t_f} x(t) dt,$$

un funcional que arroja el área bajo la curva de $x(t)$.

2.2. Funcionales en Varias Variables

Generalizando a funciones vectoriales y dominios multidimensionales, definimos el funcional

$$J[y] := \int_{\Omega} L(x, y(x), Dy(x)) dx, \quad (2.1)$$

donde

- $\Omega \subset \mathbb{R}^n$, $n \geq 1$, es un conjunto abierto y acotado cuyos puntos denotamos por $x = (x_1, x_2, \dots, x_n)$;
- $y : \Omega \rightarrow \mathbb{R}^N$ es una función expresada como $y = (y_1, \dots, y_N)$, perteneciente a un **espacio de funciones admisibles** $\mathcal{A}$ donde las condiciones de frontera se imponen como restricciones sobre $\mathcal{A}$.

- De modo que

$$Dy = \left(\frac{\partial y_j}{\partial x_i} \right)_{\substack{1 \leq j \leq N \\ 1 \leq i \leq n}} \in \mathbb{R}^{N \times n}$$

- $L : \Omega \times \mathbb{R}^N \times \mathbb{R}^{N \times n} \rightarrow \mathbb{R}$ es una función dada, denominada **Lagrangiano**, que satisface ciertas condiciones de regularidad dependiendo de cada problema (como continuidad, diferenciabilidad, entre otras).

2.3. Espacios Funcionales

En el cálculo variacional se trabaja con **espacios funcionales**, los cuales permiten imponer las condiciones de regularidad necesarias para cada problema. Aunque no existe un “espacio universal” aplicable a todos los casos, debido a la naturaleza de cada problema, sí es posible identificar ciertos espacios funcionales que aparecen con frecuencia en los problemas típicos del cálculo variacional, entre ellos:

- $C^k(\Omega)$: Espacio de funciones k -veces continuamente diferenciables en Ω (Evans 2022).
- $C_c^\infty(\Omega)$: Espacio de funciones infinitamente diferenciables con soporte compacto en Ω (Gelfand y Fomin 1963)
- $L^p(\Omega)$: Espacios de Lebesgue.
- **Espacio de Sobolev**: Denotado por $W^{k,p}(\Omega)$, es el espacio formado por funciones cuyas derivadas en el sentido débil hasta orden k existen y pertenecen a $L^p(\Omega)$ (Evans 2022).

Adicionalmente, se introducen los espacios **tangente** y **cotangente**, los cuales están asociados al conjunto de funciones admisibles $\mathcal{A}$. El *espacio tangente* $T_y \mathcal{A}$ representa las *variaciones admisibles* de $y \in \mathcal{A}$, es decir aquellas “perturbaciones” de y que preservan las restricciones impuestas por el problema. Mientras que el *espacio cotangente* $T_y^* \mathcal{A}$ es el dual del espacio tangente, es decir, contiene los funcionales lineales que actúan sobre las variaciones $\eta \in T_y \mathcal{A}$.

2.4. Continuidad de un Funcional

Un funcional $J : \mathcal{A} \rightarrow \mathbb{R}$ se dice **continuo** en $y \in \mathcal{A}$ si para toda sucesión $y_k \in \mathcal{A}$ tal que

$$\|y_k - y\|_{\mathcal{A}} \rightarrow 0,$$

se cumple que

$$J[y_k] \rightarrow J[y].$$

Además, decimos que el funcional J es **semicontinuo inferiormente** en $y \in \mathcal{A}$ si para toda sucesión $y_k \in \mathcal{A}$ tal que $\|y_k - y\|_{\mathcal{A}} \rightarrow 0$, se cumple que

$$J[y] \leq \liminf_{k \rightarrow \infty} J[y_k].$$

De manera análoga, J es **semicontinuo superiormente** en y si

$$J[y] \geq \limsup_{k \rightarrow \infty} J[y_k].$$

Ejemplo 2.2.

Ejemplo Continuidad

Consideremos el espacio de funciones continuas en $[0, 1]$ ($\mathcal{A} = C^0([0, 1])$), con la norma del supremo:

$$\|y\|_{\infty} = \sup_{x \in [0, 1]} |y(x)|.$$

Definimos el funcional

$$J[y] = \int_0^1 y(x) dx$$

Sea $y_k \in \mathcal{A}$ una sucesión que converge uniformemente a una función $y \in \mathcal{A}$, es decir, $\|y_k - y\|_{\infty} \rightarrow 0$ cuando $k \rightarrow \infty$. Dado que toda función continua en un intervalo cerrado y acotado es integrable, y la convergencia uniforme implica convergencia puntual, entonces por el teorema de convergencia uniforme, se tiene:

$$\lim_{k \rightarrow \infty} \int_0^1 y_k(x) dx = \int_0^1 \lim_{k \rightarrow \infty} y_k(x) dx = \int_0^1 y(x) dx.$$

Por lo tanto,

$$J[y_k] \rightarrow J[y]$$

Este resultado muestra que el funcional J es continuo en todo $\mathcal{A}$.

Ejemplo 2.3.

Funcional no Continuo

Sea el espacio $\mathcal{A} = C^0([0, 1])$, con la norma del supremo

$$\|y\|_{\infty} = \sup_{x \in [0, 1]} |y(x)|.$$

Definimos el funcional

$$J[y] = \begin{cases} 0 & \text{si } \sup_{x \in [0,1]} y(x) \leq 1, \\ 1 & \text{si } \sup_{x \in [0,1]} y(x) > 1. \end{cases}$$

Considerando la función constante $y(x) = 1$. Claramente,

$$\sup_x y(x) = 1 \Rightarrow J[y] = 0.$$

Ahora, definimos una sucesión de funciones continuas

$$y_k(x) = 1 + \frac{1}{k}, \quad \forall x \in [0, 1].$$

Esta sucesión converge uniformemente a la función y , ya que

$$\|y_k - y\|_\infty = \sup_{x \in [0,1]} |y_k(x) - y(x)| = \frac{1}{k} \rightarrow 0 \quad \text{cuando } k \rightarrow \infty.$$

Sin embargo, observamos que $\sup_x y_k(x) = 1 + \frac{1}{k} > 1$, así para todo k , se tendría $J[y_k] = 1$, y en consecuencia:

$$J[y_k] = 1 \nrightarrow 0 = J[y],$$

así que el funcional **no es continuo** en y .

Sin embargo, notemos que sí es semicontinuo inferiormente. Sean $\{y_k\} \subset \mathcal{A}$ una sucesión y $y \in \mathcal{A}$ una función tal que $y_k \rightarrow y$ uniformemente. Probaremos que

$$J[y] \leq \liminf_{k \rightarrow \infty} J[y_k]$$

Analizamos dos casos posibles, cuando $J[y] = 0$ y cuando $J[y] = 1$.

- **Caso 1:** $\sup_{x \in [0,1]} y(x) \leq 1$. Entonces, por definición, $J[y] = 0$. Además, como $y_k \rightarrow y$ uniformemente, para toda $\epsilon > 0$ existe $K \in \mathbb{N}$ tal que para todo $k \geq K$,

$$\sup_{x \in [0,1]} y_k(x) \leq \sup_{x \in [0,1]} y(x) + \epsilon \leq 1 + \epsilon.$$

Para ϵ lo suficientemente pequeño, se tiene que $\sup_{x \in [0,1]} y_k(x) \leq 1$, entonces $J[y_k] = 0$ para todo k suficientemente grande, así

$$\liminf_{k \rightarrow \infty} J[y_k] \geq 0 = J[y].$$

- **Caso 2:** $\sup_{x \in [0,1]} y(x) > 1$. Entonces $J[y] = 1$. Como $y_k \rightarrow y$ uniformemente, para k suficientemente grande, $\sup_{x \in [0,1]} y_k(x) > 1$, así que $J[y_k] = 1$ eventualmente. Por lo tanto, $\liminf_{k \rightarrow \infty} J[y_k] = 1 = J[y]$.

En ambos casos se cumple $J[y] \leq \liminf_{k \rightarrow \infty} J[y_k]$, es decir, el funcional es semicontinuo inferiormente.

2.5. Linealidad de un Funcional

Definición 2.1. Sea R un espacio lineal normado y sea $\varphi[h]$ un funcional en R . Decimos que $\varphi[h]$ es un funcional **lineal** (continuo) si satisface el *principio de homogeneidad*

$$\varphi[\alpha h] = \alpha \varphi[h],$$

para toda $h \in R$ y para cada número real α tal que $\alpha h \in R$, y el *principio de aditividad*

$$\varphi[h_1 + h_2] = \varphi[h_1] + \varphi[h_2]$$

para cada $h_1, h_2, h_1 + h_2 \in R$.

Ejemplo 2.4.

Ejemplo Linealidad

Consideremos el funcional $J[x] = \int_{t_0}^{t_f} x(t)dt$, donde x es una función continua. Veamos que J satisface las condiciones de linealidad:

- **Homogeneidad:** Para cualquier escalar α ,

$$J[\alpha x] = \int_{t_0}^{t_f} \alpha x(t)dt = \alpha \int_{t_0}^{t_f} x(t)dt = \alpha J[x].$$

- **Aditividad:** Para funciones x_1 y x_2 ,

$$\begin{aligned} J[x_1 + x_2] &= \int_{t_0}^{t_f} (x_1(t) + x_2(t))dt = \int_{t_0}^{t_f} x_1(t)dt + \int_{t_0}^{t_f} x_2(t)dt \\ &= J[x_1] + J[x_2]. \end{aligned}$$

Por lo tanto, J es un funcional lineal.

Ejemplo 2.5.

Ejemplo de no Linealidad

Ahora consideremos el funcional $F[x] = \int_{t_0}^{t_f} [x(t)]^2 dt$, con x una función continua.

En este caso al analizar el principio de homogeneidad observamos que

$$F[\alpha x] = \int_{t_0}^{t_f} [\alpha x(t)]^2 dt = \alpha^2 \int_{t_0}^{t_f} [x(t)]^2 dt$$

$$\alpha F[x] = \alpha \int_{t_0}^{t_f} [x(t)]^2 dt$$

Claramente $F[\alpha x] \neq \alpha F[x]$ para todo α . Así, el funcional F es no lineal.

2.6. El Incremento de un Funcional

Definimos el **incremento** de un funcional J en la dirección de una función η como:

$$\Delta J[y, \eta] := J[y + \eta] - J[y]. \quad (2.2)$$

En este contexto, η representa una *perturbación* de la función y dentro del espacio de funciones admisibles. Al analizar este incremento, particularmente en el caso donde las perturbaciones son pequeñas, nos permite evaluar la sensibilidad del funcional respecto a cambios en y . Este análisis resulta esencial para introducir el concepto de **variación** del funcional, el cual es una herramienta clave en la formulación de condiciones necesarias de optimalidad en el cálculo variacional.

Ejemplo 2.6.

Incrementos en el funcional de la Braquistócrona

Consideraremos el problema clásico de la **braquistócrona***, en el cual buscamos la curva óptima $y(x)$ por la cual una partícula, partiendo desde $A = (x_0, y(x_0))$, desciende hasta $B = (x_1, y(x_1))$ bajo el efecto de la gravedad, sin tomar en cuenta la fricción, en el menor tiempo posible. El funcional asociado al tiempo de descenso es:

$$J[y] = \int_{x_0}^{x_1} L(x, y, y') dx \quad \text{con} \quad L(x, y, y') = \frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y(x_1) - y)}}. \quad (2.3)$$

Para estudiar cómo varía este funcional bajo una perturbación de la curva, consideramos $y_\varepsilon(x) = y(x) + \varepsilon\eta(x)$, donde $\eta \in C^1([x_0, x_1])$ satisface $\eta(x_0) = \eta(x_1) = 0$.

Así, el incremento del funcional es

$$\begin{aligned}
\Delta J[y, \eta] &= J[y_\varepsilon] - J[y] \\
&= \int_{x_0}^{x_1} \left(\frac{\sqrt{1 + (y' + \varepsilon\eta')^2}}{\sqrt{2g(y(x_1) - y - \varepsilon\eta)}} \right) dx - \int_{x_0}^{x_1} \left(\frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y(x_1) - y)}} \right) dx \quad (2.4) \\
&= \int_{x_0}^{x_1} \left(\frac{\sqrt{1 + (y' + \varepsilon\eta')^2}}{\sqrt{2g(y(x_1) - y - \varepsilon\eta)}} - \frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y(x_1) - y)}} \right) dx.
\end{aligned}$$

2.7. La Primera Variación

Definición 2.2. Sea $J : \mathcal{A} \rightarrow \mathbb{R}$ un funcional definido sobre un espacio de funciones admisibles $\mathcal{A}$, y sea $y \in \mathcal{A}$. Decimos que J es **diferenciable** en y si existe una aplicación lineal

$$\delta J[y; \cdot] : T_y \mathcal{A} \rightarrow \mathbb{R},$$

tal que, para toda perturbación admisible $\eta \in T_y \mathcal{A}$, se cumple la relación

$$J[y + \eta] - J[y] = \delta J[y; \eta] + o(\eta),$$

donde

$$\lim_{\|\eta\| \rightarrow 0} \frac{o(\eta)}{\|\eta\|} = 0.$$

En este caso, $\delta J[y; \eta]$ es la **variación del funcional** J en la dirección de la perturbación η , y representa la parte lineal principal del incremento de J cuando y es perturbada por η . Es decir, mide la sensibilidad de J ante pequeñas variaciones de y en la dirección de η .

Si consideramos un funcional

$$J[y] := \int_{\Omega} L(x, y(x), Dy(x)) dx,$$

y consideremos perturbaciones de la forma:

$$y_\varepsilon(x) := y(x) + \varepsilon\eta(x), \quad \eta \in T_y \mathcal{A},$$

Entonces, el funcional evaluado en y_ε se expresa como:

$$J[y_\varepsilon] = \int_{\Omega} L(x, y(x) + \varepsilon\eta(x), Dy(x) + \varepsilon D\eta(x)) dx.$$

Calculamos la **primera variación** de J en dirección de η como:

$$\delta J[y; \eta] := \left. \frac{d}{d\varepsilon} J[y_\varepsilon] \right|_{\varepsilon=0} = \int_{\Omega} \sum_{j=1}^N \left[\frac{\partial L}{\partial y_j} \eta_j + \sum_{i=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \frac{\partial \eta_j}{\partial x_i} \right] dx. \quad (2.5)$$

Ejemplo 2.7.

Primera Variación Braquistócrona

Siguiendo con el problema de la Brachistócrona trabajado en el Ejemplo 2.6. La primera variación del funcional Ecuación 2.3 se obtiene derivando respecto a ε en $\varepsilon = 0$ el funcional evaluado en la función $y_\varepsilon(x) = y(x) + \varepsilon \eta(x)$:

$$\begin{aligned} \delta J[y; \eta] &= \left. \frac{d}{d\varepsilon} J[y_\varepsilon] \right|_{\varepsilon=0} = \int_{x_0}^{x_1} L(x, y + \varepsilon \eta, y' + \varepsilon \eta') dx \Big|_{\varepsilon=0} \\ &= \int_{x_0}^{x_1} (L_y \eta + L_{y'} \eta') dx \end{aligned}$$

donde $\eta(x_0) = \eta(x_1) = 0$ y

$$\begin{aligned} L_y &= \frac{\partial}{\partial y} \left(\frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y(x_1) - y)}} \right) = \frac{\sqrt{1 + (y')^2}}{2\sqrt{2g(y(x_1) - y)^{3/2}}} \\ L_{y'} &= \frac{\partial}{\partial y'} \left(\frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y(x_1) - y)}} \right) = \frac{(y')}{\sqrt{1 + (y')^2} \sqrt{2g(y(x_1) - y)}} \end{aligned}$$

al sustituir obtenemos que

$$\delta J[y; \eta] = \int_{x_0}^{x_1} \left[\frac{\sqrt{1 + (y')^2}}{2\sqrt{2g(y(x_1) - y)^{3/2}}} (\eta) + \frac{(y')}{\sqrt{1 + (y')^2} \sqrt{2g(y(x_1) - y)}} (\eta') \right] dx. \quad (2.6)$$

Esta expresión mide la sensibilidad del tiempo de descenso ante pequeñas perturbaciones de la curva y en la dirección de η .

Teorema 2.1. *La variación de un funcional diferenciable es única.*

Prueba. Sea $J : \mathcal{A} \rightarrow \mathbb{R}$ un funcional diferenciable definido sobre un espacio normado $\mathcal{A}$, y sea $t \in \mathcal{A}$. Supongamos que existen dos aplicaciones lineales

$$\varphi_1, \varphi_2 : T_y \mathcal{A} \rightarrow \mathbb{R}$$

tales que para toda perturbación $\eta \in T_y \mathcal{A}$, se cumple:

$$\begin{aligned}\Delta J[y + \eta] &= J[y + \eta] - J[y] = \varphi_1[\eta] + o_1(\eta), \\ \Delta J[y + \eta] &= J[y + \eta] - J[y] = \varphi_2[\eta] + o_2(\eta),\end{aligned}$$

donde se satisface

$$\lim_{\|\eta\| \rightarrow 0} \frac{o_1(\eta)}{\|\eta\|} = \lim_{\|\eta\| \rightarrow 0} \frac{o_2(\eta)}{\|\eta\|} = 0.$$

Restando ambas expresiones, obtenemos:

$$\varphi_1[\eta] - \varphi_2[\eta] = o_2(\eta) - o_1(\eta).$$

Definimos $\Phi[\eta] = \varphi_1[\eta] - \varphi_2[\eta]$, que es lineal por ser diferencia de funcionales lineales. Entonces,

$$\Phi[\eta] = o_2(\eta) - o_1(\eta).$$

Dividiendo ambos lados por $\|\eta\|$, obtenemos:

$$\left| \frac{\Phi[\eta]}{\|\eta\|} \right| = \left| \frac{o_2(\eta) - o_1(\eta)}{\|\eta\|} \right| \leq \left| \frac{o_2(\eta)}{\|\eta\|} \right| + \left| \frac{o_1(\eta)}{\|\eta\|} \right|.$$

Tomando el límite cuando $\|\eta\| \rightarrow 0$, implica que:

$$\lim_{\|\eta\| \rightarrow 0} \frac{\Phi[\eta]}{\|\eta\|} = \lim_{\|\eta\| \rightarrow 0} \frac{o_2[\eta]}{\|\eta\|} + \lim_{\|\eta\| \rightarrow 0} \frac{o_1[\eta]}{\|\eta\|} = 0 + 0 = 0$$

Y ya que Φ es lineal, el límite anterior implica que:

$$\Phi[\eta] = 0 \quad \forall \eta \in T_y \mathcal{A},$$

esto es

$$\varphi_1[\eta] = \varphi_2[\eta] \quad \forall \eta.$$

Así, la primera variación es única para cada funcional diferenciable. □

2.8. Máximo y Mínimo de un Funcional

En el cálculo variacional, el objetivo principal es encontrar funciones que hagan máximo o mínimo un funcional dado. Formalmente, dado un funcional $J : \mathcal{A} \rightarrow \mathbb{R}$ definido sobre un conjunto de funciones admisibles $\mathcal{A}$, decimos que J tiene un extremo en y^* si existe $\varepsilon > 0$ tal que para toda función $y \in \mathcal{A}$ con $\|y - y^*\|_{\mathcal{A}} < \varepsilon$ el incremento de J tiene el mismo signo. Si

$$J[y^*] \leq J[y] \quad (2.7)$$

entonces $J[y^*]$ es un **mínimo relativo**; y si

$$J[y^*] \geq J[y] \quad (2.8)$$

entonces $J[y^*]$ es un **máximo relativo**.

Diremos que $J[y^*]$ es un mínimo global (respectivamente máximo global) si Ecuación 2.7 (respectivamente Ecuación 2.8) es válida para ε arbitrariamente grandes. A la función y^* se le conoce como *extremal* y a $J[y^*]$ como un *extremo*.

Ejemplo 2.8.

Ejemplo de Mínimo de un Funcional

Consideremos el funcional de longitud de arco para una curva $y(x)$ que conecta dos puntos (x_0, y_0) y (x_1, y_1) :

$$J[y] = \int_{x_0}^{x_1} \sqrt{1 + (y'(x))^2} dx,$$

donde $y(x_0) = y_0$ y $y(x_1) = y_1$.

El problema consiste en encontrar la función $y^*(x)$ que minimiza $J[y]$, es decir, la curva de menor longitud entre los dos puntos. La solución es la línea recta:

$$y^*(x) = y_0 + \frac{y_1 - y_0}{x_1 - x_0}(x - x_0).$$

Para cualquier otra función $y(x)$ que cumpla las mismas condiciones de frontera, se tiene que $J[y] \geq J[y^*]$. Por lo tanto, la línea recta es la curva de longitud mínima entre dos puntos.

2.9. Lema y Teorema Fundamental del Cálculo Variacional

El objetivo de esta sección es presentar herramientas fundamentales para obtener las ecuaciones de *Euler-Lagrange* y dar condiciones necesarias de optimalidad. Una de las

herramientas claves en este contexto es el **Lema Fundamental del Cálculo Variacional**.

Lema 2.1 (Fundamental del Cálculo Variacional). Sea $\Omega \subset \mathbb{R}^n$ un conjunto abierto, y sea f una función continua definida en Ω , y supongamos que

$$\int_{\Omega} f(x)\eta(x)dx \geq 0 \quad \forall \eta \in C_0^\infty(\Omega) \text{ con } \eta \geq 0,$$

o

$$\int_{\Omega} f(x)\eta(x)dx = 0 \quad \forall \eta \in C_c^\infty(\Omega),$$

se cumple. Entonces tendríamos que

$$f(x) \geq 0 \quad \text{o} \quad f(x) = 0,$$

respectivamente, para todo $x \in \Omega$.

Prueba. Supongamos que se cumple la primera afirmación, y que existe un punto $x_0 \in \Omega$ tal que $f(x_0) < 0$. Por continuidad de f , existe $\epsilon > 0$ y una bola $B_r(x_0) \subset \Omega$ tal que

$$f(x) < -\epsilon \quad \text{para todo } x \in B_r(x_0).$$

Mediante la función de prueba $\eta \in C_c^\infty$, con soporte compacto en $B_r(x_0)$, definida por

$$\eta(x) := \begin{cases} \exp\left(-\frac{1}{r^2-|x-x_0|^2}\right) & \text{si } x \in B_r(x_0) \\ 0 & \text{si } x \notin B_r(x_0). \end{cases}$$

Sabemos que esta función es suave y tiene soporte compacto en $B_r(x_0) \subset \Omega$, además de ser estrictamente positiva dentro de la bola.

Ya que $f(x) < -\epsilon$ en el soporte de η , podemos deducir que:

$$\int_{\Omega} f(x)\eta(x)dx = \int_{B_r(x_0)} f(x)\eta(x)dx < -\epsilon \int_{B_r(x_0)} \eta(x)dx < 0,$$

lo cual contradice la hipótesis de que la integral es no negativa para toda función de prueba no negativa. Por lo tanto, $f(x) \geq 0$ para todo $x \in \Omega$.

La segunda afirmación se vuelve un resultado inmediato consecuencia de la que se acaba de probar. $\square$

Teorema 2.2. Sean $J[y]$ un funcional diferencial, si y^* es un extremal de J , entonces la variación del funcional se anula para toda η admisible, es decir

$$\delta J[y^*; \eta] = 0 \quad \forall \eta \text{ admisible.}$$

Prueba. Sea $J : \mathcal{A} \rightarrow \mathbb{R}$ un funcional y supongamos que $y^* \in \mathcal{A}$ es un extremal de J . Consideremos perturbaciones de la forma $y_\epsilon = y^* + \epsilon\eta$, con $\epsilon > 0$ y $\eta \in T_{y^*}\mathcal{A}$ y definimos la función $\phi : \mathbb{R} \rightarrow \mathbb{R}$ tal que

$$\phi(\epsilon) := J[y^* + \epsilon\eta]$$

Como por hipótesis y^* es un extremal, entonces ϕ tiene un extremo relativo en $\epsilon = 0$. Por cálculo diferencial, una condición necesaria para que ϕ tenga un extremo en $\epsilon = 0$ es que su derivada se anule en ese punto, es decir,

$$\left. \frac{d}{d\epsilon} \phi(\epsilon) \right|_{\epsilon=0} = 0.$$

Así

$$\left. \frac{d}{d\epsilon} \phi(\epsilon) \right|_{\epsilon=0} = \left. \frac{d}{d\epsilon} J[y^* + \epsilon\eta] \right|_{\epsilon=0} = \delta J[y^*; \eta].$$

Por lo tanto,

$$\delta J[y^*; \eta] = 0 \quad \forall \eta \text{ admisible.}$$

□

2.10. Ecuaciones de Euler

Para deducir las ecuaciones de Euler, es necesario que todos los términos η_j en la primera variación aparezcan libre de derivadas, con el fin de aplicar el Lema Fundamental (Lema 2.1). Para ello aplicamos el método de integración por partes en varias variables al segundo término de la primera variación (Ecuación 2.5):

$$\int_{\Omega} \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \frac{\partial \eta_j}{\partial x_i} dx = - \int_{\Omega} \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \eta_j dx.$$

Al repetir este procedimiento para toda $i = 1, \dots, n$ y toda $j = 1, \dots, N$, obtenemos:

$$\int_{\Omega} \sum_{i=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \frac{\partial \eta_j}{\partial x_i} dx = - \int_{\Omega} \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \eta_j dx.$$

Sustituyendo en la primera variación, se obtiene que:

$$\begin{aligned}
\delta J[y; \eta] &= \int_{\Omega} \sum_{j=1}^N \left[\frac{\partial L}{\partial y_j} \eta_j + \sum_{i=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \frac{\partial \eta_j}{\partial x_i} \right] dx \\
&= \int_{\Omega} \sum_{j=1}^N \left[\frac{\partial L}{\partial y_j} \eta_j - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \eta_j \right] dx \\
&= \int_{\Omega} \sum_{j=1}^N \left(\frac{\partial L}{\partial y_j} - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \right) \eta_j dx.
\end{aligned}$$

Ahora bien, si y representa un extremo del funcional J , entonces, de acuerdo con el (Teorema 2.2), la primera variación debe anularse para toda función $\eta \in C_c^{\infty}(\Omega; \mathbb{R}^N)$, es decir:

$$\delta J[y; \eta] = 0 \quad \forall \eta.$$

Por tanto se cumple que:

$$\int_{\Omega} \sum_{j=1}^N \left(\frac{\partial L}{\partial y_j} - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \right) \eta_j dx = 0 \quad \forall \eta.$$

Aplicando el Lema Fundamental del Cálculo Variacional (Lema 2.1) para cada $j = 1, \dots, N$, se obtiene la siguiente igualdad:

$$\frac{\partial L}{\partial y_j} - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) = 0. \quad (2.9)$$

Esta familia de ecuaciones, constituye el sistema de **ecuaciones de Euler-Lagrange**, el cual deben satisfacer todas las funciones y que hacen extremal al funcional J .

Ejemplo 2.9.

Primera Variación Braquistócrona simplificada

Aplicando integración por partes al término que involucra η' en la expresión de la primera variación (Ecuación 2.6), y usando que $\eta(x_0) = \eta(x_1) = 0$, obtenemos:

$$\int_{x_0}^{x_1} L_{y'} \eta' dx = [L_{y'} \eta]_{x_0}^{x_1} - \int_{x_0}^{x_1} \frac{d}{dx} L_{y'} \eta dx = - \int_{x_0}^{x_1} \frac{d}{dx} L_{y'} \eta dx,$$

ya que los valores de frontera se anulan.

Por lo tanto, la primera variación se puede escribir como:

$$\begin{aligned}\delta J[y; \eta] &= \int_{x_0}^{x_1} \left(L_y - \frac{d}{dx} L_{y'} \right) \eta dx \\ &= \int_{x_0}^{x_1} \left[\frac{\sqrt{1 + (y')^2}}{2\sqrt{2g}(y(x_1) - y)^{3/2}} - \frac{d}{dx} \left(\frac{(y')}{\sqrt{1 + (y')^2} \sqrt{2g}(y(x_1) - y)} \right) \right] (\eta) dx.\end{aligned}$$

Así, de acuerdo al Teorema 2.2, y por el Lema 2.1, para que y sea un extremal debe satisfacer la ecuación de Euler-Lagrange:

$$\frac{\sqrt{1 + (y')^2}}{2\sqrt{2g}(y(x_1) - y)^{3/2}} - \frac{d}{dx} \left(\frac{(y')}{\sqrt{1 + (y')^2} \sqrt{2g}(y(x_1) - y)} \right) = 0. \quad (2.10)$$

2.11. Teorema de Noether

Ya se han establecido la primera variación del funcional y las correspondientes ecuaciones de Euler–Lagrange en el marco multivariable. A partir de dichos resultados es posible abordar una cuestión fundamental: la relación entre las simetrías de un funcional y la existencia de leyes de conservación. Este vínculo, formulado de manera precisa por **Emmy Noether** en 1918, constituye uno de los resultados más profundos y de mayor alcance del cálculo variacional y sus extensiones.

De manera general, el teorema de Noether establece que si un funcional es invariante bajo un grupo uniparamétrico de transformaciones (en el tiempo, el espacio, las variables dependientes o una combinación de ellas), entonces existe una cantidad conservada asociada a tal simetría. Esta correspondencia entre invariancia y conservación no solo es esencial en física matemática, sino que también ha encontrado aplicaciones en la teoría del control óptimo y en formulaciones con retardos temporales (véase ?).

En lo que sigue presentamos el enunciado del teorema en el contexto multivariable, adoptando la notación ya introducida anteriormente.

Comencemos considerando el problema de cálculo variacional, el minimizar Ecuación 2.1 bajo condiciones de frontera dadas $y(a) = y_a$ y $y(b) = y_b$.

Definición 2.3 (Invarianza de un Funcional). Sea

$$J[y] = \int_{\Omega} L(x, y(x), Dy(x)) dx,$$

un funcional definido sobre un dominio $\Omega \subset \mathbb{R}^n$, con $y : \Omega \rightarrow \mathbb{R}^m$ y $Dy(x) \in \mathbb{R}^{m \times n}$.

Considerando la transformación

$$\begin{cases} \bar{x} = x + \varepsilon \eta(x, y) + o(\varepsilon), \\ \bar{y}(x) = y(x) + \varepsilon \xi(x, y) + o(\varepsilon), \end{cases} \quad (2.11)$$

donde

$$\eta : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}^n, \quad \xi : \Omega \times \mathbb{R}^m \rightarrow \mathbb{R}^m$$

son funciones dadas. El funcional J se dice ser **invariante** bajo un grupo uniparamétrico de transformaciones infinitesimales (Ecuación 2.11) si,

$$J[y] = J[\bar{y}], \quad (2.12)$$

para toda y admisible y para cualquier subdominio $\tilde{\Omega} \subseteq \Omega$

Teorema 2.3 (Condiciones Necesarias de Invarianza). *Si un funcional (Ecuación 2.1) es invariante bajo transformaciones (Ecuación 2.11), entonces*

$$\begin{aligned} \sum_{j=1}^n \frac{\partial L}{\partial x_j} \eta_j(x, y) + \sum_{i=1}^m \frac{\partial L}{\partial y_i} \xi_i(x, y) + \sum_{i=1}^m \sum_{j=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \left(\frac{\partial \xi_i}{\partial x_j} - \sum_{k=1}^n \frac{\partial y_i}{\partial x_k} \frac{\partial \eta_k}{\partial x_j} \right) \\ + L(x, y, Dy) \sum_{j=1}^n \frac{\partial \eta_j}{\partial x_j} = 0. \end{aligned} \quad (2.13)$$

Prueba. La invarianza de J (Ecuación 2.12) implica que, para todo subconjunto $\Omega' \subseteq \Omega$,

$$\int_{\Omega'} L(x, y(x), Dy(x)) dx = \int_{\tilde{\Omega}'} L(\tilde{x}, \tilde{y}(\tilde{x}), D\tilde{y}(\tilde{x})) d\tilde{x}, \quad (2.14)$$

donde $\tilde{\Omega}'$ es la imagen de Ω' bajo la transformación infinitesimal $\tilde{x} = x + \varepsilon \eta(x, y) + o(\varepsilon)$.

Haciendo el cambio de variable $\tilde{x} = \tilde{x}(x) = x + \varepsilon \eta(x, y) + o(\varepsilon)$ en la integral del lado derecho de la igualdad, obtenemos

$$\int_{\Omega'} L(x, y(x), Dy(x)) dx = \int_{\tilde{\Omega}'} L(\tilde{x}(x), \tilde{y}(x), D\tilde{y}(x)) \det \left(\frac{\partial \tilde{x}}{\partial x} \right) dx.$$

Dado que (Ecuación 2.14) se cumple para cualquier subconjunto Ω' de Ω , podemos remover los signos de integrales y escribir la igualdad equivalente

$$L(x, y(x), Dy(x)) = L(\tilde{x}(x), \tilde{y}(x), D\tilde{y}(x)) \det \left(\frac{\partial \tilde{x}}{\partial x} \right). \quad (2.15)$$

Derivando ambos lados de la ecuación (Ecuación 2.15) respecto a ε y evaluando en $\varepsilon = 0$, se obtiene la condición necesaria de invarianza (Ecuación 2.13).

El lado izquierdo, al no depender de ε , su derivada es cero. Así que pasamos a la derivada del lado derecho

Aplicamos la regla de la cadena para $L(\bar{x}, \bar{y}, D\bar{y})$:

$$\frac{d}{d\varepsilon} L(\bar{x}, \bar{y}, D\bar{y}) \Big|_{\varepsilon=0} = \sum_{j=1}^n \frac{\partial L}{\partial x_j} \eta_j(x, y) + \sum_{i=1}^m \frac{\partial L}{\partial y_i} \xi_i(x, y) + \sum_{i=1}^m \sum_{j=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \frac{d}{d\varepsilon} \left(\frac{\partial \bar{y}_i}{\partial \bar{x}_j} \right) \Big|_{\varepsilon=0}$$

Calculamos la derivada de $D\bar{y}$ usando la regla de la cadena:

$$\frac{\partial \bar{y}_i}{\partial \bar{x}_j} = \frac{\partial \bar{y}_i}{\partial x_k} \frac{\partial x_k}{\partial \bar{x}_j}$$

Para transformaciones infinitesimales, expandimos:

$$\frac{d}{d\varepsilon} \left(\frac{\partial \bar{y}_i}{\partial \bar{x}_j} \right) \Big|_{\varepsilon=0} = \frac{\partial \xi_i}{\partial x_j} - \sum_{k=1}^n \frac{\partial y_i}{\partial x_k} \frac{\partial \eta_k}{\partial x_j}$$

La derivada del determinante jacobiano es:

$$\frac{d}{d\varepsilon} \det \left(\frac{\partial \bar{x}}{\partial x} \right) \Big|_{\varepsilon=0} = \sum_{j=1}^n \frac{\partial \eta_j}{\partial x_j}$$

Por la regla del producto:

$$\frac{d}{d\varepsilon} \left(L(\bar{x}, \bar{y}, D\bar{y}) \det \left(\frac{\partial \bar{x}}{\partial x} \right) \right) \Big|_{\varepsilon=0} = \frac{d}{d\varepsilon} L(\bar{x}, \bar{y}, D\bar{y}) \Big|_{\varepsilon=0} + L(x, y, Dy) \sum_{j=1}^n \frac{\partial \eta_j}{\partial x_j}$$

Igualando la derivada total a cero, obtenemos:

$$\begin{aligned} \sum_{j=1}^n \frac{\partial L}{\partial x_j} \eta_j(x, y) + \sum_{i=1}^m \frac{\partial L}{\partial y_i} \xi_i(x, y) + \sum_{i=1}^m \sum_{j=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \left(\frac{\partial \xi_i}{\partial x_j} - \sum_{k=1}^n \frac{\partial y_i}{\partial x_k} \frac{\partial \eta_k}{\partial x_j} \right) \\ + L(x, y, Dy) \sum_{j=1}^n \frac{\partial \eta_j}{\partial x_j} = 0. \end{aligned}$$

Esto concluye la demostración. □

Definición 2.4 (Constante de Movimiento / Ley de Conservación). Decimos que $C : \Omega \times \mathbb{R}^m \times \mathbb{R}^{m \times n} \rightarrow \mathbb{R}^n$ es una **constante de movimiento** (o **ley de conservación**) para (Ecuación 2.1), si satisface

$$\operatorname{div} C(x, y(x), Dy(x)) = 0, \quad (2.16)$$

para todo extremal $y(\cdot)$.

Teorema 2.4 (Teorema de DuBois-Reymond (Multivariable)). *Sea $J[y]$ un funcional (Ecuación 2.1) definido sobre funciones $y : \Omega \rightarrow \mathbb{R}^m$ suficientemente regulares, donde $Dy(x)$ denota la matriz de derivadas parciales de y . Si y^* es un extremal de J , entonces satisface la condición de DuBois-Reymond*

$$\frac{\partial L}{\partial x}(x, y^*, Dy^*) = \text{Div} \left(L(x, y^*, Dy^*) I_n - \frac{\partial L}{\partial Dy}(x, y^*, Dy^*) Dy^* \right) \quad (2.17)$$

Prueba. Sea y^* un extremal del funcional (Ecuación 2.1), y consideremos la función

$$F_i(x) := L(x, y^*, Dy^*) - \sum_{j=1}^m \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)}(x, y^*, Dy^*) \frac{\partial y_j^*}{\partial x_i}, \quad i = 1, \dots, n.$$

Al calcular la derivada con respecto a x_i , obtenemos:

$$\begin{aligned} \frac{\partial F_i}{\partial x_i} &= \frac{\partial L}{\partial x_i} + \sum_{j=1}^m \left[\frac{\partial L}{\partial y_j} \frac{\partial y_j^*}{\partial x_i} + \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k} \right] \\ &\quad - \sum_{j=1}^m \left[\frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \frac{\partial y_j^*}{\partial x_i} + \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \frac{\partial^2 y_j^*}{\partial x_i^2} \right]. \end{aligned}$$

Notemos que el término con $k = i$ dentro de la suma $\sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k}$ se cancela con el último término, de modo que

$$\frac{\partial F_i}{\partial x_i} = \frac{\partial L}{\partial x_i} + \sum_{j=1}^m \left[\frac{\partial L}{\partial y_j} \frac{\partial y_j^*}{\partial x_i} + \sum_{\substack{k=1 \\ k \neq i}}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k} - \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \frac{\partial y_j^*}{\partial x_i} \right].$$

Sumando sobre $i = 1, \dots, n$. Esto produce

$$\begin{aligned} \sum_{i=1}^n \frac{\partial F_i}{\partial x_i} &= \sum_{i=1}^n \frac{\partial L}{\partial x_i} + \sum_{j=1}^m \sum_{i=1}^n \frac{\partial L}{\partial y_j} \frac{\partial y_j^*}{\partial x_i} \\ &\quad + \sum_{j=1}^m \sum_{i=1}^n \sum_{\substack{k=1 \\ k \neq i}}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k} - \sum_{j=1}^m \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \frac{\partial y_j^*}{\partial x_i} \end{aligned}$$

Ahora, notemos que el tercer término involucra una doble suma sobre i y k con $i \neq k$. Podemos reorganizarla como

$$\sum_{i=1}^n \sum_{\substack{k=1 \\ k \neq i}}^n A_{ik} = \sum_{k=1}^n \sum_{\substack{i=1 \\ i \neq k}}^n A_{ik}$$

Por lo tanto, para cada j :

$$\sum_{i=1}^n \sum_{\substack{k=1 \\ k \neq i}}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k} = \sum_{k=1}^n \sum_{\substack{i=1 \\ i \neq k}}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k}$$

Ahora, sumando el término $k = i$ (que fue excluido), notamos que:

$$\sum_{i=1}^n \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k} = \sum_{k=1}^n \sum_{i=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k}$$

Pero el término con $k = i$ es:

$$\sum_{i=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \frac{\partial^2 y_j^*}{\partial x_i^2}$$

Por lo tanto, la suma sobre $k \neq i$ más la suma sobre $k = i$ recupera la suma total:

$$\sum_{i=1}^n \sum_{\substack{k=1 \\ k \neq i}}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k} + \sum_{i=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \frac{\partial^2 y_j^*}{\partial x_i^2} = \sum_{i=1}^n \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k}$$

Ahora, agrupando todos los términos para cada j :

$$\begin{aligned} \sum_{i=1}^n \frac{\partial F_i}{\partial x_i} &= \sum_{i=1}^n \frac{\partial L}{\partial x_i} + \sum_{j=1}^m \left[\sum_{i=1}^n \frac{\partial L}{\partial y_j} \frac{\partial y_j^*}{\partial x_i} + \left(\sum_{i=1}^n \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k} - \sum_{i=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \frac{\partial^2 y_j^*}{\partial x_i^2} \right) \right. \\ &\quad \left. - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \frac{\partial y_j^*}{\partial x_i} \right] \end{aligned}$$

Pero $\sum_{i=1}^n \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \frac{\partial^2 y_j^*}{\partial x_i \partial x_k}$ es igual a $\sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_k} \right)} \sum_{i=1}^n \frac{\partial^2 y_j^*}{\partial x_i \partial x_k}$.

Ahora, por la simetría de las derivadas cruzadas (asumiendo suficiente regularidad),

$$\sum_{i=1}^n \frac{\partial^2 y_j^*}{\partial x_i \partial x_k} = \frac{\partial}{\partial x_k} \left(\sum_{i=1}^n \frac{\partial y_j^*}{\partial x_i} \right).$$

Notemos que la ecuación de Euler-Lagrange (Ecuación 2.9) nos permite cancelar los términos por lo tanto, para cada j :

$$\frac{\partial L}{\partial y_j} - \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) = 0.$$

Por lo tanto, para cada j

$$\sum_{i=1}^n \frac{\partial L}{\partial y_j} \frac{\partial y_j^*}{\partial x_i} = \sum_{i=1}^n \frac{\partial}{\partial x_i} \left(\frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)} \right) \frac{\partial y_j^*}{\partial x_i}.$$

Así, los términos se cancelan y obtenemos

$$\sum_{i=1}^n \frac{\partial F_i}{\partial x_i} = \sum_{i=1}^n \frac{\partial L}{\partial x_i},$$

que por la definición de divergencia, da paso a

$$\text{Div} \left(L(x, y^*, Dy^*) I_n - \frac{\partial L}{\partial Dy}(x, y^*, Dy^*) Dy^* \right) = \frac{\partial L}{\partial x}(x, y^*, Dy^*).$$

Por lo tanto, se cumple la identidad de DuBois-Reymond:

$$\frac{\partial L}{\partial x}(x, y^*, Dy^*) = \text{Div} \left(L(x, y^*, Dy^*) I_n - \frac{\partial L}{\partial Dy}(x, y^*, Dy^*) Dy^* \right).$$

□

Teorema 2.5 (Teorema de Noether). *Si J (Ecuación 2.1) es **invariante** bajo transformaciones infinitesimales (Ecuación 2.11), entonces*

$$C(x, y, Dy) := \sum_{i=1}^m \sum_{j=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \xi_j + \sum_{j=1}^n \left(L(x, y, Dy) - \sum_{i=1}^m \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_k} \right)} \frac{\partial y_i}{\partial x_k} \right) \eta_j$$

satisface la ley de conservación (Definición 2.4)

Prueba. Sea y^* un extremal del funcional J , y supongamos que este es invariante bajo transformaciones infinitesimales (Ecuación 2.11). Por la **condición necesaria de invarianza** (Ecuación 2.13), se cumple

$$\sum_{j=1}^n \frac{\partial L}{\partial x_j} \eta_j + \sum_{i=1}^m \frac{\partial L}{\partial y_i} \xi_i + \sum_{i=1}^m \sum_{j=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \left(\frac{\partial \xi_i}{\partial x_j} - \sum_{k=1}^n \frac{\partial y_i^*}{\partial x_k} \frac{\partial \eta_k}{\partial x_j} \right) + L \sum_{j=1}^n \frac{\partial \eta_j}{\partial x_j} = 0.$$

Usando las **ecuaciones de Euler-Lagrange** (Ecuación 2.9) para cada i , podemos reescribir el segundo término como

$$\sum_{i=1}^m \frac{\partial L}{\partial y_i} \xi_i = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \right) \xi_i.$$

Sustituyendo en la condición de invarianza

$$\begin{aligned} \sum_{j=1}^n \frac{\partial L}{\partial x_j} \eta_j + \sum_{i=1}^m \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \right) \xi_i + \sum_{i=1}^m \sum_{j=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \frac{\partial \xi_i}{\partial x_j} \\ - \sum_{i=1}^m \sum_{j=1}^n \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \frac{\partial y_i^*}{\partial x_k} \frac{\partial \eta_k}{\partial x_j} + L \sum_{j=1}^n \frac{\partial \eta_j}{\partial x_j} = 0. \end{aligned}$$

Agrupando términos que involucran a ξ_i tenemos que

$$\sum_{i=1}^m \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \xi_i \right) = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial}{\partial x_j} \left(\frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \right) \xi_i + \sum_{i=1}^m \sum_{j=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \frac{\partial \xi_i}{\partial x_j}.$$

De manera análoga, los términos que involucran a η_j , pueden escribirse como

$$\sum_{j=1}^n \frac{\partial}{\partial x_j} \left[\left(L(x, y, Dy) - \sum_{i=1}^m \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_k} \right)} \frac{\partial y_i}{\partial x_k} \right) \eta_j \right].$$

Por lo que la condición de invarianza se transforma en la siguiente expresión

$$\sum_{j=1}^n \frac{\partial}{\partial x_j} \left[\sum_{i=1}^m \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \xi_i + \left(L(x, y, Dy) - \sum_{i=1}^m \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_k} \right)} \frac{\partial y_i}{\partial x_k} \right) \eta_j \right] = 0$$

Con esto definimos el **vector conservado** $C \in \mathbb{R}^n$ por componentes

$$C_j(x, y, Dy) = \sum_{i=1}^m \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_j} \right)} \xi_i + \left(L(x, y, Dy) - \sum_{i=1}^m \sum_{k=1}^n \frac{\partial L}{\partial \left(\frac{\partial y_i}{\partial x_k} \right)} \frac{\partial y_i}{\partial x_k} \right) \eta_j,$$

para toda $j = 1, \dots, n$. Finalmente, la condición de invarianza se convierte en

$$\text{Div } C(x, y, Dy) := \sum_{j=1}^n \frac{\partial C_j}{\partial x_j} = 0.$$

Con esto concluimos que C satisface la ley de conservación. $\square$

2.12. Transformación de Legendre y Hamiltoniano

La **transformación de Legendre** es una herramienta esencial en el cálculo variacional y en la física matemática, ya que permite reformular una función (o funcional) expresada en ciertas variables y sus derivadas en función de variables duales. En el contexto del cálculo variacional, se utiliza para pasar del Lagrangiano L al Hamiltoniano H .

Definición 2.5. Sea una función $f : \mathbb{R}^n \rightarrow \mathbb{R}$. Definimos la transformada **Legendre-Fenchel** (LF) de $f(x)$ por la formula variacional

$$f^*(p) = \sup_{x \in \mathbb{R}^n} \{ \langle p, x \rangle - f(x) \}. \quad (2.18)$$

Observación 2.1. También es común definir la transformada de Legendre-Fenchel usando el ínfimo (mínimo) en lugar del supremo:

$$g^*(p) = \inf_{x \in \mathbb{R}^n} \{ \langle p, x \rangle - g(x) \}. \quad (2.19)$$

El pasar de una versión de la transformada LF a otra es simplemente cuestión de introducir un cambio de signo. Esto es,

$$-f^*(p) = -\sup_{x \in \mathbb{R}^n} \{ \langle p, x \rangle - f(x) \} = \inf_{x \in \mathbb{R}^n} \{ -\langle p, x \rangle + f(x) \}.$$

de modo que

$$g^*(q) = \inf_{x \in \mathbb{R}^n} \{ \langle q, x \rangle - g(x) \},$$

haciendo las transformaciones $p = -q$ y $g = -f$, se obtiene la equivalencia entre ambas definiciones (Touchette 2005).

Observación 2.2. Si la función $f : \mathbb{R}^n \rightarrow \mathbb{R}$ es convexa y diferenciable, el supremo (respectivamente el ínfimo) en la definición de la Ecuación 2.18 (respectivamente en la Ecuación 2.19) de la transformada de Legendre-Fenchel, se alcanza en un punto x_p que satisface la condición de primer orden

$$p = \nabla f(x_p).$$

Por lo tanto, la transformada de Legendre-Fenchel se puede definir por

$$f^*(p) = \langle p, x_p \rangle - f(x_p), \quad (2.20)$$

Esta expresión constituye la *transformación de Legendre clásica*, que aparece con frecuencia en problemas variacionales, en termodinámica y en mecánica hamiltoniana (Touchette 2005).

Aplicando la transformada de Legendre sobre el Lagrangiano $L(x, y, Dy)$ de un funcional (Ecuación 2.1) obtenemos la función transformada:

$$H(x, y, p) = \sup_{Dy} \left\{ \sum_{j=1}^N \sum_{i=1}^n p_{j,i} \frac{\partial y_j}{\partial x_i} - L(x, y, Dy) \right\}, \quad (2.21)$$

donde los *momentos conjugados* $p_{j,i}$ se definen por

$$p_{j,i} := \frac{\partial L}{\partial \left(\frac{\partial y_j}{\partial x_i} \right)}.$$

A esta nueva función (Ecuación 2.21), se le conoce como el **Hamiltoniano** asociado al problema variacional. Una vez definido, es posible describir la evolución de las trayectorias óptimas mediante las **ecuaciones de Hamilton**, las cuales constituyen un sistema equivalente a las ecuaciones de Euler-Lagrange (Ecuación 2.9). Esta reformulación es especialmente útil en física y en teoría de control óptimo.

En el caso general las ecuaciones de Hamilton asociadas se expresan como:

$$\frac{\partial y_j}{\partial x_i} = \frac{\partial H}{\partial p_{j,i}}, \quad \frac{\partial p_{j,i}}{\partial x_i} = -\frac{\partial H}{\partial y_j} \quad \forall i = 1, \dots, n; j = 1, \dots, N.$$

En el caso particular en que $n = 1 = N$, es decir, para una función $y(x)$ de una sola variable, el Hamiltoniano toma la forma

$$H(x, y, p) = \sup_{y'} \{py' - L(x, y, y')\},$$

donde $p = \frac{\partial L}{\partial y'}$.

Las ecuaciones de Hamilton asociadas se reducen al sistema clásico:

$$\begin{aligned} \frac{dy}{dx} &= \frac{\partial H}{\partial p}, \\ \frac{dp}{dx} &= -\frac{\partial H}{\partial y}. \end{aligned}$$

Estas ecuaciones constituyen la base de la mecánica hamiltoniana y también se utilizan en el análisis de problemas de control óptimo, donde el momento p se interpreta como un multiplicador asociado a las restricciones dinámicas del sistema.

Ejemplo 2.10.

La Braquistócrona y su Hamiltoniano

En secciones anteriores analizamos el problema de la braquistócrona desde una perspectiva variacional clásica. A partir del funcional de tiempo recorrido, dedujimos su incremento (Ecuación 2.4), la primera variación (Ecuación 2.6) y finalmente la correspondiente ecuación de Euler (Ecuación 2.10). Ahora, retomamos el mismo problema para ilustrar cómo se construye y se interpreta el Hamiltoniano asociado, así como las ecuaciones de Hamilton correspondientes.

Recordemos que el funcional del problema está dado por:

$$J[y] = \int_{x_0}^{x_1} L(x, y, y') dx \quad \text{con} \quad L(x, y, y') = \frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y(x_1) - y)}}.$$

Siguiendo la transformación de Legendre (Definición 2.5), definimos el momento conjugado como:

$$p := \frac{\partial L}{\partial y'} = \frac{(y')}{\sqrt{1 + (y')^2} \sqrt{2g(y(x_1) - y)}}.$$

Así, el Hamiltoniano asociado se define mediante

$$H(x, y, p) = py' - L(x, y, y').$$

Como queremos obtener una expresión explícita en términos de p y y , de la relación anterior obtenemos que

$$y' = \frac{p}{\sqrt{1 - 2g(y(x_1) - y)p^2}} \quad \text{y} \quad L(x, y, y') = \frac{1}{\sqrt{1 - 2g(y(x_1) - y)p^2}} \cdot \frac{1}{2g(y(x_1) - y)}.$$

Sustituyendo en la definición del Hamiltoniano obtenemos

$$\begin{aligned} H(x, y, p) &= \frac{p^2}{\sqrt{1 - 2g(y(x_1) - y)p^2}} - \left(\frac{1}{\sqrt{1 - 2g(y(x_1) - y)p^2}} \cdot \frac{1}{2g(y(x_1) - y)} \right) \\ &= \frac{1}{\sqrt{1 - 2g(y(x_1) - y)p^2}} \left(p^2 - \frac{1}{2g(y(x_1) - y)} \right). \end{aligned}$$

2.13. Problema Primal y Dual en Cálculo Variacional

En el cálculo variacional, el **problema primal** consiste en encontrar una función y^* dentro de un conjunto de funciones admisibles $\mathcal{A}$ que minimice (o maximice) un funcional dado $J[y]$. Formalmente, el problema primal se expresa como:

$$\min_{y \in \mathcal{A}} J[y],$$

donde $J[y]$ es típicamente de la forma

$$J[y] = \int_{\Omega} L(x, y(x), Dy(x)) dx,$$

como se definió en la sección de **Funcionales en Varias Variables**.

Las condiciones necesarias para que y^* sea óptima se obtienen a través de la anulación de la primera variación, lo que conduce a las **ecuaciones de Euler-Lagrange** (ver Ecuación 2.9), que deben satisfacer todas las funciones extremales del funcional.

El **problema dual** surge al asociar al problema primal una formulación alternativa, generalmente a través de la transformación de Legendre-Fenchel (ver Definición 2.5), que permite expresar el problema en términos de variables duales (momentos conjugados). El dual típicamente busca maximizar (o minimizar) un funcional dual $J^*[p]$ sobre un conjunto dual de funciones o campos p , y está íntimamente relacionado con el Hamiltoniano H construido a partir del Lagrangiano L .

La relación entre el problema primal y el dual se fundamenta en la teoría de la dualidad convexa, donde bajo ciertas condiciones (por ejemplo, convexidad y regularidad del funcional), ambos problemas tienen el mismo valor óptimo (sin *brecha de dualidad*).

Ejemplo 2.11.

Ejemplo: Problema Primal y Dual de la Braquistócrona

En el caso del problema de la **braquistócrona**, donde se busca la curva $y(x)$ que minimiza el tiempo de descenso de una partícula bajo gravedad entre dos puntos fijos. El **problema primal** es:

$$\min_{y \in \mathcal{A}} J[y], \quad J[y] = \int_{x_0}^{x_1} \frac{\sqrt{1 + (y')^2}}{\sqrt{2g(y(x_1) - y)}} dx,$$

donde $\mathcal{A}$ es el conjunto de funciones y tales que $y(x_0) = y_0$, $y(x_1) = y_1$ y $y(x) < y(x_1)$ para $x \in (x_0, x_1)$.

La condición de optimalidad para el problema primal es que y satisfaga la ecuación de Euler-Lagrange correspondiente (Ecuación 2.10).

Por otro lado, el **problema dual** se formula introduciendo el momento conjugado p asociado a y' , definido por

$$p = \frac{\partial L}{\partial y'} = \frac{y'}{\sqrt{1 + (y')^2} \sqrt{2g(y(x_1) - y)}},$$

y construyendo el Hamiltoniano (Ejemplo 2.10):

$$H(x, y, p) = py' - L(x, y, y').$$

El problema dual busca maximizar el funcional dual asociado, sujeto a las ecuaciones de Hamilton:

$$\begin{aligned} \frac{dy}{dx} &= \frac{\partial H}{\partial p}, \\ \frac{dp}{dx} &= -\frac{\partial H}{\partial y}, \end{aligned}$$

con las condiciones de frontera correspondientes.

De este modo, el problema primal se centra en la función y que minimiza el tiempo y satisface la ecuación de Euler-Lagrange, mientras que el dual se enfoca en el momento p y la evolución conjunta (y, p) que satisfacen las ecuaciones de Hamilton. Ambos enfoques son equivalentes bajo condiciones adecuadas y permiten analizar el problema desde perspectivas complementarias.

Parte II.

Control Óptimo

En el cálculo variacional clásico, como se expuso previamente, el objetivo consiste en encontrar funciones $y : \Omega \subset \mathbb{R}^n \rightarrow \mathbb{R}^N$ que extremen funcionales del tipo

$$J[y] := \int_{\Omega} L(x, y(x), Dy(x)) dx.$$

Esta formulación resulta adecuada cuando se trabajan problemas donde las únicas restricciones se imponen a través de condiciones de frontera sobre $y(x)$, $x \in \Omega$.

Sin embargo, en diversos contextos (como en sistemas físicos, económicos o biológicos) las trayectorias posibles de un sistema no son completamente libres, sino que se encuentran sujetas a **restricciones dinámicas** que se modelan mediante ecuaciones diferenciales. En estos casos, la evolución del sistema está gobernada por una **dinámica de estado** la cual puede verse influida por variables externas denominadas **controles** que intervienen activamente en su comportamiento. (Pontryagin 2018)

Esta situación exige una generalización del enfoque clásico, dando lugar a una formulación más amplia conocida como **cálculo variacional con restricciones**, que establece las bases para la teoría del **control óptimo**.

Para propósitos de este trabajo, nos enfocaremos en el estudio del caso donde el dominio Ω , del funcional J es un intervalo compacto $[t_0, t_f] \subset \mathbb{R}$. Esta restricción permite formular de manera clara y rigurosa los principios fundamentales del cálculo variacional con restricciones y de la teoría del control óptimo, como el Hamiltoniano, el Principio del Máximo de Pontryagin y la estructura de los sistemas dinámicos manejados por controles.

Sin embargo, es importante señalar que formulaciones más generales que consideran dominios abiertos, intervalos infinitos o espacios abstractos de evolución, son posibles y se utilizan en diversas áreas como economía dinámica, teoría de juegos diferenciales y problemas con horizonte infinito. En estos casos, el análisis requiere condiciones adicionales de regularidad, crecimiento o convergencia, y se desarrolla dentro de marcos funcionales más generales. Para una exploración más amplia se puede consultar H. Sussmann (2005) o Evans (2022).

3. Del Cálculo Variacional a la Teoría del Control Óptimo

3.1. Trayectorias y Controles

Se considera un sistema dinámico cuyo estado en el tiempo está representado mediante un vector $x(t) \in X \subset \mathbb{R}^n$. Dicho vector puede incluir, según el contexto de cada problema, variables tales como posiciones, velocidades, temperaturas, entre otras. A este espacio X se le denomina *espacio fase* o *espacio de estados* del sistema.

Desde una perspectiva dinámica, la evolución del sistema se modela mediante la trayectoria de $x(\cdot)$ a lo largo del tiempo. Esta evolución no se considera completamente autónoma, sino que puede ser influenciada por *controladores* externos $u(t) \in \mathbb{R}^m$, los cuales representan acciones que afectan directamente a la dinámica del sistema. Estos controladores están restringidos a una *región de control* $U \subset \mathbb{R}^m$ que representa las configuraciones posibles de los controladores. Es común asumir que U es un conjunto cerrado y acotado, aunque en aplicaciones reales puede tener otras formas dependiendo de las limitaciones físicas o tecnológicas del sistema (Pontryagin 2018).

Definimos un *control admisible* como una función $u : [t_0, t_f] \rightarrow U$ que satisface las restricciones impuestas al sistema, tales como acotamiento, regularidad y cualquier condición adicional establecida por el problema. El conjunto de todos los controles admisibles se denota como

$$\mathcal{U}_{adm} := \{u : [t_0, t_f] \rightarrow U \subset \mathbb{R}^m \mid u \text{ cumple las restricciones del sistema}\}.$$

Dada una función de control $u(\cdot) \in \mathcal{U}_{adm}$, la evolución del sistema la modelamos mediante una *ecuación diferencial ordinaria* (ODE) que describe cómo cambia el estado en función del tiempo y del control aplicado:

$$\dot{x}(t) = a(x(t), u(t), t), \quad t \in [t_0, t_f], \quad (3.1)$$

donde $a : X \times U \times [t_0, t_f] \rightarrow \mathbb{R}^n$ es una función que describe la dinámica del sistema. La condición inicial se especifica como

$$x(t_0) = x_0,$$

con $x_0 \in X$ dado.

Para cada control admisible $u(\cdot)$, existe una trayectoria de estado $x(\cdot)$ que satisface tanto la dinámica del sistema como las condiciones iniciales. Definimos una **trayectoria-control admisible** como el par $(x(\cdot), u(\cdot))$ cuando la trayectoria $x(\cdot)$ junto con el control $u(\cdot) \in \mathcal{U}_{adm}$, cumple con la dinámica del sistema, las condiciones de frontera, así como todas las restricciones impuestas por el problema.

Ejemplo 3.1.

Trayectorias y Controles del Problema de Goddard

El problema de Goddard consiste en determinar el perfil óptimo de empuje de un cohete que asciende verticalmente desde la superficie terrestre, de modo que alcance la máxima altitud posible con un consumo eficiente de combustible. Este problema fue planteado por primera vez por R. H. Goddard en 1919, y posteriormente formulado como problema de control óptimo por Bellman en los años 50 (Bryson y Ho 1969).

El sistema de este problema se modela mediante un vector de estado tridimensional:

$$x(t) = \begin{pmatrix} h(t) \\ v(t) \\ m(t) \end{pmatrix} \in \mathbb{R}^3,$$

donde $h(t)$ representa la altura del cohete, $v(t)$ su velocidad y $m(t)$ la masa total en el instante t . El control lo definimos como

$$u(t) \in [0, 1],$$

el cual representa el empuje utilizado en cada instante, normalizado respecto a su valor máximo. Por tanto, la región de control es el intervalo cerrado $U = [0, 1] \subset \mathbb{R}$. La dinámica del sistema toma la forma:

$$\dot{x}(t) = \begin{pmatrix} \dot{h}(t) \\ \dot{v}(t) \\ \dot{m}(t) \end{pmatrix} = F(x(t), u(t), t) = F_0(x(t)) + u(t)F_1(x(t))$$

donde F_0 describe la dinámica autónoma del sistema (es decir, la evolución del estado en ausencia de control), y F_1 representa cómo el control $u(t)$ influye en la dinámica:

$$F_0(x) = \begin{pmatrix} v \\ -\frac{D}{m} - g(h) \\ 0 \end{pmatrix}, \quad F_1(x) = \begin{pmatrix} 0 \\ \frac{T_{\max}}{m} \\ -bT_{\max} \end{pmatrix}.$$

Así, la dinámica del problema de Goddard es

$$\dot{x}(t) = \begin{pmatrix} v \\ -\frac{D}{m} - g(h) \\ 0 \end{pmatrix} + u(t) \begin{pmatrix} 0 \\ \frac{T_{\max}}{m} \\ -bT_{\max} \end{pmatrix} = \begin{pmatrix} v \\ \frac{u(t)T_{\max}}{m} - \frac{D}{m} - g(h) \\ -u(t)bT_{\max} \end{pmatrix} \quad (3.2)$$

3.2. Formulación del Problema de control óptimo

A diferencia del cálculo variacional clásico, en el que las funciones admisibles eran elegidas libremente dentro de un espacio funcional, en control óptimo toda trayectoria admisible está determinada por un control admisible a través de la dinámica del sistema.

Para involucrar el parámetro de control $u(\cdot) \in \mathcal{U}_{adm}$ y formular un problema de optimización con restricciones para hallar un control óptimo $u^*(\cdot)$, definimos el funcional de costo $J[u(\cdot)]$ asociado a cada control admisible mediante la siguiente expresión:

$$J[u(\cdot)] = \int_{t_0}^{t_f} L(x(t), u(t), t) dt + \phi(x(t_f), t_f), \quad (3.3)$$

donde L representa el **costo instantáneo** (Lagrangiano) mientras que $\phi(x(t_f), t_f)$ el **costo terminal** de la función admisible $x(\cdot)$, además a $x(t_f)$ se le conoce como **restricción terminal**.

Definición 3.1. El problema general de optimización con restricciones de este nuevo funcional J , conocido en la literatura como el **Problema de control óptimo** en cálculo variacional, se plantea como:

$$\begin{aligned} \min_{u(\cdot) \in \mathcal{U}_{adm}} \quad & J[u(\cdot)] = \int_{t_0}^{t_f} L(x(t), u(t), t) dt + \phi(x(t_f), t_f) \\ \text{sujeto a} \quad & \end{aligned} \quad (3.4)$$

$$\begin{aligned} \dot{x}(t) &= a(x(t), u(t), t), \quad t \in [t_0, t_f] \\ x(t_0) &= x_0, \\ x(t_f) &\in \mathcal{X}_f \subset \mathbb{R}^n. \end{aligned}$$

donde f describe la dinámica del sistema y x_0 es la condición de estado inicial.

Observación 3.1. A este tipo de problemas se les conoce como problemas **tipo Bolza**, o colectivamente como el **problema de Bolza**, en honor a Oskar Bolza (1857-1942). Existen dos casos importantes de los problemas tipo Bolza. El primero es el **problema de Lagrange**, el cual se deriva del calculo de variaciones y en el que no existe un costo terminal (Poznyak 2008):

$$J[u(\cdot)] = \int_{t_0}^{t_f} L(x(t), u(t), t) dt. \quad (3.5)$$

El segundo caso especial es el **problema de Mayer**, en el cual no existe el costo instantáneo y por consiguiente depende exclusivamente del costo terminal (Poznyak 2008):

$$J[u(\cdot)] = \phi(x(t_f), t_f). \quad (3.6)$$

Ejemplo 3.2.

El Problema de Goddard

Retomando el problema de Goddard de Ejemplo 3.1, este puede formularse como un problema de control óptimo en el cual se busca maximizar la altitud final $h(t_f)$ de un cohete, sometido a una dinámica gobernada por las fuerzas de empuje, arrastre y gravedad, bajo restricciones físicas del sistema sobre la masa, la velocidad y el control (empuje). Así nuestro funcional objetivo tiene la forma

$$\max_{u(\cdot) \in \mathcal{U}_{adm}} J[u(\cdot)] = h(t_f),$$

sujeto a

$$\begin{aligned} \dot{x}(t) &= F(x(t), u(t), t) = F_0(x(t)) + u(t)F_1(x(t)), \\ x(t_0) &= \begin{pmatrix} h(t_0) \\ v(t_0) \\ m(t_0) \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \\ h(t) &\geq 1, \quad 0 \leq v(t) \leq 0.1 \\ m(t_f) &= 0.6, \\ u(t) &\in [0, 1]. \end{aligned}$$

3.3. Incremento del Funcional y Primera Variación en Control Óptimo

El concepto de primera variación fue desarrollado previamente en el contexto del cálculo variacional clásico (ver Sección 2.7). En el presente contexto de control óptimo, retomamos estos conceptos, adaptándolos al caso en que la función de estado satisface una dinámica diferencial que depende del control. Aquí la **primera variación** permite analizar cómo pequeñas perturbaciones en el control $u(\cdot)$ afectan el valor del funcional de costo $J[u(\cdot)]$. Esto es fundamental para derivar condiciones necesarias de optimalidad, como el principio de Pontryagin.

Consideremos un control admisible $u(\cdot) \in \mathcal{U}_{adm}$ que genera una trayectoria $x : [t_0, t_f] \rightarrow \mathbb{R}$, solución del sistema dinámico

$$\dot{x}(t) = a(x(t), u(t), t), \quad x(t_0) = x_0.$$

Trabajando con un funcional de costo del estilo Ecuación 3.3, introducimos una perturbación admisible $\delta u(\cdot)$, y definimos controles perturbados:

$$u_\varepsilon(t) = u(t) + \varepsilon \delta u(t), \quad \varepsilon \in \mathbb{R}. \quad (3.7)$$

De modo que genera una trayectoria $x_\varepsilon(\cdot) : [t_0, t_f] \rightarrow \mathbb{R}$, que satisface:

$$\dot{x}_\varepsilon(t) = a(x_\varepsilon(t), u_\varepsilon(t), t), \quad x_\varepsilon(t_0) = x_0.$$

El funcional de costo asociado al control $u_\varepsilon(\cdot)$ lo definimos como

$$J[u_\varepsilon(\cdot)] = \int_{t_0}^{t_f} L(x_\varepsilon(t), u_\varepsilon(t), t) dt + \phi(x_\varepsilon(t_f), t_f).$$

El incremento del funcional se define como la Ecuación 2.2:

$$\Delta J[u, \delta u] := J[u_\varepsilon(\cdot)] - J[u(\cdot)].$$

Definimos la **variación del estado**, denotada por $\delta x(t)$, como la derivada respecto a ε de la trayectoria perturbada evaluada en $\varepsilon = 0$:

$$\delta x(t) := \left. \frac{d}{d\varepsilon} x_\varepsilon(t) \right|_{\varepsilon=0}. \quad (3.8)$$

Esta variación describe cómo cambia la trayectoria de estado $x(t)$ ante una pequeña perturbación $\delta u(\cdot)$ en el control. Con esto podemos obtener la **ecuación variacional** que rige la evolución de $\delta x(t)$, derivando la ecuación diferencial para la variación del estado:

$$\dot{\delta x}(t) = f_x(x(t), u(t), t) \delta x(t) + f_u(x(t), u(t), t) \delta u(t), \quad \delta x(t_0) = 0. \quad (3.9)$$

Esta ecuación linealizada describe cómo las pequeñas variaciones en el control afectan la trayectoria del sistema a primer orden. De igual forma definimos la variación del tiempo final δt_f , la cual representa una perturbación en el instante final del intervalo de control. Introducimos el tiempo final perturbado como

$$t_f^\varepsilon := t_f + \varepsilon \delta t_f, \quad (3.10)$$

de modo que el nuevo intervalo de integración queda afectado por esta variación. Esta modificación junto con las perturbaciones del control δu y del estado δx contribuyen al cálculo de la primera variación del funcional.

Definimos la **primera variación** del funcional de costo de forma análoga a la formulación del Cálculo Variacional (Ecuación 2.5) como:

$$\delta J[u; \delta u] := \left. \frac{d}{d\varepsilon} J[u_\varepsilon(\cdot)] \right|_{\varepsilon=0}.$$

Desarrollando, obtenemos:

$$\begin{aligned} \delta J[u; \delta u] = & \int_{t_0}^{t_f} (\langle L_x(x(t), u(t), t), \delta x(t) \rangle + \langle L_u(x(t), u(t), t), \delta u(t) \rangle) dt \\ & + \langle \phi_x(x(t_f)), \delta x(t_f) \rangle + [\phi_t(x(t_f), t_f) + L(x(t_f), u(t_f), t_f)] \delta t_f. \end{aligned} \quad (3.11)$$

Ejemplo 3.3.

Primera variación en el problema de Goddard

Como mencionábamos en el Ejemplo 3.2, el funcional objetivo del problema de Goddard se define como

$$J[u(\cdot)] = h(t_f),$$

donde $h(t_f)$ es la altitud final alcanzada por el cohete.

En este caso, la función de costo terminal es $\phi(x(t_f)) = h(t_f)$ y no existe un término de costo instantáneo ($L \equiv 0$). Por lo tanto, la primera variación del funcional, siguiendo la fórmula general Ecuación 3.11, se reduce a:

$$\delta J[u; \delta u] = \langle \phi_x(x(t_f)), \delta x(t_f) \rangle = \langle (1, 0, 0), \delta x(t_f) \rangle = \delta h(t_f),$$

dado que $\phi_x(x) = (1, 0, 0)$ porque $\phi(x) = h$ corresponde a la primera componente del vector de estado.

Así, la primera variación mide cómo una pequeña perturbación en el control afecta la altitud final del cohete.

3.4. Ecuaciones de Euler-Lagrange Generalizadas

En las secciones anteriores, derivamos una expresión general para la primera variación del funcional en problemas de control óptimo. Esta formulación nos permite analizar cómo una perturbación en el control influye sobre la trayectoria del sistema.

A diferencia del cálculo variacional clásico, donde las variaciones de la función y pueden elegirse libremente dentro de un espacio funcional adecuado, en el contexto del control óptimo las variaciones del estado $\delta x(t)$ no son independientes, sino que se encuentran determinadas por una ecuación diferencial conocida como **ecuación variacional** (Ecuación 3.9). Esta dependencia impide anular directamente el integrando de la variación, como se hace en las ecuaciones de Euler-Lagrange (véase Sección 2.10).

Para resolver esta dificultad, recurrimos al uso de **multiplicadores de Lagrange** en espacios funcionales. Este enfoque permite incorporar las restricciones dinámicas en el análisis variacional y deducir condiciones necesarias de optimalidad en forma de un sistema diferencial que generaliza las ecuaciones de Euler-Lagrange.

Introducimos entonces los **multiplicadores de Lagrange**, conocidos en la literatura como **variables adjuntas** o **variable de coestado**, mediante una función auxiliar

$$\lambda : [t_0, t_f] \rightarrow \mathbb{R}^n,$$

que nos permite tratar las variaciones del estado $\delta x(t)$ como si fueran independientes, incorporando una condición adicional que la variable adjunta $\lambda(t)$ debe satisfacer.

Lo que queremos lograr con esto es construir un **funcional aumentado** en el que las restricciones del sistema se incorporan directamente al integrando (Kirk (2004)) a través del término

$$\langle \lambda(t), a(x(t), u(t), t) - \dot{x}(t) \rangle.$$

Este término se anula a lo largo de trayectorias admisibles, por lo que no modifica el valor del funcional original. Sin embargo, incluirlo resulta fundamental para derivar condiciones de optimalidad más prácticas desde un enfoque variacional.

El **funcional aumentado** toma entonces la forma:

$$J_a[u] := \phi(x(t_f), t_f) + \int_{t_0}^{t_f} [L(x(t), u(t), t) + \langle \lambda(t), a(x(t), u(t), t) - \dot{x}(t) \rangle] dt. \quad (3.12)$$

Dado que para toda trayectoria admisible se cumple $a(x(t), u(t), t) - \dot{x}(t) = 0$, entonces el funcional aumentado J_a coincide en valor con el funcional original J , es decir, $J_a = J$. Sin embargo, la variación de J_a resulta significativamente más útil para el análisis, debido a que permite incorporar explícitamente las restricciones dinámicas y facilita la deducción de condiciones necesarias de optimalidad mediante herramientas del cálculo variacional.

Consideramos perturbaciones admisibles del control (Ecuación 3.7), del estado (Ecuación 3.8) y del instante final (Ecuación 3.10), procedemos a calcular la primera variación del funcional aumentado J_a respecto a dichas perturbaciones:

$$\begin{aligned} \delta J_a(u; \delta u) &:= \left. \frac{d}{d\varepsilon} J_a[u_\varepsilon(\cdot)] \right|_{\varepsilon=0} \\ &= \frac{d}{d\varepsilon} \left(\phi(x_\varepsilon(t_f^\varepsilon), t_f^\varepsilon) \right. \\ &\quad \left. + \int_{t_0}^{t_f^\varepsilon} [L(x_\varepsilon(t), u_\varepsilon(t), t) + \langle \lambda(t), a(x_\varepsilon(t), u_\varepsilon(t), t) - \dot{x}_\varepsilon(t) \rangle] dt \right) \Big|_{\varepsilon=0}. \end{aligned} \quad (3.13)$$

Al derivar, primero calculamos la variación del costo terminal ϕ :

$$\left. \frac{d}{d\varepsilon} \phi(x_\varepsilon(t_f), t_f) \right|_{\varepsilon=0} = \langle \phi_x(x(t_f), t_f), \delta x(t_f) \rangle + \phi_t(x(t_f), t_f) \delta t_f. \quad (3.14)$$

Para la variación del término integral, aplicamos la regla de Leibniz para integrales con límites variables, obteniendo:

$$\begin{aligned} & \left. \frac{d}{d\varepsilon} \int_{t_0}^{t_f^\varepsilon} [L(x_\varepsilon, u_\varepsilon, t) + \langle \lambda(t), a(x_\varepsilon, u_\varepsilon, t) - \dot{x}_\varepsilon(t) \rangle] dt \right|_{\varepsilon=0} \\ &= \int_{t_0}^{t_f} \left[\left. \frac{d}{d\varepsilon} L(x_\varepsilon(t), u_\varepsilon(t), t) + \langle \lambda(t), a(x_\varepsilon(t), u_\varepsilon(t), t) - \dot{x}_\varepsilon(t) \rangle \right|_{\varepsilon=0} \right] dt \\ &+ [L(x(t_f), u(t_f), t_f) + \langle \lambda(t_f), a(x(t_f), u(t_f), t_f) - \dot{x}(t_f) \rangle] \delta t_f. \end{aligned} \quad (3.15)$$

Desarrollando la derivada dentro del integrando, la descomponemos en tres términos:

$$\left. \frac{d}{d\varepsilon} L(x_\varepsilon(t), u_\varepsilon(t), t) \right|_{\varepsilon=0} = \langle L_x, \delta x(t) \rangle + \langle L_u, \delta u(t) \rangle, \quad (3.16)$$

$$\left. \frac{d}{d\varepsilon} \langle \lambda(t), \dot{x}_\varepsilon(t) \rangle \right|_{\varepsilon=0} = -\langle \lambda(t), \delta \dot{x}(t) \rangle, \quad (3.17)$$

$$\left. \frac{d}{d\varepsilon} \langle \lambda(t), a(x_\varepsilon, u_\varepsilon, t) \rangle \right|_{\varepsilon=0} = \langle a_x^\top \lambda(t), \delta x(t) \rangle + \langle a_u^\top \lambda(t), \delta u(t) \rangle. \quad (3.18)$$

Dado que $\delta \dot{x}(t)$ no pertenece al espacio de variaciones admisibles, aplicamos integración por partes al término correspondiente para obtener:

$$-\int_{t_0}^{t_f} \langle \lambda(t), \delta \dot{x}(t) \rangle dt = -\langle \lambda(t_f), \delta x(t_f) \rangle + \int_{t_0}^{t_f} \langle \dot{\lambda}(t), \delta x(t) \rangle dt \quad (3.19)$$

Sustituyendo las ecuaciones Ecuación 3.16, Ecuación 3.19, Ecuación 3.18 en Ecuación 3.15, y agrupando con la expresión obtenida para la variación del costo terminal (Ecuación 3.14), la primera variación del funcional aumentado queda:

$$\begin{aligned} \delta J_a[u; \delta u] &= \langle \phi_x(x(t_f), t_f), \delta x(t_f) \rangle + \phi_t(x(t_f), t_f) \delta t_f - \langle \lambda(t_f), \delta x(t_f) \rangle \\ &+ \int_{t_0}^{t_f} \langle \dot{\lambda}(t), \delta x(t) \rangle dt + \int_{t_0}^{t_f} \langle L_x, \delta x(t) \rangle + \langle L_u, \delta u(t) \rangle dt \\ &+ \int_{t_0}^{t_f} \langle a_x^\top \lambda(t), \delta x(t) \rangle + \langle a_u^\top \lambda(t), \delta u(t) \rangle dt \\ &+ [L(x(t_f), u(t_f), t_f) + \langle \lambda(t_f), a(x(t_f), u(t_f), t_f) - \dot{x}(t_f) \rangle] \delta t_f. \end{aligned}$$

Reordenando términos, obtenemos:

$$\begin{aligned}
\delta J_a[u; \delta u] &= \langle \phi_x(x(t_f), t_f) - \lambda(t_f), \delta x(t_f) \rangle \\
&+ \int_{t_0}^{t_f} \langle L_x + a_x^\top \lambda(t) + \dot{\lambda}(t), \delta x(t) \rangle dt \\
&+ \int_{t_0}^{t_f} \langle L_u + a_u^\top \lambda(t), \delta u(t) \rangle dt \\
&+ [\phi_t(x(t_f), t_f) + L(x(t_f), u(t_f), t_f) + \langle \lambda(t_f), a(x(t_f), u(t_f), t_f) - \dot{x}(t_f) \rangle] \delta t_f.
\end{aligned} \tag{3.20}$$

Al aplicar el *Teorema Fundamental de Cálculo Variacional* (Teorema 2.2), la condición de nulidad de la primera variación debe cumplirse para toda perturbación admisible, implicando que cada término debe anularse de forma independiente.

Del término asociado a la variación del estado $\delta x(t)$ obtenemos

$$\dot{\lambda}(t) = -L_x(x(t), u(t), t) - a_x(x(t), u(t), t)^\top \lambda(t), \tag{3.21}$$

conocida como **ecuaciones adjuntas** o **ecuaciones de coestado**. Esta ecuación diferencial determina la evolución de las variables de coestado $\lambda(t)$ a lo largo del intervalo de control.

Por su parte, el término asociado a la variación del control $\delta u(t)$ conduce a una condición de estacionariedad del integrando respecto al control u , que se expresa como:

$$L_u(x(t), u(t), t) + a_u(x(t), u(t), t)^\top \lambda(t) = 0. \tag{3.22}$$

Esta condición establece que el control óptimo debe hacer que el integrando del funcional aumentado sea estacionario con respecto a variaciones del control. Es decir, en un punto óptimo, pequeños cambios admisibles en el control no mejoran el resultado en primer orden, ya que los efectos directos e indirectos de esa variación se equilibran.

La variación del estado en el tiempo final $\delta x(t_f)$ da lugar a la **condición de transversalidad**:

$$\lambda(t_f) = -\phi_x(x(t_f), t_f). \tag{3.23}$$

Esta condición relaciona el valor terminal de la variable adjunta con el gradiente espacial del costo terminal.

Finalmente, el término asociado a la variación del instante final δt_f impone la condición:

$$\phi_t(x(t_f), t_f) + L(x(t_f), u(t_f), t_f) + \langle \lambda(t_f), a(x(t_f), u(t_f), t_f) - \dot{x}(t_f) \rangle = 0.$$

Dado que para trayectorias admisibles se cumple $\dot{x}(t_f) = a(x(t_f), u(t_f))$, el último término se anula y la condición de frontera se reduce a:

$$\phi_t(x(t_f), t_f) + L(x(t_f), u(t_f), t_f) = 0 \quad (3.24)$$

En conjunto estas ecuaciones constituyen las **Ecuaciones de Euler-Lagrange generalizadas** para problemas de control óptimo, y representan un sistema de condiciones necesarias de primer orden que deben satisfacerse para que una trayectoria-control sea óptima.

Ejemplo 3.4.

Funcional Aumentado y Ecuaciones de Euler Generalizadas del Problema de Goddard

Retomando el problema de Goddard (Ejemplo 3.2), construimos el correspondiente **funcional aumentado** incorporando las restricciones dinámicas del sistema mediante variables adjuntas:

$$J_F[u] := \phi(x(t_f)) + \int_{t_0}^{t_f} \langle \lambda(t), F(x(t), u(t), t) - \dot{x}(t) \rangle dt,$$

donde $\lambda(t) = \begin{pmatrix} \lambda_1(t) \\ \lambda_2(t) \\ \lambda_3(t) \end{pmatrix}$ representa la variable adjunta asociada a cada componente del estado $x(t)$.

Aplicamos el procedimiento general desarrollado en esta sección para calcular la **primera variación del funcional aumentado**. El resultado se expresa como:

$$\begin{aligned} \delta J_F[u; \delta u] = & \langle \phi_x(x(t_f)) - \lambda(t_f), \delta x(t_f) \rangle \\ & + \int_{t_0}^{t_f} \langle \dot{\lambda}(t) + F_x(x(t), u(t), t)^\top \lambda(t), \delta x(t) \rangle dt \\ & + \int_{t_0}^{t_f} \langle F_u(x(t), u(t), t)^\top \lambda(t), \delta u(t) \rangle dt. \end{aligned}$$

Al imponer que esta variación se anule para toda variación admisible δu (según el Teorema 2.2), obtenemos la **ecuación adjunta**:

$$\dot{\lambda}(t) = -F_x(x(t), u(t), t)^\top \lambda(t),$$

la **condición de transversalidad**:

$$\lambda(t_f) = \phi_x(x(t_f)),$$

y la **condición de optimalidad respecto al control**:

$$F_u(x(t), u(t), t)^\top \lambda(t) = 0.$$

Estas tres expresiones conforman el sistema de ***Ecuaciones de Euler-Lagrange generalizadas*** para el problema de Goddard, y constituyen condiciones necesarias que debe satisfacer cualquier trayectoria-control óptimo en este modelo.

3.5. Hamiltoniano en Control Óptimo

Al considerar problemas de control óptimo con restricciones dinámicas (Definición 3.1) las condiciones necesarias de optimalidad no pueden derivarse como en el cálculo variacional clásico, por lo tanto en la Sección 3.4 recurrimos al método de multiplicadores de Lagrange en espacios funcionales, y construimos un *funcional aumentado* (Ecuación 3.12) en el que las restricciones dinámicas se incorporan explícitamente mediante una función auxiliar $\lambda(t)$. Este funcional coincide en valor con el original para controles admisibles, pero su estructura permite trabajar con variaciones independientes de $\delta x(t)$, y por tanto derivar condiciones de optimalidad más potentes.

En este contexto, nos es conveniente usar la función H , llamada ***Hamiltoniano***, definida como

$$H(x(t), u(t), \lambda(t), t) \triangleq \langle \lambda, a(x(t), u(t), t) \rangle + L(x(t), u(t), t). \quad (3.25)$$

Usando esta notación, podemos escribir las condiciones necesarias como sigue:

$$\left. \begin{aligned} \dot{x}^*(t) &= \frac{\partial H}{\partial \lambda}(x^*(t), u^*(t), \lambda^*(t), t), \\ \dot{\lambda}^*(t) &= -\frac{\partial H}{\partial x}(x^*(t), u^*(t), \lambda^*(t), t), \\ 0 &= \frac{\partial H}{\partial u}(x^*(t), u^*(t), \lambda^*(t), t) \end{aligned} \right\} \quad \forall t \in [t_0, t_f]. \quad (3.26)$$

3.6. Principio del Mínimo de Pontryagin (PMP)

Hasta este punto, no hemos impuesto restricciones explícitas sobre los controles admisibles ni sobre el estado del sistema. No obstante, en aplicaciones realistas tales restricciones aparecen con frecuencia. Los controles físicamente realizables suelen estar sujetos a limitaciones de magnitud: por ejemplo, el empuje de un cohete no puede superar un valor máximo determinado, y un motor solo puede ejercer un torque hasta cierto punto. Asimismo, las restricciones sobre el estado surgen comúnmente por razones de seguridad, consideraciones físicas o limitaciones estructurales del sistema. (Kirk 2004)

En esta sección analizamos cómo la presencia de restricciones sobre el control modifica las condiciones de optimalidad derivadas previamente. En particular, examinamos las

implicaciones de dichas restricciones sobre el teorema fundamental de la primera variación (Teorema 2.2). Esta generalización del teorema fundamental nos lleva al **Principio del Mínimo de Pontryagin**, una formulación clave en la teoría del control óptimo.

La variación del funcional aumentado J_a (Ecuación 3.20), en términos del **Hamiltoniano** (Ecuación 3.25) es

$$\begin{aligned}\delta J_a[u; \delta u] &= \langle \phi_x(x(t_f), t_f) - \lambda(t_f), \delta x(t_f) \rangle \\ &+ [H(x(t_f), u(t_f), \lambda(t_f), t_f) + \phi_t(x(t_f), t_f)] \delta t_f. \\ &+ \int_{t_0}^{t_f} \left\langle \frac{\partial H}{\partial x}(x(t), u(t), \lambda(t), t) + \dot{\lambda}(t), \delta x(t) \right\rangle \\ &+ \left\langle \frac{\partial H}{\partial u}(x(t), u(t), \lambda(t), t), \delta u(t) \right\rangle \\ &+ \left\langle \frac{\partial H}{\partial \lambda}(x(t), u(t), \lambda(t), t) - \dot{x}(t), \delta \lambda(t) \right\rangle dt.\end{aligned}$$

Ahora, consideremos una trayectoria tal que las ecuaciones de estado se satisfacen, y $\lambda(t)$ es elegido de modo que el coeficiente δx en la integral es idénticamente cero, y las condiciones de frontera se satisfacen, entonces nos queda

$$\delta J_a[u^*; \delta u] = \int_{t_0}^{t_f} \left\langle \frac{\partial H}{\partial u}(x^*(t), u^*(t), \lambda^*(t), t), \delta u(t) \right\rangle dt.$$

El integrando es la aproximación de primer orden en H causada por una perturbación en u , así

$$\left\langle \frac{\partial H}{\partial u}(x^*(t), u^*(t), \lambda^*(t), t), \delta u(t) \right\rangle = H(x^*(t), u^*(t) + \delta u(t), \lambda^*(t), t) - H(x^*(t), u^*(t), \lambda^*(t), t)$$

por lo tanto

$$\delta J_a[u^*; \delta u] = \int_{t_0}^{t_f} H(x^*(t), u^*(t) + \delta u(t), \lambda^*(t), t) - H(x^*(t), u^*(t), \lambda^*(t), t) dt.$$

Si el control $u^*(t)$ es óptimo, entonces la variación $\delta J_a[u^*; \delta u]$ debe ser no negativa para todas las perturbaciones admisibles. Es decir:

$$\delta J_a[u^*; \delta u] \geq 0.$$

Sustituyendo la expresión anterior obtenida para δJ_a , esta condición implica:

$$\int_{t_0}^{t_f} H(x^*(t), u^*(t) + \delta u(t), \lambda^*(t), t) - H(x^*(t), u^*(t), \lambda^*(t), t) dt \geq 0,$$

para toda δu admisible. En orden para que esto se cumpla es necesario que

$$H(x^*(t), u^*(t) + \delta u(t), \lambda^*(t), t) \geq H(x^*(t), u^*(t), \lambda^*(t), t), \quad (3.27)$$

para toda $\delta u(t)$ admisible y toda $t \in [t_0, t_f]$. Para probar esto consideremos el control

$$\begin{aligned} u(t) &= u^*(t); & t &\notin [t_1, t_2] \\ u(t) &= u^*(t) + \delta u(t); & t &\in [t_1, t_2], \end{aligned} \quad (3.28)$$

donde $[t_1, t_2]$ es un intervalo arbitrario de tiempo distinto de cero, y $\delta u(t)$ es una variación de control admisible.

Supongamos que la desigualdad (Ecuación 3.27) no se cumple para el control descrito en Ecuación 3.28, entonces en el intervalo $[t_1, t_2]$ tenemos

$$H(x^*(t), u(t), \lambda^*(t), t) < H(x^*(t), u^*(t), \lambda^*(t), t),$$

y por lo tanto,

$$\begin{aligned} &\int_{t_0}^{t_f} H(x^*(t), u(t), \lambda^*(t), t) - H(x^*(t), u^*(t), \lambda^*(t), t) dt \\ &= \int_{t_1}^{t_2} H(x^*(t), u(t), \lambda^*(t), t) - H(x^*(t), u^*(t), \lambda^*(t), t) dt < 0. \end{aligned}$$

Dado que el intervalo $[t_1, t_2]$ es arbitrario, puede estar en cualquier parte del intervalo $[t_0, t_f]$, y así es claro que si $H(x^*(t), u(t), \lambda^*(t), t) < H(x^*(t), u^*(t), \lambda^*(t), t)$ para todo $t \in [t_0, t_f]$, entonces siempre es posible construir un control admisible, que hace que $\delta J < 0$, lo cual es una contradicción a la optimalidad del control u^* . Por lo tanto, podemos concluir que una condición necesaria para que u^* minimice el funcional J es que

$$H(x^*(t), u(t), \lambda^*(t), t) \geq H(x^*(t), u^*(t), \lambda^*(t), t), \quad (3.29)$$

para todo $t \in [t_0, t_f]$ y para todo control admisible. Así, a la Ecuación 3.29, que indica que *un control óptimo debe minimizar el Hamiltoniano*, se le llama **Principio del Mínimo de Pontryagin (PMP)**.

En resumen, se busca control u^* que causa que un sistema (Ecuación 3.1) siga una trayectoria admisible y minimiza el funcional de costo (Ecuación 3.3). En términos del Hamiltoniano (Ecuación 3.25), las condiciones necesarias para que u^* sea un con control óptimo son

$$\left. \begin{aligned} \dot{x}^*(t) &= \frac{\partial H}{\partial \lambda}(x^*(t), u^*(t), \lambda^*(t), t), \\ \dot{\lambda}^*(t) &= -\frac{\partial H}{\partial x}(x^*(t), u^*(t), \lambda^*(t), t), \\ H(x^*(t), u(t), \lambda^*(t), t) &\geq H(x^*(t), u^*(t), \lambda^*(t), t) \end{aligned} \right\} \quad \forall t \in [t_0, t_f]. \quad (3.30)$$

Ejemplo 3.5.

Pontryagin Aplicado al Problema de Goddard

En el caso del problema de Goddard (Ejemplo 3.2), el funcional objetivo consiste en maximizar la altitud final $h(t_f)$, sujeto a la dinámica del sistema y condiciones adicionales. Dado el funcional aumentado (ver Ejemplo 3.4), el Hamiltoniano asociado a este problema toma la forma

$$H(h, v, m, \lambda_h, \lambda_v, \lambda_m, t) = \lambda_h v + \lambda_v \left(\frac{T_{\max}}{m} u - \frac{D}{m} - g(h) \right) - \lambda_m b T_{\max} u$$

y con esto, las condiciones necesarias que tiene que cumplir un control u^* para que sea solución del problema de Goddard son las **condiciones del estado** (dinámica):

$$\dot{h} = v, \quad \dot{v} = \frac{T_{\max}}{m} u - \frac{D}{m} - g(h), \quad \dot{m} = b T_{\max} u,$$

las **ecuaciones adjuntas** (ver Ecuación 3.21)

$$\begin{aligned} \dot{\lambda}_h &= -\frac{\partial H}{\partial h} = \lambda_v \left(\frac{1}{m} \frac{\partial D}{\partial h} + g'(h) \right), \\ \dot{\lambda}_v &= -\frac{\partial H}{\partial v} = -\lambda_h + \lambda_v \frac{1}{m} \frac{\partial D}{\partial v}, \\ \dot{\lambda}_m &= -\frac{\partial H}{\partial m} = \lambda_v \left(\frac{T_{\max}}{m^2} u - \frac{D}{m^2} \right), \end{aligned}$$

Condición sobre el control (Pontryagin):

$$u^*(t) \in \arg \max_{u \in [0,1]} H(x^*(t), u, \lambda^*(t), t).$$

Ya que el Hamiltoniano del problema es afín en el control, es decir, es lineal respecto al control, nos fijamos en la **función de conmutación** (*switching function*):

$$\varphi(t) := \frac{\partial H}{\partial u} = \lambda_v \frac{T_{\max}}{m} - \lambda_m b T_{\max},$$

así, el control óptimo sigue la regla *bang-bang*:

$$u^*(t) = \begin{cases} 1, & \varphi(t) > 0, \\ 0, & \varphi(t) < 0. \end{cases}$$

Si $\varphi(t) = 0$ en un intervalo, se habla de un *arco singular*.

4. Existencia de Soluciones en Problemas de Control Óptimo

Hasta este punto hemos presentado la formulación de problemas de control óptimo y hemos desarrollado condiciones necesarias de optimalidad, tales como el *Principio de Pontryagin* y las *Ecuaciones de Euler-Lagrange*. Sin embargo, estas herramientas se aplican bajo el supuesto de que el problema considerado admite, al menos, una solución óptima.

Con el fin de asegurar que los problemas de control óptimo poseen soluciones, es necesario recurrir a teoremas de existencia. Entre estos resultados encontramos el teorema de Filippov, presentado en el libro clásico de Cesari (Cesari (1983)), que establece condiciones suficientes para garantizar la existencia de un control y una trayectoria óptimos en problemas de tipo Lagrange y Bolza.

4.1. Teorema de Existencia de Filippov

Para poder enunciar el teorema de existencia de *Filippov*, adoptamos en esta sección la notación utilizada por Cesari, la cual resulta especialmente conveniente por su generalidad.

Sea $A \subset \mathbb{R}^{1+n}$ un conjunto al que denominamos **conjunto de pares tiempo-estado**. Denotamos por A_0 a la proyección de A sobre el eje temporal, además para cada instante $t \in A_0$, definimos el conjunto de estados

$$A(t) := \{x \in \mathbb{R}^n \mid (t, x) \in A\}.$$

Para cada par $(t, x) \in A$ consideramos un conjunto de **controles** $U(t, x) \subset \mathbb{R}^m$. Con ello introducimos el conjunto

$$M \subset \mathbb{R}^{1+n+m},$$

conformado por todos los tríos (t, x, u) tales que $(t, x) \in A$ y $u \in U(t, x)$.

Los datos del problema se construyen a partir de dos funciones definidas sobre M ; la función $f(t, x, u) = (f_1, \dots, f_n)$, que describe la dinámica del sistema en el espacio de estados, y $f_0(t, x, u)$, que representa el integrando del funcional de costo (o Lagrangiano).

En un problema tipo *Bolza* (Ecuación 3.3) intervienen dos componentes, el **costo terminal**, descrito por una función g , definida sobre un conjunto $B \subset \mathbb{R}^{2+2n}$, que reúne las condiciones de contorno admisibles en los extremos del intervalo de tiempo. La segunda componente es el **costo integral** que depende de la trayectoria y del control a través de la función f_0 .

Finalmente, para cada $(t, x) \in A$ introducimos los **conjuntos de valores alcanzables**. En primer lugar definimos el conjunto

$$Q(t, x) := \{f(t, x, u) \mid u \in U(t, x)\} \subset \mathbb{R}^n,$$

el cual representa el rango de la dinámica f cuando el control u varía sobre el conjunto de controles $U(t, x)$.

De manera análoga, definimos el conjunto

$$\tilde{Q}(t, x) := \{(z_0, z) \in \mathbb{R}^{1+n} \mid z_0 \geq f_0(t, x, u), z = f(t, x, u), u \in U(t, x)\},$$

que reúne todos los pares (z_0, z) , con $z_0 \geq f_0(t, x, u)$ y $z = f(t, x, u)$ para algún $u \in U(t, x)$.

Estos conjuntos serán fundamentales para enunciar las condiciones de existencia del teorema de Filippov, ya que permiten analizar las propiedades de convexidad y compacidad necesarias para garantizar la existencia de soluciones óptimas.

Consideramos el problema de minimización del funcional

$$I[x, u] = g(t_1, x(t_1), t_2, x(t_2)) + \int_{t_1}^{t_2} f_0(t, x(t), u(t)) dt, \quad (4.1)$$

sujeto a

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)) & t \in [t_1, t_2], \\ (t, x(t)) &\in A, \quad u(t) \in U(t, x(t)) & t \in [t_1, t_2], \\ (t_1, x(t_1), t_2, x(t_2)) &\in B, \\ f_0(\cdot, x(\cdot), u(\cdot)) &L\text{-integrable en } [t_1, t_2]. \end{aligned} \quad (4.2)$$

Teorema 4.1. *Supongamos que $A \subset \mathbb{R}^{1+n}$ es compacto, que $M \subset \mathbb{R}^{1+n+m}$ es compacto, que $B \subset \mathbb{R}^{2+2n}$ es cerrado, que la función g es semicontinua inferiormente en B , que las funciones f_0 y f son continuas en M , y que para casi todo t los conjuntos $\tilde{Q}(t, x)$ son convexos para todo $x \in A(t)$. Entonces el funcional I (Ecuación 4.1) alcanza un mínimo absoluto en la clase no vacía Ω de todas las parejas admisibles (x, u) que satisfacen las condiciones del problema.*

Las hipótesis de este teorema no constituyen meras restricciones formales, sino condiciones naturales que reflejan la estructura de los problemas de control óptimo. La compacidad del conjunto A garantiza que los instantes de tiempo y los estados permanezcan dentro de un dominio acotado, evitando así trayectorias que se escapen al infinito y asegura que el problema posea sentido práctico. De manera análoga, la compacidad del conjunto M garantiza que el espacio de posibles combinaciones de tiempo, estado y control está acotado y cerrado, lo que impide la aparición de controles o trayectorias no admisibles que escapen a valores extremos o indefinidos.

El conjunto B , debe ser cerrado para evitar inconsistencias en los extremos del intervalo. Si B no fuera cerrado, sería posible aproximarse indefinidamente a condiciones de contorno factibles sin llegar a cumplirlas, lo que impediría la existencia de soluciones viables. La semicontinuidad inferior de g garantiza que el costo final no presente descensos abruptos en el límite.

La continuidad de las funciones f_0 y f en M asegura que la dinámica y el integrando del costo respondan de manera estable a variaciones en los controles y en los estados.

Bajo este marco, el teorema de Filippov establece que el problema de Bolza siempre admite una solución óptima, consolidando así la base teórica de la teoría del control óptimo.

Prueba. Para cada par admisible $(x, u) \in \Omega$, $(t_1, x(t_1), t_2, x(t_2))$ se encuentra en el conjunto compacto $B \cap (A \times A)$. La función g tiene un mínimo m en el compacto $B \cap (A \times A)$. Dado que M es compacto y f_0, f son continuas en M , existe una constante N tal que $|f_0(t, x, u)|, |f(t, x, u)| \leq N$ en M , y podemos tomar a N tal que también tengamos $|t|, |x|, |u| \leq N$ para $(t, x, u) \in M$. Así, el conjunto A_0 , está contenido en el intervalo $[-N, N]$. Denominemos D al diámetro de A_0 .

Note que, si $(t, x) \in A$ y $z \in Q(t, x) = f(t, x, U(t, x)) \subset \mathbb{R}^n$, entonces $(z_0, z) \in \tilde{Q}(t, x)$ implica que $z_0 \geq -N$, y $(N, z) \in \tilde{Q}(t, x)$.

De (Proposición 1.2), reemplazando x por (t, x) , obtenemos que los conjuntos $\tilde{Q}(t, x)$ son cerrados en $\mathbb{R}^{n+1}$ con proyección compacta $Q(t, x)$, y que los mismos conjuntos $\tilde{Q}(t, x)$ son semicontinuos superiormente por inclusión de conjuntos y por tanto, tienen la propiedad (K) en A . Más aún, para casi toda $\bar{t}$ y toda $\bar{x} \in A(\bar{t})$, el conjunto $\tilde{Q}(\bar{t}, \bar{x})$ es convexo, y los mismos conjuntos $\tilde{Q}(t, x)$ tienen la propiedad (Q) en $(\bar{t}, \bar{x})$.

Consideremos el siguiente problema de minimización: Determinar un par $(x(t), \eta(t))$, $t_1 \leq t \leq t_2$, x Absolutamente Continua (AC), η L-integrable, para la cual el funcional

$$J[x, \eta] = g(t_1, x(t_1), t_2, x(t_2)) + \int_{t_1}^{t_2} \eta(t) dt, \quad (4.3)$$

tiene su valor mínimo bajo las consideraciones

$$\begin{aligned} (t, x(t)) \in A, \quad (\eta(t), \dot{x}(t)) \in \tilde{Q}(t, x(t)), \quad t \in [t_1, t_2] \text{ (a.e.)}, \\ (t_1, x(t_1), t_2, x(t_2)) \in B. \end{aligned} \quad (4.4)$$

Sea Ω' la clase de todos los pares $x(t), \eta(t)$, $t_1 \leq t \leq t_2$, para los cuales todas las condiciones anteriores se cumplen. Así, sabemos que para $(x, \eta) \in \Omega'$ tenemos que $\eta(t) \geq -N$ para toda t , más aún $(x, \eta_0) \in \Omega'$ para $\eta_0(t) \equiv N$.

Finalmente, dado que $\eta(t) \geq f_0(t, x(t), u(t))$, $t_2 - t_1 \leq D$, $g \geq m$ en $B \cap (A \times A)$, y $f_0 \geq -N$ en M , tenemos, de (Ecuación 4.2) y (Ecuación 4.3), $J[x, \eta] \geq I[x, u] \geq m - DN$, y también $J = I$ siempre que $\eta(t) = f_0(t, x(t), u(t))$, $t \in [t_1, t_2]$ (a.e.). Si $j = \inf_{\Omega'} J[x, \eta]$, y $i = \inf_{\Omega'} I[x, u]$, entonces $j = i$ y ambos son finitos.

Sea $x_k(t), \eta_k(t)$, $t_{1_k} \leq t \leq t_{2_k}$, $k = 1, 2, \dots$, una sucesión minimizante para J , esto es, $(x_k, \eta_k) \in \Omega'$ y $J[x_k, \eta_k] \rightarrow j$ cuando $k \rightarrow \infty$. Así, $(t, x_k(t)) \in A$, $(\eta_k(t), \dot{x}_k(t)) \in \tilde{Q}(t, x_k(t))$, $\dot{x}_k(t) \in Q(t, x_k(t))$, $t \in [t_{1_k}, t_{2_k}]$ (a.e.), $(t_{1_k}, x_k(t_{1_k}), t_{2_k}, x_k(t_{2_k})) \in B$ para todo k , y

$$g(t_{1_k}, x_k(t_{1_k}), t_{2_k}, x_k(t_{2_k})) + \int_{t_{1_k}}^{t_{2_k}} \eta_k(t) dt \rightarrow j \text{ cuando } k \rightarrow \infty.$$

Dado que ambas partes en esta última expresión se encuentran acotadas inferiormente y j es finito, ambas partes están acotadas. Lo anterior permite tomar una subsucesión, que por la facilidad y simplicidad seguiremos denotando con $[k]$, tal que $g(t_{1_k}, x_k(t_{1_k}), t_{2_k}, x_k(t_{2_k})) \rightarrow j'$, $\int_{t_{1_k}}^{t_{2_k}} \eta_k(t) dt \rightarrow j''$, ambas j' y j'' finitas con $j' + j'' = j$. Dado que $Q(t, x_k(t)) = f(t, x_k(t), U(t, x_k(t)))$, tenemos que $|\dot{x}_k(t)| \leq N$ para toda t y k ; por lo tanto, las funciones $x_k(t)$ son equi-Lipschitzianas (ver Definición 1.14), y por tanto equicontinuas (ver Definición 1.15). Dado que $(t, x_k(t)) \in A$, tenemos que $-N \leq t_{1_k} < t_{2_k} \leq N$, $|x_k(t)| \leq N$ para toda t y k .

Por el teorema de Arzelà-Ascoli (ver Teorema 1.1), existe una subsucesión $[k_\ell]$ tal que $t_{1_{k_\ell}} \rightarrow t_1$, $t_{2_{k_\ell}} \rightarrow t_2$ cuando $\ell \rightarrow \infty$, y existe una función continua $x(t)$, $t_1 \leq t \leq t_2$, tal que $x_{k_\ell}(t) \rightarrow x(t)$ en la métrica ρ . Dado que A y B son cerrados, $(t, x(t)) \in A$ y $(t_1, x(t_1), t_2, x(t_2)) \in B$. Como g es semicontinua inferiormente, tenemos también que $g(t_1, x(t_1), t_2, x(t_2)) \leq j'$. Ya que las funciones x_k son equi-Lipschitzianas, x también es Lipschitziana (ver Definición 1.14) y por lo tanto es absolutamente continua. Por el teorema de la cerradura inferior (ver @), con $\phi(t) = -N$, existe una función L -integrable $\eta(t)$, $t_1 \leq t \leq t_2$, tal que $(\eta(t), \dot{x}(t)) \in \tilde{Q}(t, x(t))$, $t \in [t_1, t_2]$ (a.e.), y $\int_{t_1}^{t_2} \eta(t) dt \leq j''$. Por lo que

$$J[x, \eta] = g(t_1, x(t_1), t_2, x(t_2)) + \int_{t_1}^{t_2} \eta(t) dt \leq j' + j'' = j.$$

Pero $(x, \eta) \in \Omega'$, así que $J[x, \eta] \geq j$, y esto demuestra que $J[x, \eta] = j$.

Notemos que $\tilde{Q}(t, x)$ es la imagen continua de $U(t, x) \times [0, +\infty)$ bajo el mapeo $(u, v) \rightarrow [f(t, x, u), f_0(t, x, u) + v]$, y $(\eta(t), \dot{x}(t)) \in \tilde{Q}(t, x(t))$. También sabemos que los conjuntos $\tilde{Q}(t, x)$ tiene la propiedad (K), y más aún, para casi toda $\bar{t} \in [t_1, t_2]$, los conjuntos $\tilde{Q}(t, x)$ tienen la propiedad (Q) con respecto a (t, x) en $(\bar{t}, x(\bar{t}))$. Por el teorema de la función implícita (ver Teorema 1.2), deducimos que existen funciones medibles $u(t)$, $v(t)$, $t_1 \leq t \leq t_2$, $u(t) \in U(t, x(t))$, $v(t) \geq 0$, tales que $\dot{x}(t) = f(t, x(t), u(t))$, $\eta(t) = f_0(t, x(t), u(t)) + v(t)$. Debido a que j es mínimo, debemos tener que $v(t) = 0$ a.e. en $[t_1, t_2]$. Así

$$\begin{aligned} J[x, \eta] &= I[x, u] \\ &= g(t_1, x(t_1), t_2, x(t_2)) + \int_{t_1}^{t_2} f_0(t, x(t), u(t)) dt = j = i. \end{aligned}$$

□

Parte III.

Caso de Estudio

Problema de Goddard

El problema de maximizar la altitud final de un cohete en ascenso vertical constituye un desafío fundamental en el campo de la ingeniería aeroespacial, así como en la teoría de control óptimo. Goddard (1919) planteó este problema por primera vez, introduciendo las formulaciones matemáticas que resultaron esenciales para optimizar las trayectorias de cohetes y promover los avances en la exploración espacial.

Posteriormente, numerosos autores abordaron este problema con distintos métodos matemáticos, como Hamel (1927), el cual propuso la existencia de una solución mediante el Cálculo Variacional. Con el tiempo, el problema de Goddard ha evolucionado para abarcar escenarios más complejos, incorporando restricciones adicionales y límites finales.

En este trabajo, se analiza el problema considerando diversos factores que afectan el rendimiento del cohete, tales como el empuje, la pérdida de masa debido al consumo de combustible y las fuerzas externas que actúan sobre su trayectoria, entre las cuales destacan el límite de la presión dinámica y la gravedad.



Figura 4.1.: Cohete

5. Problema Original

5.1. Formulación

El problema se enfoca en maximizar la altitud final de un cohete en ascenso vertical, considerando los efectos del arrastre atmosférico y el inverso cuadrado del campo gravitacional, tomando el empuje (T) como el único control. Los parámetros empleados corresponden a los presentados por Seywald y Cliff (1993), mientras que los datos relativos al vehículo y su aerodinámica provienen de Zlatskii y Kiforenko (1975). Dichos datos describen las características del misil *Soviet SA-2 surface-to-air missile*, **NATO**.

A continuación, se presenta la formulación matemática del problema (vista como un problema de Cálculo Variacional), basada en los parámetros y datos previamente descritos.

$$\max h(t_f) \quad (5.1)$$

sujeto a

$$\begin{aligned} \dot{h} &= v \\ \dot{v} &= \frac{T - D}{m} - \frac{1}{h^2} \\ \dot{m} &= -\frac{T}{c} \end{aligned} \quad (5.2)$$

el control (Thrust)

$$T \in [0, T_{max}] \quad (5.3)$$

y las condiciones de frontera

$$\begin{aligned} r(0) &= 1 \\ v(0) &= 0 \\ m(0) &= 1 \\ m(t_f) &= m_f \end{aligned} \quad (5.4)$$

5.2. Resultados Numéricos

La existencia de la solución es clara debido a las condiciones del problema y lo discutido en Capítulo 4. Resolver el problema de Goddard implica determinar la trayectoria óptima $\mathbf{x}^*(t)$ y el control $u^*(t)$ que maximicen la altitud final $h(t_f)$, cumpliendo con las ecuaciones de dinámica, las restricciones de frontera, y las limitaciones en el control. Esto se puede lograr mediante dos enfoques principales: el método directo y el método indirecto. En este trabajo, implementamos ambos enfoques utilizando el lenguaje de programación Julia y la paquetería OptimalControl.jl (s. f.), presentando las soluciones y resultados obtenidos.

5.2.1. Método Directo

El método directo consiste en discretizar las ecuaciones de estado y de control, convirtiendo el problema de un problema de Calculo Variacional en un problema de programación no lineal (PNL). Se utiliza IPOPT para resolver el NLP, con MUMPS como el solver lineal para hacer más eficiente el manejo de grandes sistemas dispersos.

Las paqueterías que se usarán son las siguientes

```
using OptimalControl
using NLPModelsIpopt
using DataFrames
```

Definimos los datos del problema (las restricciones, condiciones de frontera y la dinámica)

```
t0 = 0          # Tiempo Inicial
h0 = 1          # Altitud Inicial
v0 = 0          # Velocidad Inicial
m0 = 1          # Masa Inicial
vmax = 0.1      # Velocidad Máxima Autorizada
mf = 0.6        # Masa Final Objetivo
const Cd = 310  # Coeficiente de Arrastre
const Tmax = 3.5 # Empuje Máximo
const rho = 500 # Tasa de caída de Densidad
const b = 2     # Constante positiva de acuerdo al motor
```

Definimos el problema de control optimo

```
@def ocp begin
    tf R, variable
```

```

t    [ t0, tf ], time
x = (h, v, m)   R³, state
u    R, control

x(t0) == [ h0, v0, m0 ]
m(tf) == mf,          (1)
0  u(t)    1
h(t)    h0
0  v(t)    vmax

ẋ(t) == F0(x(t)) + u(t) * F1(x(t))

h(tf) → max
end;

F0(x) = begin
    h, v, m = x
    D = Cd * v^2 * exp(-(h - 1)) # Drag
    return [ v, -D/m - 1/h^2, 0 ]
end

F1(x) = begin
    h, v, m = x
    return [ 0, Tmax/m, -b*Tmax ]
end

```

Se reuelve utilizando IPOPT (Métodos de Punto Interior)

```
direct_sol = solve(ocp; grid_size=100)
```

This is Ipopt version 3.14.17, running with linear solver MUMPS 5.7.3.

Number of nonzeros in equality constraint Jacobian...	2104
Number of nonzeros in inequality constraint Jacobian..	0
Number of nonzeros in Lagrangian Hessian.....	1111
 Total number of variables.....	 405
variables with only lower bounds:	101
variables with lower and upper bounds:	202
variables with only upper bounds:	0
Total number of equality constraints.....	304
Total number of inequality constraints.....	0

inequality constraints with only lower bounds:	0
inequality constraints with lower and upper bounds:	0
inequality constraints with only upper bounds:	0

iter	objective	inf_pr	inf_du	lg(mu)	d	lg(rg)	alpha_du	alpha_pr	ls
0	-1.0100000e+00	9.00e-01	2.00e+00	0.0	0.00e+00	-	0.00e+00	0.00e+00	0
1	-1.0090670e+00	8.99e-01	6.67e+01	1.3	1.67e+02	-	3.64e-03	5.93e-04f	1
2	-1.0000907e+00	8.74e-01	1.83e+02	1.0	6.64e+00	-	3.63e-02	2.83e-02h	1
3	-1.0023670e+00	8.37e-01	1.34e+04	1.0	6.91e+00	-	2.11e-01	4.19e-02f	1
4	-1.0025025e+00	7.70e-01	9.45e+03	1.5	4.04e+00	-	1.00e+00	8.09e-02f	1
5	-1.0033626e+00	7.16e-01	1.48e+05	2.3	3.56e+00	-	3.73e-01	6.94e-02f	1
6	-1.0142503e+00	9.62e-03	3.99e+04	2.3	7.16e-01	-	4.49e-01	9.90e-01h	1
7	-1.0101264e+00	4.21e-03	4.24e+05	1.8	5.32e-01	-	5.03e-01	9.90e-01h	1
8	-1.0068427e+00	3.20e-04	2.87e+06	0.9	2.44e-01	-	6.73e-01	9.91e-01h	1
9	-1.0067336e+00	2.64e-06	2.30e+07	0.1	7.39e-02	-	7.07e-01	1.00e+00f	1
iter	objective	inf_pr	inf_du	lg(mu)	d	lg(rg)	alpha_du	alpha_pr	ls
10	-1.0067340e+00	1.13e-10	6.50e+05	-5.0	2.90e-04	-	9.89e-01	1.00e+00h	1
11	-1.0067350e+00	2.81e-10	7.20e+03	-7.0	4.26e-04	-	9.89e-01	1.00e+00h	1
12	-1.0078967e+00	9.07e-04	5.95e+03	-3.0	7.61e-01	-	6.55e-01	7.21e-01f	1
13	-1.0081866e+00	3.67e-06	9.38e+03	-9.0	1.31e-02	-	8.60e-01	1.00e+00h	1
14	-1.0091814e+00	1.57e-04	1.79e+02	-4.6	2.18e-01	-	1.00e+00	9.28e-01h	1
15	-1.0105115e+00	2.55e-04	7.17e+02	-4.4	3.09e-01	-	1.00e+00	5.75e-01h	1
16	-1.0114149e+00	2.34e-05	7.98e-04	-5.1	4.29e-02	-	1.00e+00	1.00e+00h	1
17	-1.0122891e+00	7.90e-05	1.24e+02	-5.8	1.25e-01	-	9.98e-01	7.92e-01h	1
18	-1.0125091e+00	2.83e-05	4.85e-04	-6.1	1.06e-01	-	1.00e+00	1.00e+00h	1
19	-1.0125585e+00	1.21e-05	8.67e-05	-6.8	1.01e-01	-	1.00e+00	1.00e+00h	1
iter	objective	inf_pr	inf_du	lg(mu)	d	lg(rg)	alpha_du	alpha_pr	ls
20	-1.0125675e+00	4.40e-06	3.01e-05	-7.3	9.80e-02	-	1.00e+00	1.00e+00h	1
21	-1.0125707e+00	1.75e-06	3.70e-02	-8.0	1.14e-01	-	1.00e+00	9.91e-01h	1
22	-1.0125714e+00	9.50e-07	4.98e-03	-8.6	1.37e-01	-	1.00e+00	9.94e-01h	1

Number of Iterations.....: 22

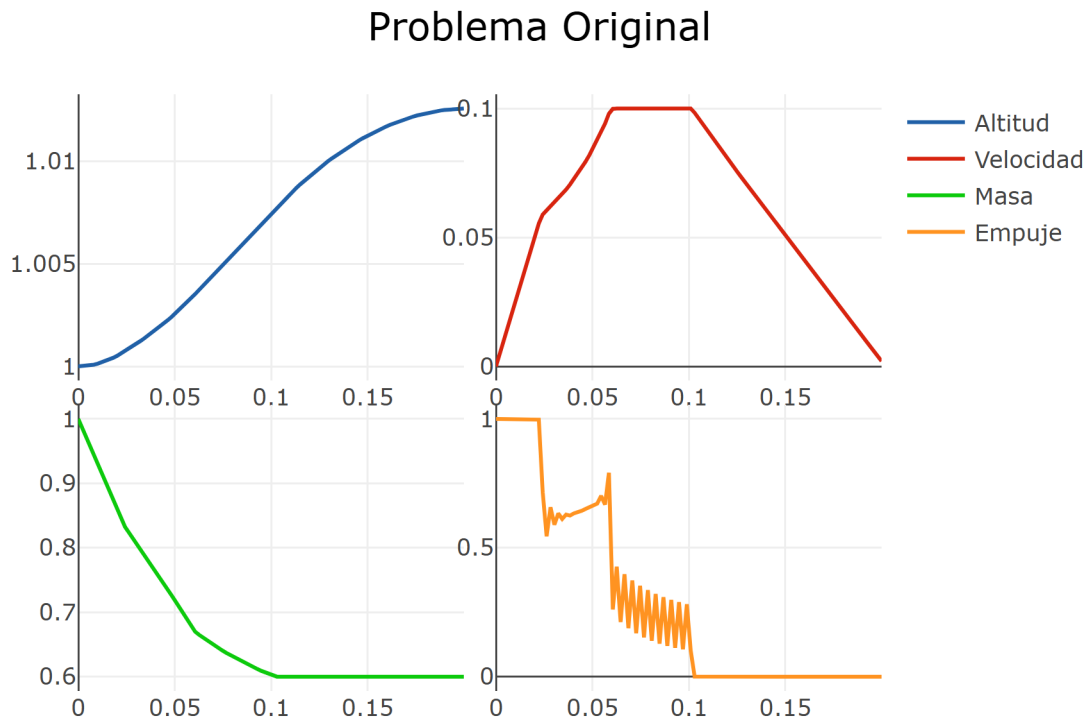
	(scaled)	(unscaled)
Objective.....:	-1.0125714428815755e+00	-1.0125714428815755e+00
Dual infeasibility.....:	4.9826923912073525e-03	4.9826923912073525e-03
Constraint violation.....:	3.5941849951814930e-07	9.4995894299454420e-07
Variable bound violation:	9.9621924876114321e-09	9.9621924876114321e-09
Complementarity.....:	8.1796266157923928e-09	8.1796266157923928e-09
Overall NLP error.....:	5.1484604087044977e-07	4.9826923912073525e-03

Number of objective function evaluations	= 23
Number of objective gradient evaluations	= 23

Number of equality constraint evaluations	= 23
Number of inequality constraint evaluations	= 0
Number of equality constraint Jacobian evaluations	= 23
Number of inequality constraint Jacobian evaluations	= 0
Number of Lagrangian Hessian evaluations	= 22
Total seconds in IPOPT	= 4.288

EXIT: Optimal Solution Found.

OptimalControlSolution



5.2.2. Método Indirecto

El método indirecto se basa en la teoría del mínimo de Pontryagin y requiere resolver un sistema de ecuaciones diferenciales en términos de las variables de estado, de coestado y las condiciones de transversalidad. Para verificar que las condiciones de frontera y las restricciones del problema se cumplen, se define una función de disparo. Se resuelve un sistema no lineal para encontrar los valores óptimos de los parámetros usando métodos numéricos. Finalmente, se construye la trayectoria óptima mediante la concatenación de flujos de los diferentes regímenes de control.

```

t = time_grid(direct_sol)
x = state(direct_sol)
u = control(direct_sol)
p = costate(direct_sol)

H1 = Lift(F1)           #  $H_1(x, p) = p' * F_1(x)$ 
(t) = H1(x(t), p(t))   # Función "switching"
g(x) = vmax - x[2]      # Restricción del estado v    vmax

# Controles

u0 = 0                  # Control Apagado
u1 = 1                  # Control "Bang"

H0 = Lift(F0)           #  $H_0(x, p) = p' * F_0(x)$ 
H01 = @Lie { H0, H1 }
H001 = @Lie { H0, H01 }
H101 = @Lie { H1, H01 }
us(x, p) = -H001(x, p) / H101(x, p)   # Control Singular

ub(x) = -(F0 g)(x) / (F1 g)(x)         # Frontera del Control
(x, p) = H01(x, p) / (F1 g)(x)         # Multiplicador asociado a
                                        # la restricción de estado g

# Flujo

f0 = Flow(ocp, (x, p, tf) -> u0)
f1 = Flow(ocp, (x, p, tf) -> u1)
fs = Flow(ocp, (x, p, tf) -> us(x, p))
fb = Flow(ocp, (x, p, tf) -> ub(x), (x, u, tf) -> g(x), (x, p, tf)
    ↪ -> (x, p))

x0 = [ r0, v0, m0 ] # Estado Inicial

#Función de Disparo

function shoot!(s, p0, t1, t2, t3, tf)

    x1, p1 = f1(t0, x0, p0, t1)
    x2, p2 = fs(t1, x1, p1, t2)
    x3, p3 = fb(t2, x2, p2, t3)
    xf, pf = f0(t3, x3, p3, tf)

```

```

    s[1] = constraint(ocp, :eq1)(x0, xf, tf) - mf # Restricción de
↪ Masa Final
    s[2:3] = pf[1:2] - [ 1, 0 ]                # Condiciones de
↪ Transversalidad
    s[4] = H1(x1, p1)                          # H1 = H01 = 0
    s[5] = H01(x1, p1)                        # Entrada del
↪ Arco Singular
    s[6] = g(x2)                              # g = 0 al entrar
↪ en el Arco Limite
    s[7] = H0(xf, pf)                         # Ya que el
↪ Tiempo Final es libre

end

# Suposición Inicial

    = 1e-3
t13 = t[ abs.(.(t)) . ]
t23 = t[ 0 . (g x).(t) . ]
p0 = p(t0)
t1 = min(t13...)
t2 = min(t23...)
t3 = max(t23...)
tf = t[end]

println("p0 = ", p0)
println("t1 = ", t1)
println("t2 = ", t2)
println("t3 = ", t3)
println("tf = ", tf)

# Norma de la función de disparo en la solución
using LinearAlgebra: norm
s = similar(p0, 7)
shoot!(s, p0, t1, t2, t3, tf)
println("\nNorma de la función de disparo: ||s|| = ", norm(s), "\n")

    = [ p0 ; t1 ; t2 ; t3 ; tf ] # Suposición Inicial

# Función auxiliar con entradas agregadas
nle! = (s, , ) -> shoot!(s, [1:3], [4], [5], [6], [7])

```

```

# Problema NLE con Suposición Inicial
prob = NonlinearProblem(nle!, )

# Resolución de  $S() = 0$ 
indirect_sol = solve(prob; abstol=1e-8, reltol=1e-8,
    ↪ show_trace=Val(true))

# Recuperamos la solución de coestado junto con los tiempos
p0 = indirect_sol.u[1:3]
t1 = indirect_sol.u[4]
t2 = indirect_sol.u[5]
t3 = indirect_sol.u[6]
tf = indirect_sol.u[7]

println("")
println("p0 = ", p0)
println("t1 = ", t1)
println("t2 = ", t2)
println("t3 = ", t3)
println("tf = ", tf)

# Norma de la función de disparo en la solución
s = similar(p0, 7)
shoot!(s, p0, t1, t2, t3, tf)
println("\nNorma de la función de disparo: ||s|| = ", norm(s), "\n")

f = f1 * (t1, fs) * (t2, fb) * (t3, f0) # concatenación de los
    ↪ flujos
flow_sol = f((t0, tf), x0, p0)      # Calcular la solución:
                                    # estado, coestado, control...

```

A continuación se presenta una gráfica comparando el método directo (Sección 5.2.1) y el método indirecto (Sección 5.2.2)

5.2.3. Análisis de Sensibilidad

Se realiza un análisis de sensibilidad del problema variando el parámetro de la masa final en un intervalo $[0.45, 0.65]$ tomando valores con distancia de 0.05. Este estudio permite determinar el límite máximo de altitud en función de diferentes valores de masa final, evaluando como afectan las restricciones dinámicas y el consumo de combustible.

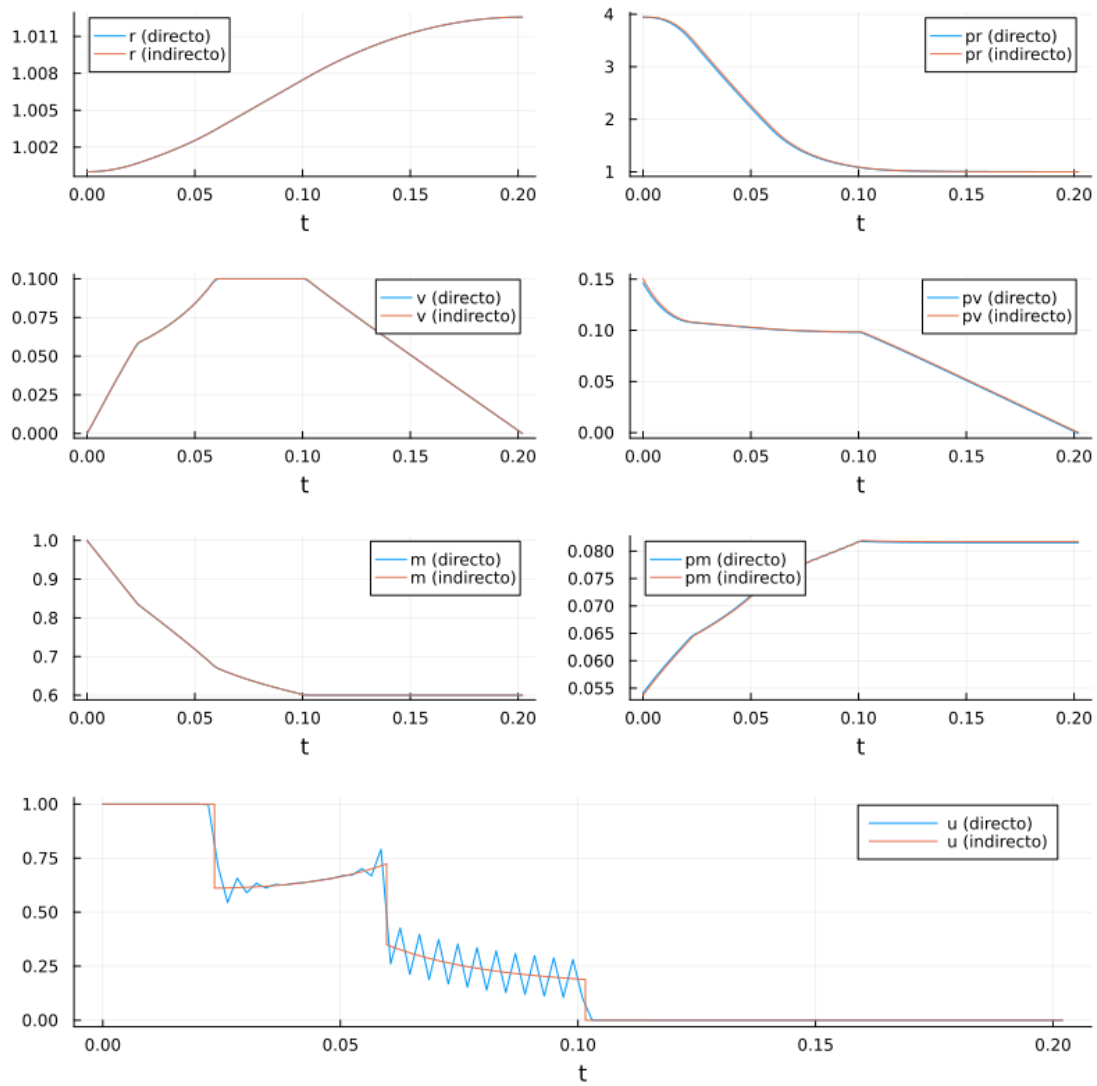


Figura 5.1.: Comparación entre ambos métodos

A través de simulaciones numéricas (haciendo uso del Sección 5.2.1), se puede examinar la relación entre la eficiencia del control y la altitud final alcanzada.

```

using OptimalControl
using NLPModelsIpopt
using DataFrames

t0 = 0      # initial time
h0 = 1      # initial altitude
v0 = 0      # initial speed
m0 = 1      # initial mass
mass = [0.45, 0.5, 0.55, 0.6, 0.65] # final mass to target

Op_1 = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
Op_2 = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
Op_3 = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
Op_4 = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
Op_5 = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
List_dataf = [Op_1, Op_2, Op_3, Op_4, Op_5]

const Cd = 310      # Drag Coeficient
const Tmax = 3.5    # Maximal Thrust
const ρ = 500       # density decay rate
const b = 2         # Positive real number depending on the engine

F0(x) = begin
    h, v, m = x
    D = Cd * v^2 * exp(-(h - 1)) # Drag force
    return [ v, -D/m - 1/h^2, 0 ]
end

F1(x) = begin
    h, v, m = x
    return [ 0, Tmax/m, -b*Tmax ]
end

for i in [1,2,3,4,5]
    @def ocp_as begin # definition of the optimal control problem

        tf  R, variable
        t   [ t0, tf ], time
    end
end

```

```

x = (h, v, m)   R3, state
u   R, control

x(t0) == [ h0, v0, m0 ]
m(tf) == mass[i],          (1)
0   u(t)   1
h(t)   h0

ẋ(t) == F0(x(t)) + u(t) * F1(x(t))

h(tf) → max

end;

direct_sol_as = solve(ocp_as; grid_size=100)

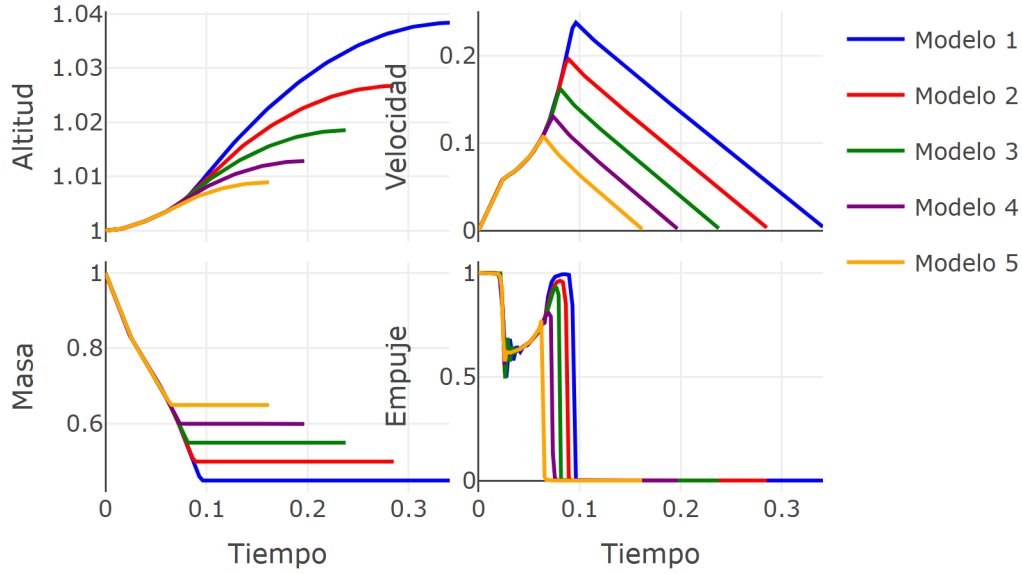
t_as = direct_sol_as.time_grid

x_as = direct_sol_as.state
u_as = direct_sol_as.control

for j in 1:100
    List_dataf[i].t[j] = t_as[j]
    List_dataf[i].x_h[j] = x_as(t_as[j])[1]
    List_dataf[i].x_v[j] = x_as(t_as[j])[2]
    List_dataf[i].x_m[j] = x_as(t_as[j])[3]
    List_dataf[i].u_t[j] = u_as(t_as[j])
end
end

```

Análisis de Sensibilidad



5.2.4. Interpretación de los resultados

La solución óptima obtenida revela una estructura de control característica en tres fases claramente diferenciadas. En la gráfica del empuje (control $u(t)$), observamos que el cohete tiene un inicio con empuje máximo durante el intervalo $[t_0, t_1]$, lo que corresponde a la fase de ascenso inicial donde se busca vencer la atracción gravitacional.

Posteriormente, en el siguiente intervalo $[t_1, t_2]$, el control muestra un comportamiento **singular**, donde $u(t)$ toma valores intermedios entre 0 y 1. Esta fase corresponde a un arco singular, donde la función de conmutación $\phi(t)$ se anula en un intervalo no trivial. Durante esta fase, el cohete mantiene un equilibrio entre el consumo de combustible y la ganancia de altitud.

En el intervalo $[t_2, t_3]$, observamos un arco límite donde la restricción de velocidad $v(t) \leq v_{\max}$ se activa. El control $u(t)$ ajusta automáticamente el empuje para mantener la velocidad en su límite superior sin sobrepasarla, evidenciado por el comportamiento plano de la curva de velocidad en este intervalo.

Finalmente, en $[t_3, t_f]$, el control se apaga por completo permitiendo que el cohete continúe su ascenso por inercia hasta alcanzar la altitud máxima, momento en el cual la velocidad se anula simultáneamente con el agotamiento del combustible hasta alcanzar la masa final $m_f = 0.6$.

La solución obtenida se caracteriza por la secuencia de arcos: **Bang-Singular-Límite-Bang**, denotada como (1-S-B-0) (Seywald y Cliff (1993)).

La comparación que vemos entre el método directo e indirecto (Figura 5.1) muestra una concordancia notable, validando ambos enfoques. Las pequeñas discrepancias observables en las transiciones entre fases se deben a la discretización empleada en el método directo (*grid_size=100*), mientras que el método indirecto captura estas transiciones con mayor precisión al resolver el sistema continuo de ecuaciones diferenciales.

6. Problema Alterno

6.1. Método Directo

Se formula un problema alternativo en el que la función objetivo consiste en maximizar la masa final del cohete al alcanzar una altitud final fija. Este enfoque permite evaluar la eficiencia en el consumo de combustible. Para garantizar la coherencia en el análisis, se emplean los mismos datos y restricciones del problema original, preservando la dinámica del sistema y las condiciones operativas.

$$\left\{ \begin{array}{l} \max m(t_f) \\ \dot{h} = v \\ \dot{v} = \frac{T-D}{m} - \frac{1}{h^2} \\ \dot{m} = -\frac{T}{c} \\ T \in [0, T_{max}] \\ h(0) = 1, v(0) = 0, m(0) = 1 \\ h(t_f) = 1.01257 \\ t_f \text{ libre} \end{array} \right.$$

Para obtener los resultados numéricos de este problema, resolvemos el problema mediante el método directo, implementado en el lenguaje de programación Julia.

```
using OptimalControl
using NLPModelsIpopt
using DataFrames
```

Definimos los datos del problema

```
t0_m = 0           # Tiempo Inicial
h0_m = 1           # Altitud Inicial
v0_m = 0           # Velocidad Inicial
m0_m = 1           # Masa Inicial
h_target = 1.01283 # Altitud Final Objetivo
```

```

mf_min = 0.2      # Masa Final Mínima Autorizada
const Cd = 310    # Coeficiente de Arrastre
const Tmax = 3.5  # Empuje Máximo
const  = 500      # Tasa de caída de Densidad
const b = 2       # Constante positiva de acuerdo al motor

```

```

@def ocp_mass begin
    tf_m  R, variable
    t     [ t0_m, tf_m ], time
    x = (h, v, m)  R3, state
    u  R, control

    x(t0_m) == [ h0_m, v0_m, m0_m ]
    h(tf_m) == h_target, (1)
    m(tf_m)  mf_min
    0  u(t)  1
    h(t)  h0_m

     $\dot{x}(t) == F0(x(t)) + u(t) * F1(x(t))$ 

    m(tf_m) → max
end;

```

```

F0(x) = begin
    r, v, m = x
    D = Cd * v2 * exp(-*(r - 1)) # Drag force
    return [ v, -D/m - 1/r2, 0 ]
end

```

```

F1(x) = begin
    r, v, m = x
    return [ 0, Tmax/m, -b*Tmax ]
end

```

```

direct_sol_mass = solve(ocp_mass; grid_size=100)

```

This is Ipopt version 3.14.17, running with linear solver MUMPS 5.7.3.

```

Number of nonzeros in equality constraint Jacobian...:    2104
Number of nonzeros in inequality constraint Jacobian.:      1
Number of nonzeros in Lagrangian Hessian.....:    1111

```

```

Total number of variables.....:      405
      variables with only lower bounds:    101
      variables with lower and upper bounds: 101
      variables with only upper bounds:     0
Total number of equality constraints.....:    304
Total number of inequality constraints.....:     1
      inequality constraints with only lower bounds:    1
      inequality constraints with lower and upper bounds: 0
      inequality constraints with only upper bounds:    0

```

iter	objective	inf_pr	inf_du	lg(mu)	d	lg(rg)	alpha_du	alpha_pr	ls
0	-1.0000000e-01	9.00e-01	1.00e+00	0.0	0.00e+00	-	0.00e+00	0.00e+00	0
1	-5.2232164e-01	1.04e-02	5.64e+01	-0.5	9.00e-01	-	7.98e-01	9.90e-01f	1
2	-5.5285947e-01	5.46e-03	7.09e+02	-0.6	3.28e-01	2.0	9.98e-01	9.90e-01h	1
3	-6.4577962e-01	2.33e-03	6.46e+02	-1.1	2.23e-01	1.5	1.00e+00	1.00e+00h	1
4	-6.0124963e-01	1.62e-03	1.33e+07	-0.5	1.31e+00	1.0	1.00e+00	5.75e-01f	1
5	-5.4253619e-01	8.12e-04	6.06e+06	-0.6	5.78e-01	-	7.01e-01	4.81e-01h	1
6	-5.2059521e-01	5.98e-04	2.71e+06	-1.3	1.98e-01	-	5.63e-01	2.50e-01f	3
7	-4.9742555e-01	4.68e-04	7.92e+06	-2.0	1.17e-01	0.6	6.55e-01	2.50e-01h	3
8	-4.4368540e-01	1.85e-04	2.09e+01	-2.0	5.37e-02	1.0	1.00e+00	1.00e+00h	1
9	-4.9559548e-01	2.91e-04	3.46e+00	-2.3	1.90e-01	-	1.00e+00	1.00e+00f	1
iter	objective	inf_pr	inf_du	lg(mu)	d	lg(rg)	alpha_du	alpha_pr	ls
10	-5.0093351e-01	2.44e-06	2.49e-02	-4.0	5.34e-03	0.5	1.00e+00	1.00e+00h	1
11	-5.0940839e-01	1.32e-05	1.79e-02	-5.5	1.63e-02	0.0	1.00e+00	1.00e+00f	1
12	-5.2708055e-01	5.86e-05	4.89e+01	-6.7	4.69e-02	-0.4	1.00e+00	8.30e-01h	1
13	-5.4994473e-01	1.12e-04	6.34e+00	-6.2	1.34e-01	-0.9	1.00e+00	4.89e-01h	1
14	-5.6065923e-01	1.10e-04	3.24e+01	-6.5	4.39e-01	-1.4	1.00e+00	8.76e-02h	1
15	-5.6809749e-01	1.04e-04	2.54e+01	-6.4	1.94e+00	-1.9	1.00e+00	2.15e-02f	1
16	-5.7634469e-01	8.77e-05	3.47e+01	-6.8	2.25e-01	-1.4	1.00e+00	2.36e-01h	1
17	-5.8121261e-01	8.55e-05	3.61e+01	-7.0	6.93e-01	-1.9	1.00e+00	1.03e-01h	1
18	-5.8419394e-01	8.01e-05	2.70e+01	-6.7	2.36e+00	-2.4	1.00e+00	3.62e-02h	1
19	-5.8792501e-01	7.02e-05	2.89e+01	-7.0	1.44e-01	-2.0	1.00e+00	2.01e-01h	1
iter	objective	inf_pr	inf_du	lg(mu)	d	lg(rg)	alpha_du	alpha_pr	ls
20	-5.8983597e-01	6.28e-05	2.71e+01	-7.1	4.17e-01	-2.4	1.00e+00	1.05e-01h	1
21	-5.9070972e-01	6.00e-05	2.63e+01	-7.1	1.21e+00	-2.9	1.00e+00	4.51e-02h	1
22	-5.9127966e-01	5.82e-05	1.08e+00	-6.5	2.56e+00	-3.4	1.00e+00	2.99e-02h	1
23	-5.9161565e-01	5.71e-05	2.40e+02	-5.6	8.45e+00	-	1.00e+00	1.94e-02h	1
24	-6.0422424e-01	1.91e-03	1.65e+01	-5.6	7.61e-01	-	1.00e+00	9.16e-01h	1
25	-6.0177706e-01	8.27e-04	5.46e+00	-5.6	6.45e-01	-	9.29e-01	5.67e-01h	1
26	-5.9948065e-01	6.25e-05	1.74e-02	-5.1	7.21e-01	-	1.00e+00	1.00e+00h	1
27	-5.9983571e-01	1.38e-05	3.94e-04	-5.6	2.93e-01	-	1.00e+00	1.00e+00h	1
28	-5.9999480e-01	5.28e-06	1.69e-04	-6.5	1.45e-01	-	1.00e+00	1.00e+00h	1
29	-6.0001951e-01	1.32e-06	4.56e-05	-7.6	1.01e-01	-	1.00e+00	1.00e+00h	1

Number of Iterations.....: 29

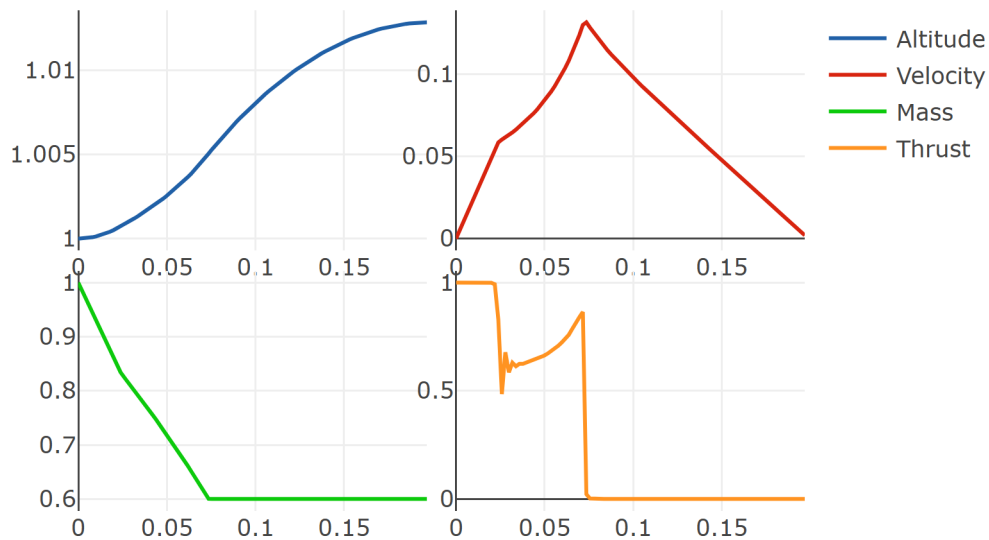
	(scaled)	(unscaled)
Objective.....:	-6.0001951350455163e-01	-6.0001951350455163e-01
Dual infeasibility.....:	4.5634917827996978e-05	4.5634917827996978e-05
Constraint violation....:	5.0046210509080424e-09	1.3167889997001136e-06
Variable bound violation:	0.0000000000000000e+00	0.0000000000000000e+00
Complementarity.....:	2.6395316602666028e-07	2.6395316602666028e-07
Overall NLP error.....:	2.6395316602666028e-07	4.5634917827996978e-05

Number of objective function evaluations	= 36
Number of objective gradient evaluations	= 30
Number of equality constraint evaluations	= 36
Number of inequality constraint evaluations	= 36
Number of equality constraint Jacobian evaluations	= 30
Number of inequality constraint Jacobian evaluations	= 30
Number of Lagrangian Hessian evaluations	= 29
Total seconds in IPOPT	= 3.725

EXIT: Optimal Solution Found.

OptimalControlSolution

Problema Alterno



6.2. Análisis de Sensibilidad

Al igual que con el Capítulo 5, se realiza un análisis de sensibilidad del Capítulo 6 variando el parámetro de la altura final tomando los datos de la salida de la Sección 5.2.3 del problema original. Este análisis se lleva a cabo a través de simulaciones numéricas (haciendo uso de la Sección 5.2.1), se puede examinar la relación entre la eficiencia del control y la masa final alcanzada.

```
using OptimalControl
using NLPModelsIpopt
using DataFrames

t0_ms = 0          # Tiempo Inicial
h0_ms = 1          # Altitud Inicial
v0_ms = 0          # Velocidad Inicial
m0_ms = 1          # Masa Inicial
h_targets = [1.038372, 1.026699, 1.018521, 1.012829] # Altitud
               ↪ Final Objetivo
mf_min = 0.2       # Masa Final Mínima Autorizada
const Cd = 310     # Coeficiente de Arrastre
const Tmax = 3.5   # Empuje Máximo
const rho = 500    # Tasa de caída de Densidad
const b = 2        # Constante positiva de acuerdo al motor

Op_1_ms = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
Op_2_ms = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
Op_3_ms = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
Op_4_ms = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
#Op_5_ms = DataFrame(rand(100, 5), [:t, :x_h, :x_v, :x_m, :u_t])
List_dataf_ms = [Op_1_ms, Op_2_ms, Op_3_ms, Op_4_ms]

F0(x) = begin
    h, v, m = x
    D = Cd * v^2 * exp(-(h - 1)) # Drag force
    return [ v, -D/m - 1/h^2, 0 ]
end

F1(x) = begin
    h, v, m = x
    return [ 0, Tmax/m, -b*Tmax ]
end
```

```

for i in [1,2,3,4]
    @def ocp_ms begin # definition of the optimal control problem

        tf_ms R, variable
        t [ t0_ms, tf_ms ], time
        x = (h, v, m) R3, state
        u R, control

        x(t0_ms) == [ h0_ms, v0_ms, m0_ms ]
        h(tf_ms) == h_targets[i], (1)
        m(tf_ms) mf_min
        0 u(t) 1
        h(t) h0_ms

        xdot(t) == F0(x(t)) + u(t) * F1(x(t))

        m(tf_ms) → max

    end;

sol_ms = solve(ocp_ms; grid_size=100)

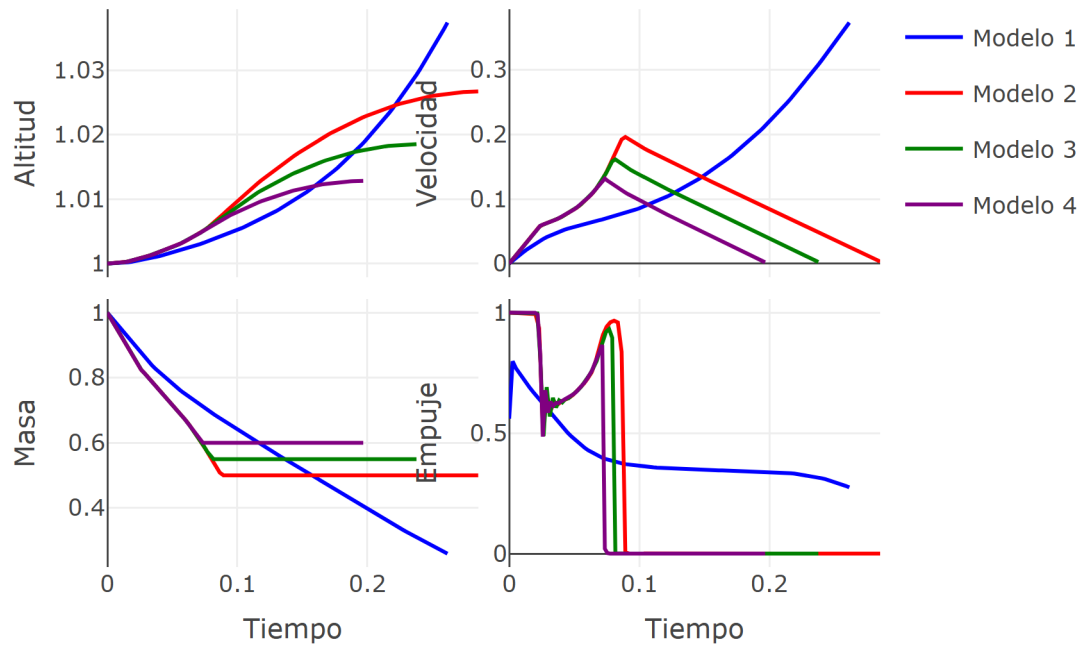
t_ms = sol_ms.time_grid

x_ms = sol_ms.state
u_ms = sol_ms.control

for j in 1:100
    List_dataf_ms[i].t[j] = t_ms[j]
    List_dataf_ms[i].x_h[j] = x_ms(t_ms[j])[1]
    List_dataf_ms[i].x_v[j] = x_ms(t_ms[j])[2]
    List_dataf_ms[i].x_m[j] = x_ms(t_ms[j])[3]
    List_dataf_ms[i].u_t[j] = u_ms(t_ms[j])
end
end
end

```

Análisis de Sensibilidad



7. Conclusiones

Este trabajo estableció el vínculo estructural entre el cálculo variacional clásico y la teoría del control óptimo, destacando cómo ambos marcos comparten una base conceptual común con base en la noción de variación. Se mostró que las ecuaciones de Euler-Lagrange constituyen el punto de partida natural para la formulación de las condiciones de optimalidad que surgen al introducir restricciones dinámicas en forma de ecuaciones diferenciales.

Se presentó el Teorema de Noether en su versión multivariable, mostrando así la relación entre simetrías de un funcional y leyes de conservación. Esta formulación resultó fundamental para comprender la estructura matemática subyacente tanto en problemas variacionales como en sistemas controlados, mostrando que las invariantes del sistema conducen naturalmente a cantidades conservadas que caracterizan la dinámica óptima.

La transición del enfoque Lagrangiano al Hamiltoniano, mediante la transformación de Legendre-Fenchel, permitió reformular los problemas de optimización en términos de variables duales. Esta dualidad se manifestó particularmente útil al introducir las variables adjuntas o de coestado, que desempeñaron un papel central en la derivación del Principio del Mínimo de Pontryagin. Se mostró que este principio constituye una generalización natural de las ecuaciones de Euler-Lagrange cuando los controles están sujetos a restricciones de conjunto.

El estudio de los problemas de control se centró en aquellos de tipo Bolza, por ser la forma más general de los problemas de control óptimo, pues contiene como casos particulares a las formulaciones de Lagrange y de Mayer. Esta elección nos permitió abordar un marco en el que la función terminal desempeña un papel esencial en la formulación de las condiciones de frontera y en la estructura del funcional aumentado. A partir de esta generalidad, fue posible derivar las condiciones de optimalidad en su forma más completa, incluyendo las ecuaciones adjuntas, la condición de estacionariedad y las condiciones de transversalidad.

Asimismo, la versión del teorema de Filippov adoptada en el trabajo, resulta particularmente adecuada para garantizar la existencia de soluciones óptimas en este contexto. Este enfoque permite extender el análisis a sistemas donde las funciones de control pueden presentar discontinuidades o pertenecer a conjuntos no convexos, preservando la validez de las condiciones necesarias. La elección de esta versión del teorema subraya el interés del trabajo por mantener la mayor generalidad posible dentro del plano determinista.

El estudio del problema de Goddard como caso de aplicación validó la teoría desarrollada y permitió contrastar dos enfoques metodológicos complementarios: el método directo, basado en la discretización del problema continuo y su conversión en un problema de programación no lineal, y el método indirecto, fundamentado en las condiciones necesarias de optimalidad derivadas del Principio de Pontryagin. Ambos métodos condujeron a soluciones consistentes, observándose la estructura bang-bang característica del control óptimo en sistemas con Hamiltonianos afines respecto al control. La implementación numérica en Julia, utilizando las paqueterías `OptimalControl.jl` y `NLPModelsIpopt.jl`, demostró la viabilidad computacional de ambos enfoques.

La perspectiva multivariable adoptada a lo largo del trabajo permitió formular resultados generales, aunque las aplicaciones se restringieron al caso donde el dominio temporal es un intervalo compacto. Esta restricción, si bien limitó el alcance inmediato de los resultados, facilitó la exposición de los conceptos fundamentales y su implementación numérica, sentando las bases para extensiones futuras a problemas con horizontes infinitos o dominios más generales.

Los resultados obtenidos abrieron diversas líneas de investigación futura que permitirían extender y profundizar el alcance de este trabajo. En primer lugar, resultaría natural incorporar incertidumbre en la formulación mediante control óptimo estocástico, donde las ecuaciones de estado incluirían términos de ruido modelados por procesos estocásticos, conduciendo a formulaciones basadas en la ecuación de Hamilton-Jacobi-Bellman. Asimismo, la extensión a problemas de control robusto permitiría considerar perturbaciones estocásticas en los parámetros del sistema, garantizando desempeño óptimo ante escenarios adversos. El estudio de problemas con retardos temporales en la dinámica o en el control extendería la teoría desarrollada a sistemas con memoria, fenómeno común en sistemas de comunicación y control distribuido. La formulación de problemas de control óptimo en espacios de dimensión infinita, utilizando teoría de semigrupos y ecuaciones en derivadas parciales, permitiría abordar sistemas gobernados por ecuaciones de difusión, ondas o transporte. Finalmente, la exploración de métodos de aprendizaje por refuerzo profundo como alternativa a los enfoques clásicos presentados abriría posibilidades para resolver problemas de alta dimensionalidad donde los métodos tradicionales enfrentan limitaciones computacionales, estableciendo puentes entre la teoría clásica del control óptimo y los desarrollos recientes en inteligencia artificial.

Referencias

- Bernoulli, Johann. 1996. «Problema novum ad cujus solutionem mathematici invitantur». *Acta Eruditorum* 15 (1696): 269.
- Bryson, Arthur E., y Yu-Chi Ho. 1969. *Applied Optimal Control*. New York: Blaisdell Publishing Company.
- Caillau, Jean-Baptiste, Olivier Cots, Joseph Gergaud, Pierre Martinon, y Sophia Sed. s. f. «OptimalControl.jl: a Julia package to model and solve optimal control problems with ODE's». <https://doi.org/10.5281/zenodo.13336563>.
- Cesari, Lamberto. 1983. *Optimization—theory and applications: problems with ordinary differential equations*. Springer New York, NY.
- Evans, Lawrence C. 2022. *Partial differential equations*. Vol. 19. American Mathematical Society.
- Ferguson, James. 2004. «A brief survey of the history of the calculus of variations and its applications». *arXiv preprint math/0402357*.
- Gelfand, IM, y SV Fomin. 1963. «Calculus of variations prentice-Hall». Inc., Englewood Cliffs 7.
- Goddard, Robert H. 1919. *A Method of Reaching Extreme Altitudes: The seminal text on rocket science that foretold the Space Age*. Sibyl Publishing.
- Hamel, G. 1927. «Über eine mit dem Problem der Rakete zusammenhängende Aufgabe der Variationsrechnung». *Zeitschrift für angewandte Mathematik und Mechanik* 7 (6): 451-52.
- Kirk, Donald E. 2004. *Optimal control theory: an introduction*. Courier Corporation.
- McShane, E. J. 2015. *Integration*. Princeton Legacy Library. Princeton University Press. <https://books.google.com.mx/books?id=zjXWCgAAQBAJ>.
- Munkres, J. R. 2000. *Topology*. Prentice Hall. <https://books.google.com.mx/books?id=OogGyAEACAAJ>.
- Pontryagin, Lev Semenovich. 2018. *Mathematical theory of optimal processes*. Routledge.
- Poznyak, Alexander S. 2008. *Advanced Mathematical Tools for Control Engineers: Volume 1: Deterministic Systems*. Vol. 1. Elsevier.
- Rockafellar, R. Tyrrell. 197d. C. *Convex analysis*. Princeton Mathematical Series. Princeton, N. J.: Princeton University Press.
- Rudin, W. 1976. *Principles of Mathematical Analysis*. International series en pure y applied mathematics. McGraw-Hill. <https://books.google.com.mx/books?id=kwqzPAAACAAJ>.
- Seywald, Hans, y Eugene M Cliff. 1993. «Goddard problem in presence of a dynamic pressure limit». *Journal of Guidance, Control, and Dynamics* 16 (4): 776-81.

- Spivak, M. 2008. *Calculus*. Publish or Perish. <https://books.google.com.mx/books?id=6jNLPwAACAAJ>.
- Sussmann, Hector J, y Jan C Willems. 1997. «300 years of optimal control: from the brachystochrone to the maximum principle». *IEEE Control Systems Magazine* 17 (3): 32-44.
- Sussmann, HJ. 2005. «Optimal control». *Three Decades of Mathematical System Theory: A Collection of Surveys at the Occasion of the 50th Birthday of Jan C. Willems*, 409-25.
- Touchette, Hugo. 2005. «Legendre-Fenchel transforms in a nutshell». URL <http://www.maths.qmul.ac.uk/ht/archive/lfth2.pdf>, 25.
- Zlatskii, VT, y BN Kiforenko. 1975. «Optimum trajectories with singular arcs». *Automation and Remote Control* 35 (12): 1885-90.