UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y
DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA.

FUNCIONES DE PENALIZACIÓN PARA EL PROBLEMA
DE LAGRANGE CON RESTRICCIONES

TESIS
QUE PARA OPTAR POR EL GRADO DE:
MAESTRO EN CIENCIAS

PRESENTA:
EMANUEL LÓPEZ CERDA

DIRECTOR DE LA TESIS
Dr. JAVIER FERNANDO ROSENBLUETH LAGUETTE
I.I.M.A.S.

MÉXICO, D. F. 19 DE ENERO DE 2015.

Índice general

Introducción **1**

1. Aumentabilidad en dimensión finita **5**
1.1. Recordatorio 5
1.2. Aumentabilidad en dimensión finita 6
1.3. Funciones penalizadoras 9
1.4. Método numérico de multiplicadores 10

2. Cálculo variacional **13**
2.1. Planteamiento del problema 13
2.2. Ecuaciones de Euler-Lagrange 17
2.3. Función de Weierstrass 20
2.4. Condición de Jacobi 24

3. Condiciones de suficiencia **31**
3.1. Resultados auxiliares 31
3.2. Campos de Mayer 36

4. Problemas con restricciones **39**
4.1. Problemas isoperimétricos 39
4.2. Problema con restricciones tipo Lagrange 44
4.3. Conos tangentes, conos derivados, conjuntos derivados y resultados técnicos 45
4.4. Principio del máximo 48
4.4.1. Formulación Hamiltoniana 58
4.4.2. Formulación Lagrangiana 59

5. Aumentabilidad en dimensión infinita **63**
5.1. Un nuevo enfoque 63
5.2. Suficiencia y aumentabilidad 65

Conclusiones **73**

Apéndice **75**

Introducción

El cálculo variacional se originó hace poco más de trescientos años a partir del problema que planteó Johann Bernoulli a la comunidad científica en 1696 para encontrar la curva por la cual una partícula desciende de un punto a otro en el menor tiempo posible, sujeto a la gravedad y sin considerar la fricción.

Al problema anterior se le conoce como el problema de la braquistócrona (la palabra braquistócrona proviene de las palabras griegas *brachistos* y *chronos* que signfican *el más corto* y *tiempo* respectivamente). El problema fue resuelto por el mismo Johann Bernoulli, por su hermano Jakob, y por Leibniz, L'Hôpital y Newton, entre otros. Poco tiempo después, Euler y Lagrange comenzaron a desarrollar la teoría para resolver este tipo de problemas, dando origen al cálculo variacional (o cálculo de variaciones).

El problema básico del cálculo variacional consiste en, dada una función $L\colon [a,b] \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ llamada el *Lagrangiano*, minimizar una funcional del tipo

$$I(x) := \int_a^b L(t, x(t), \dot{x}(t))dt$$

sobre todas las curvas $x\colon [a,b] \to \mathbb{R}^n$ que pertenezcan a cierto espacio de funciones y tengan sus puntos inicial y final fijos.

Euler y Lagrange mostraron que una condición necesaria para que una curva x sea un mínimo consiste en satisfacer la ecuación que lleva sus nombres, la ecuación de Euler-Lagrange:

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0.$$

Esta condición es análoga a la condición de que en el punto donde se alcanza un mínimo de una función $f\colon \mathbb{R}^n \to \mathbb{R}$ diferenciable se necesita cumplir que su primera derivada se anule en ese punto ($f'(x) = 0$). El siguiente paso es encontrar la condición análoga a la condición de que la segunda derivada sea no negativa. Varias personas dieron respuesta a este problema como lo fueron Weierstrass, Jacobi, entre otros; proporcionaron condiciones necesarias y suficientes (las cuales llevan sus nombres) para que un extremo (es decir una curva que satisfaga las ecuaciones de Euler-Lagrange) sea un mínimo para problemas con ciertas condiciones.

Desde los primeros resultados de Euler y Lagrange hasta nuestros días se han investigado condiciones necesarias y suficientes para determinar cuándo una curva x es un mínimo para I de acuerdo con diferentes condiciones impuestas a x, las cuales incluyen, por ejemplo, que sea continuamente diferenciable o absolutamente continua y que satisfaga ciertas restricciones de tipo isoperimétrico o mixtas (impuestas sobre la curva y su derivada) en forma de igualdades y/o desigualdades.

Por otra parte, en la teoría de control óptimo, nuestro funcional I no depende directamente de la derivada $\dot{x}$ sino de una función u (que puede ser continua a trozos) por lo que el funcional ahora toma la forma

$$I(x,u) := \int_a^b L(t, x(t), u(t))dt$$

en presencia de ciertas dinámicas del tipo

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

También se pueden introducir restricciones como las de cálculo variacional solo que ahora dependerán de u. Pontryagin, en colaboración con otros investigadores, publicó en 1962 el libro titulado *The Mathematical Theory of Optimal Processes* en el que se derivan condiciones de primer orden para ciertos problemas bastante generales de control óptimo en términos de un *principio máximo*.

El concepto de aumentabilidad se ha utilizado, en particular, al estudiar problemas de optimización de funciones $f\colon \mathbb{R}^n \to \mathbb{R}$ sujetas a restricciones en forma de igualdades $g_i(x) = 0$ y desigualdades $g_j(x) \leq 0$. Las condiciones clásicas necesarias (suponiendo una condición de normalidad) y suficientes para este tipo de problemas se expresan generalmente en términos de multiplicadores de Lagrange. El propósito de la teoría de aumentabilidad consiste en transformar este problema en uno nuevo sin restricciones a través del cual se puedan derivar las condiciones necesarias sin la hipótesis de normalidad. Para este propósito, en particular, es que se introducen las llamadas funciones penalizadoras.

Para problemas en espacios de dimensión infinita, como los problemas de cálculo de variaciones y control óptimo mencionados anteriormente, es posible generalizar el concepto de aumentabilidad de manera que las condiciones obtenidas para problemas con restricciones sean consecuencia de las condiciones conocidas para problemas sin restricciones. En este trabajo estudiaremos con detalle este aspecto de la teoría para problemas de cálculo variacional cuando las restricciones están expresadas en forma de igualdades

$$\phi_i(t, x, \dot{x}) = 0, \quad i = 1, \ldots, m.$$

En el primer capítulo se hará un breve resumen de la teoría para el caso de dimensión finita. Veremos las condiciones que debe satisfacer un punto x_0 en $\mathbb{R}^n$ que minimiza una función $f\colon \mathbb{R}^n \to \mathbb{R}$ dada sujeta a restricciones del tipo $g(x) = 0$. Por medio del concepto de aumentabilidad veremos cómo es posible obtener dichas condiciones a través de un problema asociado sin restricciones (ver [9] y referencias incluidas). Es importante mencionar que esta teoría es la base de métodos de multiplicadores desarrollados inicialmente por Hestenes [9] para resolver numéricamente problemas de optimización con restricciones (ver [1–4, 12] para una idea de posibles aplicaciones de dichos métodos). Una generalización ampliamente conocida en la literatura se debió a Rockafellar [14] para el caso de análisis convexo.

En el siguiente capítulo se desarrollará la teoría básica de cálculo de variaciones la cual abarca las ecuaciones de Euler-Lagrange y las condiciones de Legendre, Weierstrass y Jacobi. Daremos algunos ejemplos donde se puede ver la importancia del cálculo variacional en problemas reales. En el tercer capítulo daremos condiciones de suficiencia en cálculo de variaciones mediante la introducción de campos de Mayer. En el cuarto capítulo estudiaremos problemas variacionales que involucran restricciones en forma de igualdades. Como corolario del principio máximo en control óptimo obtendremos las condiciones necesarias clásicas para problemas con restricciones del cálculo variacional.

En el último capítulo se introducirá una definición de aumentabilidad para el problema de Lagrange con restricciones en forma de igualdades. Dicha definición se basa en aumentar el Lagrangiano a través de funciones de penalización que permiten generalizar los resultados de dimenisión finita de manera que las condiciones necesarias clásicas (que requieren la hipótesis de normalidad) se puedan obtener como consecuencia de ese tipo de aumentabilidad (prescindiendo de dicha hipótesis). Por otro lado, para mínimos locales fuertes o débiles, probaremos que las condiciones suficientes clásicas implican el tipo de aumentabilidad correspondiente. Este resultado es crucial en la teoría ya que, además de justificar ampliamente la búsqueda de una solución al problema aumentable sin restricciones, permite desarrollar una nueva demostración de suficiencia con posibles generalizaciones a problemas de control óptimo.

Capítulo 1

Aumentabilidad en dimensión finita

1.1. Recordatorio

Para generar interés en el lector y le sea comprensible el concepto de aumentabilidad en cálculo de variaciones se desarrollará, antes que nada, el concepto de aumentabilidad en dimensión finita, es decir, nuestras funciones a minimizar son de la forma $f\colon \mathbb{R}^n \to \mathbb{R}$. Se podrán observar las ideas principales en las que se sustenta esta teoría para después tratar de generalizarlas al caso de dimensión infinita.

Como nuestro interés es optimizar funcionales, recordaremos el teorema que establece condiciones necesarias y suficientes para encontrar mínimos locales para una función $f\colon U \subset \mathbb{R}^n \to \mathbb{R}$.

Teorema 1.1.1. *Sea $U \subset \mathbb{R}^n$ abierto y $f\colon U \to \mathbb{R}$ de clase C^2. Si f alcanza un mínimo local en $x_0 \in U$ entonces $f'(x_0) = 0$ y $f''(x_0) \geq 0$. Inversamente, si $x_0 \in U$ es tal que $f'(x_0) = 0$ y $f''(x_0) > 0$ entonces existen δ y m tales que si $\|x - x_0\| < \delta$ entonces $f(x) \geq f(x_0) + m\|x - x_0\|^2$.*

Ahora generalicemos este problema. Queremos que $x \in U$ satisfaga restricciones del tipo

$$g_i(x) = 0, \quad (x \in \mathbb{R}^n,\ i = 1, \ldots, m < n) \tag{1.1.1}$$

con $g_i\colon \mathbb{R}^n \to \mathbb{R}$. Supondremos que f, g_i son de clase C^2. Condiciones necesarias de primer y segundo orden para este problema se pueden enunciar de la siguiente manera.

Teorema 1.1.2. *(Multiplicadores de Lagrange) Supongamos que f alcanza un mínimo local en $x_0 \in U$ bajo las restricciones*

$$g_i(x) = 0, \qquad i = 1, \ldots, m < n$$

y que x_0 es normal, es decir, la matriz

$$\left(\frac{\partial g_i(x)}{\partial x_j}\right)$$

evaluada en x_0 tiene rango m. Entonces existe un único $\lambda = (\lambda_1, \ldots, \lambda_m) \in \mathbb{R}^m$ tal que para

$$F(x) = f(x) + \langle \lambda, g(x) \rangle$$

se satisface $\nabla F(x_0) = 0$. Además, $F''(x_0; h) \geq 0$ para toda $h \in \mathbb{R}^n$ tal que $g_i'(x_0; h) = 0$.

Nota 1. *$\langle \cdot, \cdot \rangle$ representa el producto interior euclideo y $g'(x; h)$ la derivada direccional de x en la dirección del vector h, es decir, $g'(x; h) = \langle \nabla g(x), h \rangle$.*

Para facilitar la escritura se introducirán los siguientes conjuntos.

Definición 1.1.3.

$$S := \{x \in \mathbb{R}^n \mid g_i(x) = 0,\ i = 1, \ldots, m\},$$

$$R_s(x_0) := \{h \in \mathbb{R}^n \mid g_i'(x_0; h) = 0,\ i = 1, \ldots, m\},$$

$$L(\lambda) := \{x_0 \in S \mid F'(x_0) = 0 \ y \ F''(x_0; h) \geq 0 \ \forall h \in R_s(x_0)\}.$$

Diremos que un punto x_0 satisface la regla de multiplicadores de Lagrange si pertenece al conjunto $L(\lambda)$.

1.2. Aumentabilidad en dimensión finita

Ahora que tenemos una condición necesaria para mínimos cuando el problema está sujeto a resctricciones, introduciremos el concepto de aumentabilidad. La idea es sumar una función no negativa a la función F de tal manera que para los elementos de S la función se anule. El propósito de esto es que si tenemos un mínimo en x_0 para la nueva función entonces tendremos un mínimo para f tal que satisface las restricciones $g_i(x) = 0$.

Definición 1.2.1. *Sean $\lambda \in \mathbb{R}^n$ y $\sigma \in \mathbb{R}^+ \setminus \{0\}$. Definimos*

$$H(x) := f(x) + \langle \lambda, g(x) \rangle + \sigma G(x)$$

donde

$$G(x) = \frac{1}{2} \sum_{i=1}^{m} g_i^2(x).$$

A la función H le llamaremos función aumentada con respecto a (λ, σ).

Definición 1.2.2. *Diremos que para el par $(\lambda, \sigma) \in \mathbb{R}^n \times \mathbb{R}^+ \setminus \{0\}$, el problema es aumentable en x_0 si $x_0 \in S$ y la función H alcanza un mínimo local en x_0.*

Notemos que $H(x) = f(x)$ para $x \in S$. Por lo tanto si el problema es aumentable en x_0 entonces x_0 minimiza localmente a f en S. Veamos que si el problema es aumentable entonces se satisface el teorema de multiplicadores de Lagrange.

Teorema 1.2.3. *Si $A(\lambda, \sigma)$ es el conjunto de $x_0 \in S$ tal que nuestro problema es aumentable en x_0 con respecto a (λ, σ) entonces $A(\lambda, \sigma) \subset L(\lambda)$.*

Demostración. Sea $x_0 \in A(\lambda, \sigma)$. Notemos que

$$H(x) = F(x) + \sigma G(x).$$

Como x_0 minimiza localmente a H, se sigue del Teorema 1.1.1 que

$$H'(x_0) = 0 \ \text{ y } \ H''(x_0) \geq 0.$$

Como G es no negativa y $G(x_0) = 0$, x_0 minimiza a G. Por el Teorema 1.1.1, $G'(x_0) = 0$, concluimos que

$$0 = H'(x_0) = F'(x_0) + \sigma G'(x_0) = F'(x_0),$$
$$0 \leq H''(x_0) = F''(x_0) + \sigma G''(x_0)$$

de manera que $F''(x_0; h) \geq 0$ siempre que $G''(x_0; h) = 0$. Ahora, como

$$G''(x_0; h) = \sum_1^m g_i'(x_0; h)^2,$$

tenemos que $F''(x_0; h) \geq 0$ siempre que $g_i'(x_0; h) = 0$ $(i = 1, \ldots, m)$ y por lo tanto $x_0 \in L(\lambda)$. $\square$

Veamos ahora que las condiciones suficientes clásicas en términos de multiplicadores de Lagrange implican aumentabilidad. La demostración se basa en el siguiente resultado (ver [9]).

Teorema 1.2.4. *Sean V un cono cerrado y P y Q dos formas cuadráticas tales que $P(x) > 0$ para toda $x \neq 0$ en V con $Q(x) = 0$. Supongamos que $Q(x) \geq 0$ en V. Entonces existe $\sigma_0 > 0$ tal que $P(x) + \sigma Q(x)$ es positiva para toda $\sigma \geq \sigma_0$ y $x \neq 0$ en V.*

Consideremos el conjunto

$$L'(\lambda) := \{x_0 \in S \mid F'(x_0) = 0 \text{ y } F''(x_0; h) > 0 \ \forall h \in R_s(x_0),\ h \neq 0\}$$

cuyos elementos satisfacen la regla de multiplicadores de Lagrange reforzada.

Teorema 1.2.5. *Supongamos que $x_0 \in L'(\lambda)$ para alguna $\lambda \in \mathbb{R}^m$. Entonces existen constantes positivas σ_0, τ y una vecindad N_0 de x_0 tales que, para toda $x \in N_0$,*

$$H(x) \geq H(x_0) + \tau \|x - x_0\|^2 \quad (\sigma \geq \sigma_0). \tag{1.2.1}$$

En particular, si $x \in N_0 \cap S$, entonces

$$f(x) \geq f(x_0) + \tau \|x - x_0\|^2.$$

Demostración. Definamos $P(h) = F''(x_0; h)$ y $Q(h) = G''(x_0, h)$. Aplicando el Teorema 1.2.4 obtenemos que existe σ_0 tal que

$$F''(x_0; h) + \sigma_0 G''(x_0; h) > 0 \quad (h \neq 0)$$

por lo que la función $H(x) = F(x) + \sigma_0 G(x)$ cumple

$$\nabla H(x_0) = \nabla F(x_0) + \sigma_0 \nabla G(x_0) = 0,$$

$$H''(x_0; h) = F''(x_0; h) + \sigma_0 G''(x_0; h) > 0.$$

Por el Teorema 1.1.1, existen $\tau > 0$ y una vecindad N_0 de x_0 tal que se cumple la desigualdad de (1.2.1) para $x \in N_0$. Como $G(x) \geq 0$ sigue cumpliendo esta desigualdad para $\sigma \geq \sigma_0$ se obtiene el resultado. □

Ejemplo 1.2.6. *La función* $f\colon \mathbb{R}^2 \to \mathbb{R}$ *dada por*

$$f(x, y) = x^2 + 2x + y^4$$

tiene un mínimo local estricto en $(x_0, y_0) = (0, 0)$ *sujeta a la restricción*

$$g(x, y) = xy - x = 0.$$

Como $\nabla g(0, 0) = (-1, 0) \neq (0, 0)$, *el problema es normal en* $(0, 0)$. *Probaremos ahora que el problema no es aumentable en ese punto. La función aumentada es*

$$\begin{aligned} H(x, y; \sigma) &= x^2 + 2x + y^4 + \lambda(xy - x) + \frac{\sigma}{2}(xy - y)^2 \\ &= x^2 \left[1 + \frac{\sigma}{2}(y - 1)^2\right] + (2 - \lambda)x + \lambda xy + y^4. \end{aligned}$$

Si el problema fuera aumentable en $(0, 0)$ *se seguiría que* $\nabla H(0, 0) = 0$ *por lo que* $\lambda = 2$. *En ese caso,*

$$H(x, y; \sigma) = x^2 \left[1 + \frac{\sigma}{2}(y - 1)^2\right] + 2xy + y^4.$$

Podemos suponer que $\sigma \geq 0$. *Para*

$$y^2 < \frac{1}{1 + 2\sigma}, \quad x = -\frac{y}{1 + 2\sigma},$$

se tiene que

$$H(x, y) \leq y^2 \left(y^2 - \frac{1}{1 + 2\sigma}\right) < 0 = H(0, 0)$$

y se llega a una contradicción. Por lo tanto el problema no es aumentable en $(0, 0)$.

1.3. Funciones penalizadoras

Veamos algunas propiedades que cumple la función $\sigma G(x)$ en H. A $\sigma G(x)$ se le conoce como función penalizadora.

Supongamos que tenemos funciones F, G continuas en un conjunto compacto N tales que $G(x) \geq 0$ en N,

$$S = \{x \in N \mid G(x) = 0\}$$

es no vacío y existe un único $x_0 \in N$ que minimiza a F en S.

Lema 1.3.1. *Sean $(\sigma_n)_{n\in\mathbb{N}}$ una sucesión de números y $(x_n)_{n\in\mathbb{N}}$ una sucesión de puntos en N tales que*

$$\lim_{n\to\infty} \sigma_n = \infty, \quad \limsup_{n\to\infty} \, [F(x_n) + \sigma_n G(x_n)] \leq F(x_0). \tag{1.3.1}$$

Entonces $\lim_{n\to\infty} x_n = x_0$.

Demostración. Como $G(x) \geq 0$, la relación (1.3.1) implica que

$$\limsup_{n\to\infty} F(x_n) \leq F(x_0), \quad \lim_{n\to\infty} G(x_n) = 0.$$

Por lo tanto cualquier límite $\bar{x}$ de una subsucesión convergente de (x_n) satisface $F(\bar{x}) \leq F(x_0)$, $G(\bar{x}) = 0$. Además por compacidad $\bar{x} \in N$ y la unicidad de x_0 implica que $\bar{x} = x_0$. De lo anterior se deduce que toda subsucesión convergente de (x_n) converge a x_0. Por último, como la sucesión (x_n) está acotada, esto es posible solo si $x_0 = \lim_{n\to\infty} x_n$. □

Teorema 1.3.2. *Sea x_σ un punto mínimo en N de la función aumentada*

$$H(x, \sigma) = F(x) + \sigma G(x).$$

Entonces

$$\lim_{\sigma\to\infty} x_\sigma = x_0, \quad \lim_{\sigma\to\infty} \sigma G(x_\sigma) = 0$$

donde x_0 es el punto mínimo de F en S. Además si $0 \leq \sigma < \bar{\sigma}$ entonces

$$H(x_\sigma, \sigma) \leq H(x_{\bar{\sigma}}, \bar{\sigma}) \leq F(x_0),$$

$$F(x_\sigma) \leq F(x_{\bar{\sigma}}) \leq F(x_0), \quad G(x_\sigma) \geq G(x_{\bar{\sigma}}) \geq 0.$$

Demostración. Como $G(x_0) = 0$ y x_σ minimiza a $H(x, \sigma)$, se tienen las desigualdades

$$H(x_\sigma, \sigma) = F(x_\sigma) + \sigma G(x_\sigma) \leq H(x_0, \sigma) = F(x_0),$$

$$F(x_\sigma) \leq F(x_0).$$

Supongamos que $\sigma < \bar{\sigma}$. Entonces, si $\bar{x} = x_{\bar{\sigma}}$ y $x = x_\sigma$, se tiene

$$0 \leq H(\bar{x}, \sigma) - H(x, \sigma) = F(\bar{x}) - F(x) + \sigma[G(\bar{x}) - G(x)], \tag{1.3.2}$$

$$0 \leq H(x, \bar{\sigma}) - H(\bar{x}, \bar{\sigma}) = F(x) - F(\bar{x}) + \bar{\sigma}[G(x) - G(\bar{x})].$$

Por lo tanto

$$0 \leq (\bar{\sigma} - \sigma)[G(x) - G(\bar{x})]$$

lo cual implica que $G(x) \geq G(\bar{x})$. Usando 1.3.2 obtenemos $F(x) \leq F(\bar{x})$.

Consideremos por último una sucesión $(\sigma_n)_{n\in\mathbb{N}}$ que tienda a $+\infty$ y sea $x_n = x_{\sigma_n}$. Por el Lema 1.3.1 tenemos que

$$\lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{\sigma_n} = x_0$$

lo cual implica que $\lim_{n\to\infty} x_\sigma = x_0$. $\square$

Corolario 1.3.3. *Si existe un número σ_0 tal que una de las condiciones*

$$F(x_{\sigma_0}) = F(x_0), \quad G(x_{\sigma_0}) = 0$$

se satisface, entonces $x_\sigma = x_0$ $(\sigma \geq \sigma_0)$, o sea, x_0 minimiza a $H(x, \sigma)$ en N.

Demostración. Si $F(x_{\sigma_0}) = F(x_0)$ entonces, por el Teorema 1.3.2, se tiene $G(x_{\sigma_0}) = 0$. Si $G(x_{\sigma_0}) = 0$ entonces $G(x_\sigma) = 0$ $(\sigma \geq \sigma_0)$. Como $F(x_\sigma) \leq F(x_0)$ y $G(x_\sigma) = 0$ $(\sigma \geq \sigma_0)$ tenemos que $x_\sigma = x_0$ $(\sigma \geq \sigma_0)$ por unicidad de x_0 como un punto mínimo de F en N sujeto a $G = 0$. $\square$

Corolario 1.3.4. *Si x_0 minimiza localmente a F en N entonces existe σ_0 tal que x_0 es un mínimo estricto de $H(x, \sigma)$ para toda $\sigma \geq \sigma_0$.*

Demostración. Sea N_0 vecindad de x_0 tal que $F(x) \geq F(x_0)$ para $x \in N_0$. Como $\lim_{\sigma\to\infty} x_\sigma = x_0$, existe σ_0 tal que $x_{\sigma_0} \in N_0$. Por lo tanto $F(x_{\sigma_0}) = F(x_0)$. Por el Corolario 1.3.3 $x_\sigma = x_0$ si $\sigma \geq \sigma_0$. Concluimos que en x_0 la función $H(x, \sigma)$ alcanza un mínimo si $\sigma \geq \sigma_0$. $\square$

1.4. Método numérico de multiplicadores

Ya sabemos cómo obtener mínimos mediante aumentabilidad, lo cual se realizó al suponer la existencia de multiplicadores de Lagrange, por lo tanto es de nuestro interés encontrar los multiplicadores λ y la constante σ de la función aumentada.

Sea $x_0 \in L(\lambda)$. Supongamos que la matriz

$$\left(\frac{\partial g_i(x)}{\partial x_j}\right)$$

evaluada en x_0 tiene rango máximo.

Nota 2. *Recordemos que nuestros multiplicadores de Lagrange los denotamos por $\lambda_1, \ldots, \lambda_n$.*

Definamos

$$H(x,\mu,\sigma) = f(x) + \sum_{i=1}^{m} \Big(\mu_i g_i(x) + \frac{\sigma}{2} g_i^2(x) \Big)$$

con $\mu = (\mu_1, \ldots, \mu_m) \in \mathbb{R}^m$. Para poder presentar el método numérico con el que se obtiene a λ_i, σ necesitamos del siguiente resultado ([9]).

Lema 1.4.1. *Sea M un conjunto acotado de multiplicadores $\mu = (\mu_1, \ldots, \mu_m)$. Si $H(x,\mu,\sigma)$ alcanza un mínimo en $x = x(\mu,\sigma)$ entonces*

$$\lim_{\sigma \to \infty} x(\mu,\sigma) = x_0$$

y

$$\lim_{\sigma \to \infty} (\mu_i + \sigma g_i(x(\mu,\sigma))) = \lambda_i$$

donde el límite es uniforme para $\mu \in M$ con $i = 1, \ldots, m$.

En la sección de funciones penalizadoras obtuvimos que si $(\sigma_n)_{\mathbb{N}}$ es una sucesión que tiende a infinito entonces podemos encontrar a x_0 como límite de la sucesión $(x_n)_{\mathbb{N}}$ donde en x_n la función $H(x,\sigma_n)$ alcanza un mínimo (Lema 1.3.1). Por el Lema 1.4.1 tenemos que

$$\lambda = \lim_{n \to \infty} \sigma_n g(x_n)$$

Mediante este método, conocido como el método de funciones penalizadoras, encontramos los multiplicadores de Lagrange y el punto x_0 donde se alcanza el mínimo. Sin embargo dicho procedimiento presenta dificultades numéricas cuando los multiplicadores son de magnitud grande; a continuación mostraremos un algoritmo más eficiente (Para su demostración véase [9]).

Teorema 1.4.2. *Sea $(\eta_n)_{\mathbb{N}}$ tal que $0 < \eta_0 \leq \eta_i$. Dado μ_0 existe σ_0 tal que para x_n y μ_n seleccionados de la siguiente manera:*

$$x_n \text{ minimiza a } H(x, \mu_{n-1}, \sigma_0 + \eta_{n-1}) \text{ en } N$$

y

$$\mu_n = \mu_{n-1} + \eta_{n-1} g(x_n)$$

entonces

$$\lim_{n \to \infty} x_n = x_0, \qquad \lim_{n \to \infty} \mu_n = \lambda.$$

Además existe n_0 y $\alpha, \beta, \gamma > 0$ tales que para $n \geq n_0$ tenemos

$$\|g(x_{n+1})\| \leq \frac{\|g(x_n)\|}{1 + \alpha \eta_n}$$

$$\|x_n - x_0\| \leq \beta \|g(x_n)\|, \qquad \|\mu_n - \lambda\| \leq \gamma \|g(x_n)\|.$$

Capítulo 2

Cálculo variacional

2.1. Planteamiento del problema

Antes de introducir el concepto de aumentabilidad en el cálculo variacional es indispensable conocer cómo encontrar mínimos, por lo que necesitamos establecer condiciones necesarias y suficientes que deben satisfacer.

Primero necesitamos introducir los objetos con los que estaremos trabajando.

Sean $T := [a, b]$ un intervalo cerrado en $\mathbb{R}$, $\xi_1, \xi_2 \in \mathbb{R}^n$, $A \subset T \times \mathbb{R}^n \times \mathbb{R}^n$ un conjunto abierto y $L\colon T \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ un funcional de clase C^2. A L comúnmente se le llama Lagrangiano.

Una función u es continua a trozos en un intervalo $[a, b]$ si este se puede dividir en un número finito de subintervalos $[t_i, t_{i+1}]$, $i = 1, ..., n$ tales que $t_1 = a$, $t_{n+1} = b$, la función u es continua en (t_i, t_{i+1}) y existen los límites derechos e izquierdos en cada subintervalo que se denominarán respectívamente $u(t_i + 0)$ y $u(t_i - 0)$.

Definamos a X como el conjunto de funciones $x\colon T \to \mathbb{R}^n$ que poseen derivada continua a trozos. Este conjunto es un espacio vectorial el cual se denota comúnmente como PWC^1.

Definición 2.1.1.

$$X(A) := \{x \in X \mid (t, x(t), \dot{x}(t)) \in A, \ \ t \in T\}$$

$$X_e(A) := \{x \in X(A) \mid x(a) = \xi_1, \ x(b) = \xi_2\}.$$

A los elementos de $X(A)$ les llamaremos trayectorias y a los de $X_e(A)$ trayectorias admisibles.

Definición 2.1.2. *Sea $I\colon X \to \mathbb{R}$ definido por*

$$I(x) := \int_a^b L(t, x(t), \dot{x}(t))dt.$$

A este funcional $I(x)$ se le conoce como el funcional acción.

Nuestro problema es encontrar $x_0 \in X_e(A)$ tal que minimice a I en $X_e(A)$.

Definición 2.1.3. *Diremos que una trayectoria x_0 es solución de nuestro problema si pertenece a*

$$S(A) := \{x_0 \in X_e(A) \mid I(x_0) \leq I(x),\ \forall x \in X_e(A)\}.$$

Los mínimos los clasificaremos en dos categorías, para esto consideraremos las siguientes normas en X:

$$\|x\|_0 := \sup\{|x(t)| : t \in T\},$$

$$\|x\|_1 := \|x\|_0 + \|\dot{x}\|_0.$$

Diremos que $x_0 \in X_e(A)$ es un mínimo fuerte si es un mínimo local de I en $X_e(A)$ con respecto a la norma $\|\cdot\|_0$, y es un mínimo débil si es un mínimo local de I con respecto a $\|\cdot\|_1$. Por lo anterior a la norma $\|\cdot\|_0$ se le denomina norma fuerte y a $\|\cdot\|_1$ norma débil.

Definición 2.1.4.

$$X_0 := \{x \in X \mid x(a) = x(b) = 0\}.$$

Este conjunto es un subespacio vectorial de X y es de utilidad ya que si tomamos $x \in X_e(A)$ y $y \in X_0$ entonces $x + y \in X_e(A)$. Por lo que podemos obtener elementos de $X_e(A)$ a partir de un $x_0 \in X_e(A)$ fijo.

Sea $x \in X_e(A)$ y $y \in X_0$. Como A es abierto entonces existe $\epsilon > 0$ tal que $x + \epsilon y \in X_e(A)$. Definamos a la función $F\colon \mathbb{R} \to \mathbb{R}$ de la siguiente manera

$$F(\epsilon) := \int_a^b L(t, x(t) + \epsilon y(t), \dot{x}(t) + \epsilon \dot{y}(t))dt.$$

Utilizando el teorema de Leibniz (veáse Apéndice Teorema 5.2.5),

$$\left.\frac{dF}{d\epsilon}\right|_{\epsilon=0} = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} \int_a^b L(t, x(t) + \epsilon y(t), \dot{x}(t) + \epsilon \dot{y}(t))dt =$$

$$\int_a^b \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} L(t, x(t) + \epsilon y(t), \dot{x}(t) + \epsilon \dot{y}(t))dt =$$

$$\int_a^b \left(\frac{\partial L}{\partial x} y + \frac{\partial L}{\partial \dot{x}} \dot{y}\right) dt.$$

Nota 3. *Para simplificar la escritura a veces escribiremos $L(x(t))$ en vez de*

$$L(t, x(t), \dot{x}(t))$$

y de igual manera con sus derivadas.

Definición 2.1.5. *Sea $x \in X_e(A)$ y $y \in X_0$,*

$$I'(x, y) := \int_a^b (L_x(x(t))y(t) + L_{\dot{x}}(x(t))\dot{y}(t))dt$$

es la primera variación de I a lo largo de x en la dirección y.

De la misma manera definimos la segunda variación mediante la primera variación de la primera variación, es decir:

$$I''(x, y, z) = \left.\frac{d}{d\epsilon}\right|_{\epsilon=0} I'(x + \epsilon y, z)$$

donde ahora la variación la hacemos en la dirección z. Cuando $z = y$ escribiremos $I''(x, y)$ en vez de $I''(x, y, y)$.

Definición 2.1.6. *Dados $x \in X_e(A)$, $y \in X_0$,*

$$I''(x, y) := \int_a^b (\langle y, L_{xx}(x(t))y\rangle + 2\langle y, L_{x\dot{x}}(x(t))\dot{y}\rangle + \langle \dot{y}, L_{\dot{x}\dot{x}}(x(t))\dot{y}\rangle)dt \quad (2.1.1)$$

es la segunda variación de I a lo largo de x en la dirección y. Para fines prácticos al integrando lo denotaremos como $2\Omega(t, y, \dot{y})$.

Sea $x_0 \in X_e(A)$, $y \in X_0$ y $F(\epsilon) = I(x_0 + \epsilon y)$. Si x_0 minimiza a I sobre $X_e(A)$ entonces $\epsilon = 0$ minimiza a F, por lo tanto

$$F'(0) = I'(x_0, y) = 0, \qquad F''(0) = I''(x_0, y) \geq 0$$

Por lo que tenemos el siguiente resultado.

Teorema 2.1.7. *Si x_0 minimiza a I sobre $X_e(A)$ entonces*

$$I'(x_0, y) = 0 \text{ y } I''(x_0, y) \geq 0 \qquad \forall y \in X_0.$$

Sea $x \in X$, veamos el comportamiento de las variaciones de $I(x)$ respecto a la norma $\|x\|_1$.
Sea $x_0 \in X$ fijo,

$$S_0 := \{(t, x_0(t), \dot{x}_0(t)) \mid t \in [a, b]\}.$$

S_0 es un conjunto compacto en A ya que $T = [a, b]$ también lo es y $x \in PWC^1([a, b])$. Como A es abierto entonces existe una vecindad S de $S_0 \subset A$ de radio δ tal que $\overline{S} \in A$. Sea $x(t)$ en esta vecindad, es decir, $\|x - x_0\|_1 < \delta$.

Recordemos que $L \in C^2$ por lo tanto sus derivadas parciales son uniformemente continuas en $\overline{S}$, si $D(t, x, \dot{x})$ es una derivada parcial de L entonces:

$$\lim_{\|x - x_0\| \to 0} D(t, x(t), \dot{x}(t)) = D(t, x_0(t), \dot{x}_0(t)), \quad \forall t \in [a, b].$$

Por lo tanto si $x \in S$

$$\begin{aligned} |I'(x, y) - I'(x_0, y)| \quad \leq \quad & \int_a^b \{(|L_x(x(t)) - L_x(x_0(t))|)|y|_1 \\ & +(|L_{\dot{x}}(x(t)) - L_{\dot{x}}(x_0(t))|)|\dot{y}|\}dt \leq \epsilon\|y\|_1. \end{aligned}$$

De la misma manera,

$$|I''(x,y,z) - I''(x_0,y,z)| \leq \epsilon \|y\|_1 \|z\|_1.$$

Por las desigualdades anteriores tenemos que

$$\lim_{x \to x_0} I'(x,y) = I'(x_0,y) \quad \text{uniformemente si} \quad \|y\|_1 \leq 1$$

$$\lim_{x \to x_0} I''(x,y,z) = I''(x_0,y,z) \quad \text{uniformemente si} \quad \|y\|_1 \leq 1, \|z\|_1 \leq 1.$$

Notemos que si x está en una vecindad de x_0 entonces $x_0 + \lambda(x - x_0)$ también está en la vecindad para $\lambda \in [0,1]$. Desarrollando por Taylor a la función

$$F(\lambda) = I(x_0 + \lambda(x - x_0))$$

alrededor de $\lambda = 0$ obtenemos:

$$I(x_0 + \lambda(x - x_0)) = I(x_0) + I'(x_0, x - x_0)\lambda + \frac{1}{2} I''(x_0, x - x_0)\lambda^2 + R(x_0, x - x_0)$$

y tomando $\lambda = 1$,

$$I(x) = I(x_0) + I'(x_0, x - x_0) + R_1(x_0, x - x_0)$$

$$I(x) = I(x_0) + I'(x_0, x - x_0) + \frac{1}{2} I''(x_0, x - x_0) + R_2(x_0, x - x_0).$$

Por lo anterior tenemos

$$R_1(x_0, x - x_0) = I(x) - I(x_0) - I'(x_0, x - x_0) =$$

$$\int_0^1 (I'(x_0 + \theta(x - x_0), x - x_0) - I'(x_0, x - x_0))d\theta =$$

$$\int_0^1 (1 - \theta) I''(x_0 + \theta(x - x_0), x - x_0)d\theta.$$

De la misma manera:

$$R_2(x_0, x - x_0) = \int_0^1 (1 - \theta)(I''(x_0 + \theta(x - x_0), x - x_0) - I''(x_0, x - x_0))d\theta.$$

Por lo tanto

$$\lim_{x \to x_0} \frac{R_1(x_0, x - x_0)}{\|x - x_0\|} = 0, \qquad \lim_{x \to x_0} \frac{R_2(x_0, x - x_0)}{\|x - x_0\|^2} = 0.$$

Así deducimos que $I'(x_0, y)$ y $I''(x_0, y)$ son la primera y segunda derivadas respecto a la norma $\| \cdot \|_1$ de I en $x = x_0$. En vista de lo anterior obtenemos:

Teorema 2.1.8. *Sea x_0 admisible y supongamos que existe $\delta > 0$ tal que*

$$I'(x_0, y) = 0, \quad I''(x, y) > 0$$

$\forall y \neq 0$ en X_0 y $\forall x \in X_e(A)$ tal que $\|x - x_0\|_1 < \delta$. Entonces $I(x) > I(x_0)$ para toda $x \neq x_0$ en $X_e(A)$ con $\|x - x_0\|_1 < \delta$.

2.2. Ecuaciones de Euler-Lagrange

Sabemos, por lo llevado a cabo, que para obtener un posible mínimo la primera variación se debe anular, de tal manera debemos conocer cómo encontrar estas trayectorias. Los siguientes resultados nos resuelven este problema.

Lema 2.2.1. *Dada una trayectoria admisible x, existe una única trayectoria $z \in X_0$ tal que*

$$I'(x,y) = \langle y, z\rangle_1 \qquad \forall y \in X_0 \tag{2.2.1}$$

con $\langle y,z\rangle_1 = \int_a^b \langle \dot{y}(t), \dot{z}(t)\rangle dt$*, donde* $z = (z_1(t), \dots, z_n(t))$ *satisface*

$$z(a) = 0, \quad \dot{z}(t) = L_{\dot{x}}(x(t)) - \textstyle\int_a^t L_x(x(s))ds - c, \qquad i = 1, \dots, n$$

con $c \in \mathbb{R}^n$ *constante tal que satisface* $z(b) = 0$.

Demostración. Como $z \in X_0$, fijándonos en una componente z_i tenemos

$$0 = z_i(t)|_a^b = \int_a^b \dot{z}_i(t)dt = \int_a^b L_{\dot{x}_i}dt - \int_a^b \left(\int_a^t L_{x_i}ds\right)dt - c_i\int_a^b dt$$

por lo tanto tenemos

$$c_i = \frac{1}{b-a}\left(\int_a^b L_{\dot{x}_i}dt - \int_a^b\left(\int_a^t L_{x_i}ds\right)dt\right).$$

Sea $y \in X_0$ entonces

$$0 = \int_a^b \frac{d}{dt}\left(y_i\left(\int_a^t L_{x_i}ds + c_i\right)\right)dt = \int_a^b\left(\dot{y}_i\left(\int_a^t L_{x_i}ds + c_i\right) + y_iL_{x_i}\right)dt$$

$$= \int_a^b \left(\dot{y}_i(L_{\dot{x}_i} - \dot{z}_i) + y_iL_{x_i}\right)dt = \int_a^b (y_iL_{x_i} + \dot{y}_iL_{\dot{x}_i})dt - \int_a^b \dot{y}_i\dot{z}_i dt$$

Por lo tanto

$$\langle y,z\rangle_1 = \sum_{i=1}^n \int_a^b \dot{y}_i\dot{z}_i dt = I'(x,y).$$

Veamos que z es único. Supongamos que existen z_1 y z_2 que satisfacen la ecuación (2.2.1), entonces

$$0 = \langle y, z_1\rangle_1 - \langle y, z_2\rangle_1 = \langle y, z_1 - z_2\rangle_1$$

Tomando $y = z_1 - z_2 \in X_0$ tenemos

$$\int_a^b |\dot{y}(t)|^2dt = 0 \Rightarrow \dot{y}(t) = 0 \Rightarrow y(t) = k = cte \qquad \forall t \in [a,b]$$

pero $y(a) = 0$ esto implica que $y(t) = 0$ por lo tanto $z_1(t) = z_2(t)$. □

Teorema 2.2.2. *Sea x_0 en $X_e(A)$ entonces*

$$I'(x_0, y) = 0 \quad \forall y \in X_0 \Leftrightarrow \exists c \in \mathbb{R}^n \text{ tal que } L_{\dot{x}}(x_0(t)) = \int_a^t L_x(x_0(s))ds + c.$$

Demostración. $\underline{\Rightarrow}|$ Sea $z \in X_0$ relacionado a x_0 por el Lema 2.2.1 entonces

$$I'(x_0, z) = \langle z, z \rangle_1 = \int_a^b \mid \dot{z}(t) \mid^2 dt$$

pero $\forall y \in X_0$ tenemos

$$I'(x_0, y) = \langle y, z \rangle_1 = 0 \Rightarrow \dot{z}(t) = 0.$$

Por el Lema 2.2.1 resulta

$$L_{\dot{x}}(x_0(t)) = \int_a^t L_x(x_0(s))ds + c. \tag{2.2.2}$$

$\underline{\Leftarrow}|$ Para todo $y \in X_0$ se cumple

$$0 = \int_a^b \frac{d}{dt} \langle y(t), \int_a^t L_x(x_0(s))ds + c \rangle dt =$$

$$\int_a^b \langle y(t), L_x(x_0(t)) \rangle + \langle \dot{y}(t), \int_a^t L_x(x_0(s))ds + c \rangle dt$$

Usando (2.2.2) tenemos

$$0 = \int_a^b (L_x(x_0(t))y(t) + L_{\dot{x}}(x_0(t))\dot{y}(t))dt = I'(x_0, y).$$

□

Las ecuaciones anteriores son la forma integral de las ecuaciones de Euler-Lagrange, es decir de

$$\frac{d}{dt}\left(\frac{\partial L}{\partial \dot{x}}\right) - \frac{\partial L}{\partial x} = 0. \tag{2.2.3}$$

De tal forma que para hallar un candidato a mínimo es indispensable encontrar x_0 tal que anule la primera variación, es decir, que satisfaga las ecuaciones (2.2.2) ó (2.2.3) dependiendo del espacio en que estemos trabajando.

A lo largo del trabajo se introducirán conjuntos los cuales nos ayudarán a indicar que una trayectoria $x(t)$ satisface ciertas propiedades.

Definición 2.2.3.

$$\mathcal{E} := \{x \in X \mid I'(x, y) = 0 \ \forall y \in X_0\}.$$

A los elementos de $\mathcal{E}$ les denominaremos extremales.

Ejemplo 2.2.4. *Sea una curva en* $\mathbb{R}^2$ *(el mismo análisis se puede extender a* $\mathbb{R}^n$*) definida por* $f : [a, b] \to \mathbb{R}^2$ *continua. Si* $f \in \mathcal{C}^1([a, b], \mathbb{R}^2)$ *entonces la curva determinada por* f *es rectificable y su longitud es igual a*

$$\mathcal{L} = \int_a^b \|f'(t)\|\, dt = \int_a^b \sqrt{f_1'(t)^2 + f_2'(t)^2}\mathrm{d}t.$$

Dados dos puntos (t_1, x_1)*,* $(t_2, x_2) \in \mathbb{R}^2$*, ¿cuál es la curva* $f : [a, b] \to \mathbb{R}^2$ *tal que* $f(a) = (t_1, x_1)$*,* $f(b) = (t_2, x_2)$ *con longitud mínima? Nuestra curva esta paramétrizada por* $(t, x(t))$*; entonces queremos* $x = x(t)$ *tal que minimice*

$$I(x) = \int_{t_1}^{t_2} \sqrt{1 + \dot{x}(t)^2}\mathrm{d}t$$

cumpliendo además las condiciones $x(t_1) = x_1$*,* $x(t_2) = x_2$*. Las ecuaciones de Euler-Lagrange en este caso se reducen a*

$$\frac{\dot{x}}{\sqrt{1 + (\dot{x})^2}} = C,$$

La solución a esta ecuación es de la forma

$$x(t) = \frac{C}{\sqrt{1 - C^2}}t + C_2$$

con $C, C_2 \in \mathbb{R}$ *constantes. A partir de las condiciones iniciales obtenemos que la curva que las satisface es*

$$x(t) = \frac{x_1 - t_2}{t_1 - t_2}t + \frac{t_1 x_2 - x_1 t_2}{t_1 - t_2},$$

que es la ecuación de la recta euclidea que une a los puntos; por lo tanto es un extremal y un probable mínimo. Más adelante se probará que en efecto es un mínimo.

Nota 4. *En física existe la formulación Lagrangina de la mecánica, la cual se sustenta en la teoría del cálculo variacional.*

Supongamos que una partícula se desplaza en una dimensión, la trayectoria está descrita por una curva $x = x(t)$*; y se mueve en una región en la que existe un potencial* $V = V(x)$ *el cual sólo depende la posición. El Principio de Maupertuis (o de mínima acción) establece que la trayectoria que seguirá el sistema es la que minimiza el funcional acción.*

$$J(x) = \int_a^b \left(\frac{m}{2}(\dot{x})^2 - V(x)\right) dt$$

donde m *representa la masa de la partícula, el primer sumando del integrando, la energía cinética y el segundo la potencial.*

Las ecuaciones de Euler-Lagrange se reducen a:

$$m\ddot{x} = -\frac{dV}{dx} = F \tag{2.2.4}$$

donde $F = -\frac{dV}{dx}$ *es la fuerza originada por el potencial* $V(x)$*. Se observa que las ecuaciones* (2.2.4) *son las ecuaciones de Newton.*

2.3. Función de Weierstrass

Hasta aquí hemos visto cómo obtener extremales los cuales son candidatos a mínimos, por lo que ahora necesitamos saber cuándo la segunda variación es no negativa. La siguiente función servirá para este propósito.

Definición 2.3.1. *A la función* $E : T \times \mathbb{R}^{3n} \to \mathbb{R}$ *definida por*

$$E(t, x, \dot{x}, u) = L(t, x, u) - L(t, x, \dot{x}) - \langle u - \dot{x}, L_{\dot{x}}(t, x, \dot{x})\rangle$$

se le conoce como función exceso de Weierstrass o la E-función de Weierstrass.

En el siguiente resultado se mostrará la utilidad de esta función.

Teorema 2.3.2. *(Condición de Weierstrass) Si* x_0 *minimiza a* I *en* $X_e(A)$ *entonces*

$$E(t, x_0(t), \dot{x}_0(t), u) \geq 0$$

para toda $(t, x_0(t), u) \in A$*.*

Demostración. Sea $\bar{t} \in (a, b)$ tal que no sea esquina de $x_0(t)$, es decir, $\bar{t}$ está en un intervalo de continuidad de $\dot{x}_0$. Sea $\overline{x}_i = x_0^i(\bar{t})$, $\overline{u}_i = \dot{x}_0^i(\bar{t})$ y sea u_i de manera que $(\bar{t}, \overline{x}, u)$ sea admisible. Tomemos $\delta_0 > 0$ que satisfaga $\bar{t} + \delta_0 < b$. Para $0 \leq \delta \leq \delta_0$ y $0 < \epsilon < 1$, definamos $x(t, \epsilon, \delta)$ de la siguiente forma:

$$x(t, \epsilon, \delta) = \begin{cases} x_0(t) & t \in [a, \bar{t}] \cup [\bar{t} + \delta, b] \\ x_0(t) + (t - \bar{t})(u - \overline{u}) & t \in [\bar{t}, \bar{t} + \epsilon\delta] \\ x_0(t) + \frac{\epsilon}{1+\epsilon}(u - \overline{u})(\bar{t} + \delta - t) & t \in [\bar{t} + \epsilon\delta, \bar{t} + \delta] \end{cases}$$

la cual es continua en $[a, b]$. Como A es abierto entonces podemos escoger ϵ_0 tal que $0 < \epsilon < \epsilon_0$ y $0 < \delta < \delta_0$ de manera que la trayectoria $x(t, \epsilon, \delta)$ sea admisible, por lo que

$$0 \leq I(x(t, \epsilon, \delta)) - I(x_0) = F(\epsilon, \delta) + G(\epsilon, \delta) \tag{2.3.1}$$

donde

$$F(\epsilon, \delta) = \int_{\bar{t}}^{\bar{t}+\epsilon\delta} (L(t, x(t, \epsilon, \delta), \dot{x}_0(t) + u - \overline{u}) - L(t, x_0(t), \dot{x}_0(t)))dt$$

$$G(\epsilon, \delta) = \int_{\bar{t}+\epsilon\delta}^{\bar{t}+\delta} (L(t, x(t, \epsilon, \delta), \dot{x}_0(t) - \frac{\epsilon}{1 - \epsilon}(u - \overline{u})) - L(t, x_0(t), \dot{x}_0(t)))dt.$$

Sea $\widehat{t}$ tal que $\bar{t} \leq \widehat{t} \leq \bar{t} + \epsilon\delta$, aplicando el teorema de valor medio para integrales obtenemos

$$\lim_{\delta\to 0} \frac{F(\epsilon,\delta)}{\epsilon\delta} = \lim_{\delta\to 0} \frac{(L(\widehat{t}, x(\widehat{t},\epsilon,\delta), \dot{x}_0(\widehat{t}) + u(\widehat{t}) - \dot{u}(\widehat{t})) - L(\widehat{t}, x(\widehat{t}), \dot{x}(\widehat{t}))\epsilon\delta}{\epsilon\delta} =$$

$$L(\bar{t}, \overline{x}, u) - L(\bar{t}, \overline{x}, \overline{u})$$

$$\lim_{\delta\to 0} \frac{G(\epsilon,\delta)}{\epsilon\delta} = \frac{1-\epsilon}{\epsilon}\left(L(\bar{t}, \overline{x}, \overline{u} - \frac{\epsilon}{1-\epsilon}(u - \overline{u})) - L(\bar{t}, \overline{x}, \overline{u})\right).$$

Por (2.3.1) tenemos la desigualdad

$$L(\bar{t}, \overline{x}, u) - L(\bar{t}, \overline{x}, \overline{u}) +$$

$$\frac{1-\epsilon}{\epsilon}\left(L(\bar{t}, \overline{x}, \overline{u} - \frac{\epsilon}{1-\epsilon}(u - \overline{u})) - L(\bar{t}, \overline{x}, \overline{u})\right) \geq 0$$

Utilizando el teorema de valor medio en la tercera entrada de L tenemos

$$L(\bar{t}, \overline{x}, u) - L(\bar{t}, \overline{x}, \overline{u}) +$$

$$\left(L_u(\bar{t}, \overline{x}, \overline{u} - \frac{\epsilon}{1-\epsilon}(u - \overline{u})) - L(\bar{t}, \overline{x}, \overline{u})\right)(u - \overline{u}) \geq 0$$

Tomando el límite cuando ϵ tiende a cero nos resulta

$$0 \leq E(\bar{t}, \overline{x}, \overline{u}, u).$$

Por lo tanto $E \geq 0$ para todos los puntos excepto en los extremos o esquinas pero por continuidad esta desigualdad se mantiene para todo $(t, x, \dot{x}, u)$. □

Definición 2.3.3.

$$\mathcal{W}(A) := \{x \in X(A) \mid E(t, x(t), \dot{x}(t), u) \geq 0 \ \forall (t, x(t), u) \in A\}$$

A partir de la condición de Weierstrass se obtienen resultados con varias aplicaciones.

Corolario 2.3.4. *(Condición de esquina) Si x_0 satisface las ecuaciones de Euler-Lagrange, la condición de Weierstrass y $L \in C^1$ entonces las funciones*

$$L - L_{\dot{x}}\dot{x}, \quad L_{\dot{x}}$$

son continuas a lo largo de x_0 y como consecuencia en cada esquina de x_0.

Demostración. Sea

$$p(t) = L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)),$$

como en x_0 se satisfacen las ecuaciones de Euler-Lagrange entonces $p(t) = \int_a^t L_x(x_0(t)) + c$, por lo cual $p(t)$ es continua ya que es suma de dos funciones continuas. A esta condición se le conoce generalmente como la condición esquina de Weierstrass-Erdmann.

Claramente la continuidad de la primera función se sigue de la condición de Weierstrass y la continuidad de p. Para probar el resultado, sea

$$F(t,u) = L(t,x_0(t),u) - \langle p(t),u\rangle.$$

Por la definición de la función de Weierstrass y de la continuidad de p tenemos

$$0 \leq E(t,x_0(t),\dot{x}_0(t-0),\dot{x}_0(t+0)) = F(t,\dot{x}_0(t+0)) - F(t,\dot{x}_0(t-0)),$$

$$0 \leq E(t,x_0(t),\dot{x}_0(t+0),\dot{x}_0(t-0)) = F(t,\dot{x}_0(t-0)) - F(t,\dot{x}_0(t+0)).$$

Por lo anterior resulta $F(t,\dot{x}_0(t+0)) = F(t,\dot{x}_0(t-0))$. □

Corolario 2.3.5. *En un punto esquina de una trayectoria admisible x_0 que satisface la condición de Weierstrass se satisface*

$$E(t,x_0(t),\dot{x}_0(t-0),\dot{x}_0(t+0)) = E(t,x_0(t),\dot{x}_0(t+0),\dot{x}_0(t-0)) = 0.$$

Corolario 2.3.6. *(Condición de Legendre) Si $x \in \mathcal{W}(A)$ entonces*

$$\langle \pi, L_{\dot{x}\dot{x}}(t,x(t),\dot{x}(t))\pi\rangle \geq 0 \qquad \forall \pi = (\pi_1,...,\pi_n) \in \mathbb{R}^n \tag{2.3.2}$$

Demostración. Sea $t \in [a,b]$ fijo y definamos $G(u) = E(t,x(t),\dot{x}(t),u)$, por lo que $G(\dot{x}(t)) = 0$ y

$$G_u(u) = E_u = L_{\dot{x}}(t,x(t),u) - L_{\dot{x}}(t,x(t),\dot{x}(t)),$$

$$G_{uu}(u) = L_{\dot{x}\dot{x}}(t,x(t),u).$$

Como $G(u) \geq 0$, $\dot{x}(t)$ es un mínimo local por lo que

$$0 \leq \langle \pi, G_{uu}(\dot{x})\pi\rangle = \langle \pi, L_{\dot{x}\dot{x}}(t,x(t),\dot{x}(t))\pi\rangle$$

para $\pi \in \mathbb{R}^n$. □

Definición 2.3.7. *Utilizaremos la condición de Legendre más adelante por lo tanto definiremos el siguiente conjunto:*

$$\mathcal{L} := \{x \in X \mid L_{\dot{x}\dot{x}}(x(t)) \geq 0 \ \forall t \in T\}.$$

Definición 2.3.8. *Diremos que $x \in X$ satisface la condición de Legendre reforzada si la desigualdad (2.3.2) es estricta, es decir si pertenece al conjunto*

$$\mathcal{L}' := \{x \in X \mid L_{\dot{x}\dot{x}}(x(t)) > 0 \ \forall t \in T\}.$$

Definición 2.3.9. *Sea $x \in X$. Si $|L_{\dot{x}\dot{x}}(t,x(t),\dot{x}(t))| \neq 0 \ \forall t \in T$ diremos que x es no singular.*

Notemos que x satisface la condición de Legendre reforzada si y solo si x es no singular y satisface la condición de Legendre.

Definición 2.3.10. *Se dice que L es positivamente regular en A si A es convexo en $\dot{x}$ y $L_{\dot{x}\dot{x}}(t,x,\dot{x}) > 0$ $\forall (t,x,\dot{x}) \in A$.*

Lo anterior nos indica que para (t,x) fija la función $L(t,x,\cdot)$ es estrictamente convexa.

Proposición 2.3.11. *Sea $L \in C^2$ positivamente regular en A. Entonces:*
a. *$E(t,x,\dot{x},u) > 0$, $\forall (t,x,\dot{x}) \in A$, $(t,x,u) \in A$ con $\dot{x} \neq u$.*
b. *Si $x \in X(A) \cap \mathcal{E}$ entonces $x \in C^1$.*

Demostración.
(a): Desarrollando por Taylor a $L(t,x,u)$ alrededor de $\dot{x}$, resulta:

$$L(t,x,u) = L(t,x,\dot{x}) + L_{\dot{x}}(t,x,\dot{x})(u-\dot{x}) +$$
$$\frac{1}{2}\langle u-\dot{x}, L_{\dot{x}\dot{x}}(t,x,\lambda\dot{x}+(1-\lambda)u)(u-\dot{x})\rangle$$

para $\lambda \in (0,1)$. Como A es convexo entonces $\lambda\dot{x}+(1-\lambda)u \in A$ y

$$E(t,x,\dot{x},u) = L(t,x,u) - L(t,x,\dot{x}) - L_{\dot{x}}(t,x,\dot{x})(u-\dot{x}) =$$
$$\frac{1}{2}\langle u-\dot{x}, L_{\dot{x}\dot{x}}(t,x,\lambda\dot{x}+(1-\lambda)u)(u-\dot{x})\rangle > 0$$

ya que $L_{\dot{x}\dot{x}} > 0$.
(b): Por el Corolario 2.3.5, tenemos que en las esquinas de $\dot{x}(t)$ la función de Weierstrass es igual a cero, pero el inciso (a) nos dice que es estrictamente positiva, por lo tanto $x \in C^1$. □

Como nos interesa determinar cuándo tenemos un mínimo local débil o fuerte introduciremos los siguientes conjuntos:

$$T_0(x;\epsilon) := \{(t,y) \in (T,\mathbb{R}^n) : |x(t)-y| < \epsilon\},$$
$$T_1(x;\epsilon) := \{(t,y,v) \in T_0(x;\epsilon) \times \mathbb{R}^n : |\dot{x}(t)-v| < \epsilon\}.$$

Definición 2.3.12. *Las trayectorias que satisfacen la condición reforzada de Weierstrass corresponden a los elementos del conjunto*

$$\mathcal{W}(A,\epsilon) := \{x_0 \in X(A) \mid E(t,x,\dot{x},u) \geq 0 \ \forall (t,x,\dot{x}) \in T_1(x_0,\epsilon),\ (t,x,u) \in A\}.$$

Proposición 2.3.13. *Supongamos que $L \in C^2$ y x_0 es una trayectoria no singular en $\mathcal{W}(A,\epsilon)$ para alguna $\epsilon > 0$. Entonces ϵ se puede disminuir de tal manera que la desigualdad en la definición de $\mathcal{W}(A,\epsilon)$ sea estricta.*

Demostración. Por continuidad podemos tomar $\epsilon_0 < \epsilon$ tal que $|L_{\dot{x}\dot{x}}(t,x,\dot{x})| \neq 0$ $\forall (t,x,\dot{x}) \in T_1(x_0,\epsilon_0)$. Supongamos que la conclusión del teorema es falsa con respecto a ϵ_0. Entonces existe $(t,x,\dot{x},u)$ con $(t,x,\dot{x}) \in T_1(x_0;\epsilon_0)$, $(t,x,u) \in A$ y $u \neq \dot{x}$, tal que $E(t,x,\dot{x},u) = 0$. Sea $f(v) := E(t,x,v,u)$ para toda $v \in \mathbb{R}^n$. Entonces f tiene un mínimo local en $v = \dot{x}$ y por lo tanto

$$0 = f'(\dot{x}) = -L_{\dot{x}\dot{x}}(t,x,\dot{x})(u-\dot{x})$$

por lo que tenemos que $u = \dot{x}$, pero esta es una contradicción. □

Ejemplo 2.3.14. *En el ejemplo 2.2.4 queremos encontrar la curva más corta entre dos puntos del plano* $\mathbb{R}^2$*; en este problema aparece el Lagrangiano*

$$L(x, y, y') = \sqrt{1 + (y')^2}.$$

Este Lagrangiano es convexo, aplicando la Proposición 2.3.11 tenemos que el segmento de recta que une dos puntos (el único extremal de este problema) es un mínimo.

2.4. Condición de Jacobi

Ya tenemos algunas condiciones necesarias que debe de satisfacer $x \in X(A)$ para que sea un mínimo, x debe pertenecer a $\mathcal{E} \cap \mathcal{W} \cap \mathcal{L}$. Ahora desarrollaremos otra condición necesaria.

Recordemos que la segunda variación

$$I''(x_0, y) = \int_a^b 2\Omega(t, x(t), \dot{x}(t))dt$$

está dada por (2.1.1), además si x_0 minimiza a I entonces

$$I''(x_0, y) \geq 0 \qquad \forall y \in X_0.$$

Esta condición equivale a decir que $y(t) = 0$ minimiza $I''(x_0, y)$ en X_0. El problema de minimizar a la segunda variación sobre X_0 se llama el problema del mínimo auxiliar.

Supongamos que x es no singular y $x \in C^1$. Sea

$$J(y) = \int_a^b 2\Omega(t, y(t), \dot{y}(t))dt$$

el cual es una forma cuadrática en y; su correspondiente forma bilineal es

$$J(y, z) = \int_a^b (\Omega_y z + \Omega_{\dot{y}} \dot{z})dt.$$

Definición 2.4.1. *Sea* $\mathcal{E}_x$ *el conjunto de trayectorias tales que satisfacen la condición de Euler-Lagrange para* Ω*, es decir,*

$$\begin{aligned}\mathcal{E}_x &= \{y \in X \mid J'_x(y; z) = 0, \forall z \in X_0\} \\ &= \{y \in X \mid \exists c \in \mathbb{R}^n \text{ tal que } \Omega_{\dot{y}} = \int_a^t \Omega_y ds + c\}\end{aligned}$$

Notemos, que dado $x \in X$ tenemos

$$\Omega_y(t, y, \dot{y}) = L_{xx}(x(t))y + L_{x\dot{x}}(x(t))\dot{y}$$

$$\Omega_{\dot{y}}(t, y, \dot{y}) = L_{\dot{x}x}(x(t))y + L_{\dot{x}\dot{x}}(x(t))\dot{y}$$

Proposición 2.4.2. *Si $L \in C^2$ y $x \in X(A) \cap \mathcal{L}' \cap C^1$, entonces $\mathcal{E}_x \subset C^1$.*

Demostración. Es consecuencia de la Proposición 2.3.11 ya que $\Omega_{\dot{y}\dot{y}} = L_{\dot{x}\dot{x}}(x(t))$. □

Si pedimos que $x \in C^1 \cap X(A) \cap \mathcal{L}'$ entonces tenemos: $y \in \mathcal{E}_x$ si y sólo si se cumple

$$\frac{d}{dt}\Omega_{\dot{y}} = \Omega_y \tag{2.4.1}$$

Desarrollada esta ecuación es de la forma

$$\frac{d}{dt}\left(L_{\dot{x}x}(x(t))y(t) + L_{\dot{x}\dot{x}}(x(t))\dot{y}(t)\right) = L_{xx}(x(t))y(t) + L_{x\dot{x}}(x(t))\dot{y}(t)$$

Definición 2.4.3. *La ecuación (2.4.1) recibe el nombre de ecuación auxiliar de Jacobi y sus soluciones se denominan extremales auxiliares.*

Recordando el teorema de existencia y unicidad de soluciones de ecuaciones diferenciales ordinarias y usando el hecho de que $y(t)$ es solución de una ecuación diferencial de segundo grado inferimos el siguiente lema:

Lema 2.4.4. *Un extremal auxiliar y está únicamente determinado por sus valores $y(a)$, $\dot{y}(a)$. En particular si $y(a) = \dot{y}(a) = 0$ entonces $y(t) = 0$.*

Introduciremos los puntos conjugados con la finalidad de determinar el intervalo de tiempo donde nuestros extremales son mínimos.

Definición 2.4.5. *(Puntos Conjugados) Sea $c \in (a, b]$. Se dice que c es un punto conjugado de a en $x_0(t)$ si existe $y(t)$ extremal auxiliar tal que $y(a) = y(c) = 0$ con $y(t) \not\equiv 0$ para $t \in (a, c)$.*

Teorema 2.4.6. *(Condición de Jacobi) Supongamos $L \in C^2$ y $x \in X(A) \cap C^1 \cap \mathcal{L}'$, si la segunda variación es no negativa entonces x no tiene puntos conjugados de a en T.*

Demostración. Supongamos que existe un punto conjugado, es decir, $\exists c \in (a, b)$ tal que $y(a) = y(c) = 0$. Sea $z(t)$ definido como

$$z(t) = \begin{cases} y(t) & t \in [a, c] \\ 0 & t \in [c, b] \end{cases}$$

por lo que tenemos que $z(t) \in X_0$, ya que $y(t)$ es de clase C^1 por la proposición 2.4.2, usando (2.4.1) tenemos

$$J_x(z) = \int_a^c \left(\langle \Omega_y(t, z(t), \dot{z}(t)), z(t)\rangle + \langle \Omega_{\dot{y}}(t, z(t), \dot{z}(t)), \dot{z}(t)\rangle\right)dt =$$

$$\int_a^c \frac{d}{dt}\langle \Omega_{\dot{y}}(t, z(t), \dot{z}(t)), z(t)\rangle dt = 0$$

Como la segunda variación de x es no negativa, entonces z minimiza J_x en X_0. Ya que $x \in C^1$ entonces Ω, Ω_y y $\Omega_{\dot{y}}$ son continuas y $z \in \mathcal{E}_x$, pero $z(t) = \dot{z}(t) = 0$ en (c, b) por lo tanto $z(t) \equiv 0$ en T, sin embargo esto es una contradiccón. □

La importancia de los puntos conjugados es asegurar que $x(t) \in X(A) \cap C^1 \cap \mathcal{E} \cap \mathcal{L}'$ sea un mínimo para el funcional $I(x)$ en el intervalo $[a, b)$ si éste no contiene puntos conjugados.

Definición 2.4.7.

$$\mathcal{J} := \{x \in X \mid x \ \textit{no tiene puntos conjugados a a en} \ (a, b)\}$$

Al igual que las condiciones que hemos definido anteriormente la condición de Jacobi también tiene su versión reforzada.

Definición 2.4.8.

$$\mathcal{J}' := \{x \in X \mid x \ \textit{no tiene puntos conjugados a a en} \ (a, b]\}$$

Los siguientes lemas nos ayudan a determinar cuándo un punto $c \in (a, b]$ es un punto conjugado.

Lema 2.4.9. *Sea $x \in X$ y $y_1, \ldots, y_{2n} \in \mathcal{E}_x$ linealmente independientes. Definamos $\forall s, t \in T$,*

$$D(s,t) := \begin{vmatrix} y_1(s) & \ldots & y_{2n}(s) \\ y_1(t) & \ldots & y_{2n}(t) \end{vmatrix}.$$

Entonces tenemos:
a. *$s \in (a, b]$ es conjugado a a en $x \Leftrightarrow D(s, a) = 0$.*
b. *Existe $\delta > 0$ tal que $D(s,t) \neq 0$, $\forall s, t \in T$ con $0 <\mid s - t \mid< \delta$.*

Demostración.
(a): $\underline{\Rightarrow}|$ Sea $y \in \mathcal{E}_x$ tal que $y(a) = y(s) = 0$, $y(t) \neq 0$ en (a, s). Sean $\alpha_i \in \mathbb{R}$ tales que

$$y(t) = \sum_{i=1}^{2n} \alpha_i y_i(t).$$

Como $y \neq 0$ existen constantes α_i no todas cero. Ya que $y(a) = y(s) = 0$ se presenta el sistema de ecuaciones

$$0 = \sum_{i=1}^{2n} \alpha_i y_i(s)$$

$$0 = \sum_{i=1}^{2n} \alpha_i y_i(a)$$

por lo tanto $D(s, a) = 0$.
$\underline{\Leftarrow}|$ Si $D(s, a) = 0$ entonces existen $\alpha_i \in \mathbb{R}$ no todas cero tales que

$$0 = \sum_{i=1}^{2n} \alpha_i y_i(a) = \sum_{i=1}^{2n} \alpha_i y_i(s)$$

luego

$$y(t) := \sum_{i=1}^{2n} \alpha_i y_i(s)$$

pertenece a $\mathcal{E}_x$ y satisface $y(a) = y(s) = 0$, $y(t) \neq 0$ para $t \in (a,s)$.

(b): Por Taylor tenemos

$$y_i(s) - y_i(t) = (s-t)\int_0^1 \dot{y}_i(t + \lambda(s-t))d\lambda$$

por lo que

$$D(s,t) = \begin{vmatrix} y_1(s) - y_1(t) & \dots & y_{2n}(s) - y_{2n}(t) \\ y_1(t) & \dots & y_{2n}(t) \end{vmatrix} = (s-t)^n \Delta(s,t)$$

con

$$\Delta(s,t) = \begin{vmatrix} \int_0^1 \dot{y}_1(t + \lambda(s-t))d\lambda & \dots & \int_0^1 \dot{y}_{2n}(t + \lambda(s-t))d\lambda \\ y_1(t) & \dots & y_{2n}(t) \end{vmatrix}.$$

Como $y_1, \dots, y_{2n}$ son linealmente independientes se obtiene

$$\Delta(t,t) = \begin{vmatrix} \dot{y}_1(t) & \dots & \dot{y}_{2n}(t) \\ y_1(t) & \dots & y_{2n}(t) \end{vmatrix} \neq 0.$$

Además Δ es continua, entonces existe $\delta > 0$ tal que $\Delta(s,t) \neq 0, \forall\, s,t \in T$ con $\mid s - t \mid < \delta$. Por lo que $D(s,t) \neq 0$. □

Lema 2.4.10. *Sea $x \in X$ y supongamos que $z_1, \dots, z_n \in \mathcal{E}_x$ son linealmente independientes y se anulan en $t = a$. Sea*

$$\Delta(t) = |z_1(t) \dots z_n(t)|.$$

Entonces $s \in (a,b]$ es conjugado a $a \Leftrightarrow \Delta(s) = 0$.

Demostración. Sean $y_1, \dots y_n$ n extremales auxiliares tales que

$$|y_1(a) \dots y_n(a)| = 1.$$

Entonces tenemos

$$D(s,t_0) = \begin{vmatrix} y_1(s) & \dots & y_n(s) & z_1(s) & \dots & z_n(s) \\ y_1(a) & \dots & y_n(a) & 0 & \dots & 0 \end{vmatrix} = (-1)^n \Delta(s).$$

Usando el Lema 2.4.9 obtenemos el resultado. □

Los siguiente resultados nos serán de utilidad más adelante.

Lema 2.4.11. *Sea $x \in X$, si $y, z \in \mathcal{E}_x$ entonces existe $c \in R^n$ tal que*

$$\langle \Omega_{\dot{y}}(y(t)), z(t) \rangle - \langle \Omega_{\dot{y}}(z(t)), y(t) \rangle = c \quad (t \in [a,b]).$$

Demostración. Definamos, $\forall y, z \in X$,

$$K(y,z) := \int_a^b (\langle \Omega_y(y(t)), z(t)\rangle + \langle \Omega_{\dot{y}}(y(t)), \dot{z}(t)\rangle)dt.$$

Notemos que $K(y,z) = K(z,y)$. Si $y \in \mathcal{E}_x$ tenemos

$$K(y,z) = \int_a^b \frac{d}{dt}\langle \Omega_{\dot{y}}(y(t)), z(t)\rangle dt = \left.\langle \Omega_{\dot{y}}(y(t)), z(t)\rangle\right|_a^b$$

Si z también está en $\mathcal{E}_x$ podemos intercambiar los roles de y y z por lo tanto

$$K(y,z) - K(z,y) = \left.\langle \Omega_{\dot{y}}(y(t)), z(t)\rangle - \langle \Omega_{\dot{y}}(z(t)), y(t)\rangle\right|_a^b = 0.$$

Como b se puede tomar como cualquier punto en T tenemos el resultado. $\square$

Proposición 2.4.12. *Supongamos que $L \in C^2$, $x \in X(A) \cap C^1 \cap \mathcal{L}'$ y $(y,q) \in X \times X$. Entonces son equivalentes:*
 a. $y \in \mathcal{E}_x$ *y* $q(t) = L_{x\dot{x}}(x(t))y(t) + L_{\dot{x}\dot{x}}(x(t))\dot{y}(t)$.
 b. (y,q) *satisface para* $t \in T$

$$\begin{aligned} \dot{y}(t) &= A(t)y(t) + B(t)q(t) \\ \dot{q}(t) &= C(t)y(t) - A^T(t)q(t) \end{aligned} \tag{2.4.2}$$

donde

$$\begin{aligned} A(t) &= -L_{\dot{x}\dot{x}}^{-1}(x(t))L_{\dot{x}x}(x(t)) \\ B(t) &= L_{\dot{x}\dot{x}}^{-1}(x(t)) \\ C(t) &= L_{xx}(x(t)) - L_{x\dot{x}}(x(t))L_{\dot{x}\dot{x}}^{-1}(x(t))L_{\dot{x}x}(x(t)). \end{aligned}$$

Nota 5. *Para el caso en el que estemos trabajando en $\mathbb{R}^n$ con $n = 1$ la ecuación auxiliar de Jacobi es de la forma ([10]):*

$$-\frac{d}{dt}\left(P\frac{df}{dt}\right) + Qf = 0, \tag{2.4.3}$$

con

$$\begin{aligned} P &= L_{\dot{x}\dot{x}} \\ Q &= L_{xx} - \frac{d}{dt}(L_{\dot{x}x}) \end{aligned}$$

Además notemos que si x satisface la condición de Legendre entonces $P \geq 0$.

Ejemplo 2.4.13. *En modelos de gravedad cuántica aparece el siguiente funcional acción:*

$$J(x) = \frac{1}{2}\int \sqrt{g(t)}(g^{-1}(t)\dot{x}(t)^2 + \omega x^2(t))dt,$$

donde $g : \mathbb{R} \to \mathbb{R}$ *es una función positiva que actúa como métrica en* $\mathbb{R}$ *y* ω *es una constante. Tomaremos* $g(t) = e^t$ *por simplicidad. Entonces*

$$L(t, x, \dot{x}) = \frac{1}{2} e^{t/2} (e^{-t} (\dot{x})^2 + \omega x^2).$$

Las ecuaciones de Euler-Lagrange se reducen a

$$\ddot{x} - \frac{1}{2}\dot{x} - \omega e^t x = 0.$$

La solución a esta ecuación es

$$x(t) = k_1 \exp(2\sqrt{\omega} e^{t/2}) + k_2 \exp(-2\sqrt{\omega} e^{t/2}).$$

Recordando la nota 7 tenemos $P = e^{-t/2} > 0$. *La ecuación auxiliar de Jacobi (2.4.3) se reduce a*

$$\frac{d^2 f}{dt^2} - \frac{1}{2}\frac{df}{dt} - \omega e^t f = 0,$$

Si le imponemos condiciones iniciales $f(0) = 0$, $f'(0) = 1$ *tenemos que la solución a la ecuación de Jacobi es*

$$f(t) = \frac{1}{\sqrt{\omega}} \sinh(2\sqrt{\omega}(e^{t/2} - 1)).$$

que claramente no tiene puntos conjugados a $t = 0$. *Como consecuencia del Teorema 2.4.6 la solución* $x(t)$ *es un mínimo en el intervalo en el cual está definida.*

Capítulo 3

Condiciones de suficiencia

Hasta este punto hemos mostrado condiciones necesarias para que un extremal sea mínimo; en este apartado se proporcionarán condiciones suficientes para determinar tanto un mínimo débil como uno fuerte.

3.1. Resultados auxiliares

Veamos primero que el teorema clásico de suficiencia para mínimos débiles es una consecuencia inmediata del teorema clásico de suficiencia para mínimos fuertes.

Teorema 3.1.1. *Si $L \in C^2$ y $x_0 \in X_e(A) \cap C^1 \cap \mathcal{E} \cap \mathcal{J}' \cap \mathcal{L}' \cap \mathcal{W}(A, \epsilon)$ entonces x_0 es un mínimo fuerte.*

Teorema 3.1.2. *Si $L \in C^2$ y $x_0 \in X_e(A) \cap C^1 \cap \mathcal{E} \cap \mathcal{J}' \cap \mathcal{L}'$ entonces x_0 es un mínimo débil.*

Demostración. Como $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0 \ \forall t \in T$, existe $\epsilon > 0$ tal que

$$L_{\dot{x}\dot{x}}(t, x, \dot{x}) \geq 0 \ \forall (t, x, \dot{x}) \in T_1(x_0; \epsilon).$$

Si es necesario sea ϵ tal que $T_1(x_0; \epsilon) \subset A$. Sean $(t, x, \dot{x})$ y (t, x, u) en $T_1(x_0, \epsilon)$. Como para toda $\lambda \in [0, 1]$,

$$(t, x, \lambda u + (1 - \lambda)\dot{x}) = (t, x, \dot{x} + \lambda(u - \dot{x})) \in T_1(x_0, \epsilon)$$

se sigue del Teorema de Taylor que

$$E(t, x, \dot{x}, u) = \int_0^1 (1 - \lambda)\langle u - \dot{x}, L_{\dot{x}\dot{x}}(t, x, \dot{x} + \lambda(u - \dot{x}))(u - \dot{x})\rangle d\lambda \geq 0$$

y por lo tanto $x_0 \in \mathcal{W}(T_1(x_0, \epsilon); \epsilon)$. Por el Teorema 3.1.1, x_0 es un mínimo fuerte estricto con respecto a $T_1(x_0; \epsilon)$, es decir, existe $\delta > 0$ tal que $I(x_0) < I(x)$ para toda $x \neq x_0$ con $x \in X_e((T_1(x_0; \epsilon))$ y $\|x - x_0\|_0 < \delta$. Tomando $\mu = \text{mín}\{\epsilon, \delta\}$ tenemos que $I(x_0) < I(x) \ \forall x \neq x_0$ con $x \in X_e(A)$ y $\|x - x_0\|_1 < \mu$. $\square$

Prosigamos a desarrollar la maquinaria necesaria para demostrar el teorema 3.1.1.

Teorema 3.1.3. *Supongamos que $L \in C^2$ y $x_0 \in X(A) \cap C^1$ es no singular. Sean $p_0(t) := L_{\dot{x}}(x_0(t))$ y*

$$T_2(x_0;\epsilon) := \{(t,x,p) \in T_0(x_0,\epsilon) \times \mathbb{R}^n : |p - p_0(t)| < \epsilon\}.$$

Entonces existen $\epsilon > 0$ y una función $l : T_2(x_0;\epsilon) \to \mathbb{R}$ (de clase C^r si $L \in C^r$) tales que, para cualesquiera $x, p \in X$ con $(t, x(t), p(t)) \in T_2(x_0,\epsilon)$, si (x,p) satisface el sistema

$$\dot{x}(t) = l_p(t,x(t),p(t)), \quad \dot{p}(t) = -l_x(t,x(t),p(t)) \tag{3.1.1}$$

para toda $t \in T$ entonces $x \in \mathcal{E} \cap C^1$ y $p(t) = L_{\dot{x}}(x(t))$.

Recíprocamente, si $x, p \in X$ con $(t,x(t),p(t)) \in T_2(x_0;\epsilon)$ son tales que $(t,x(t),\dot{x}(t)) \in T_1(x_0;\epsilon) \cap A$ entonces $x(t) \in \mathcal{E} \cap C^1$ y $p(t) = L_{\dot{x}}(x(t))$ implican que (x,p) satisface el sistema (3.1.1).

Demostración. Definamos

$$G(t,x,p,v) := L_{\dot{x}}(t,x,v) - p, \quad \forall (t,x,v) \in A, \ p \in \mathbb{R}^n.$$

Notemos que

$$G(t,x_0(t),p_0(t),\dot{x}_0(t)) = L_{\dot{x}}(x_0(t)) - p_0(t) = 0.$$

Como x_0 es no singular tenemos

$$|G_v(t,x_0,p_0,x_0)| = |L_{\dot{x}\dot{x}}(x_0(t))| \neq 0.$$

Como $L_{\dot{x}}$ y $L_{\dot{x}\dot{x}}$ son continuas entonces también lo son G y G_v. Por el teorema de la función implícita existen $\delta, \rho > 0$ y $U : T_2(x_0;\rho) \to \mathbb{R}^n$ continua tales que

$$U(t,x_0(t),p_0(t)) = \dot{x}_0(t) \quad (t \in [a,b]),$$

$$p = L_{\dot{x}}(t,x,U(t,x,p)) \quad ((t,x,p) \in T_2(x_0;\rho)).$$

Además las relaciones

$$p = L_{\dot{x}}(t,x,v), \quad |v - U(t,x,p)| < \delta, \quad (t,x,p) \in T_2(x_0,\rho)$$

se satisfacen solo si $v = U(t,x,p)$. Como $L_{\dot{x}} \in C^r$ entonces $U \in C^r$. Definamos

$$l(t,x,p) := \langle p, U(t,x,p)\rangle - L(t,x,U(t,x,p)) \tag{3.1.2}$$

$\forall (t,x,p) \in T_2(x_0;\rho)$. Entonces

$$\begin{aligned} l_t(t,x,p) &= -L_t(t,x,U(t,x,p)) \\ l_x(t,x,p) &= -L_x(t,x,U(t,x,p)) \\ l_p(t,x,p) &= U(t,x,p). \end{aligned}$$

Como $L \in C^r$ tenemos que $l \in C^r$.

Sea $0 < \epsilon < \text{mín}\{\rho, \delta/2\}$ tal que, para toda $(t, x, p) \in T_2(x_0; \epsilon)$,

$$|U(t, x_0(t), p_0(t)) - U(t, x, p)| < \delta/2 \tag{3.1.3}$$

y sean $x, p \in X$ tales que $(t, x(t), p(t)) \in T_2(x_0; \epsilon)$. Entonces

$$\dot{x}(t) = U(t, x(t), p(t))$$

y por lo tanto

$$\dot{p}(t) = L_{\dot{x}}(t, x(t), U(t, x(t), p(t)) = L_{\dot{x}}(t, x(t), \dot{x}(t)).$$

Como $(t, x(t), p(t)) \in T_2(x_0; \epsilon) \subset T_2(x_0; \rho)$ entonces $p(t) = L_{\dot{x}}(x(t))$.

Para el recíproco, supongamos que $(t, x, \dot{x}) \in T_1(x_0; \epsilon) \cap A$. El resultado se sigue si probamos que $\dot{x}(t) = U(t, x(t), p(t))$, lo cual es consecuencia de

$$\begin{aligned} |\dot{x}(t) - U(t, x(t), p(t))| \quad &\leq \quad |\dot{x}(t) - \dot{x}_0(t)| \\ &\quad + |U(t, x_0(t), p_0(t)) - U(t, x(t), p(t))| < \delta. \end{aligned}$$

□

A (3.1.2) se le conoce como Hamiltoniano y al sistema (3.1.1) como las ecuaciones de Hamilton; se trabajarán con ellas en las siguientes secciones.

Supongamos que $L \in C^2$ y $x \in X(A) \cap C^1 \cap \mathcal{E}$ es no singular. Por el teorema 3.1.3 existe $\delta > 0$ tal que $\overline{x} : (a-\delta, b+\delta) \to \mathbb{R}^n$ satisface $\overline{x}(t) = x(t)$ para $t \in T$. Además $\overline{x}(t)$ también es extremal.

Definición 3.1.4. *A $\overline{x}$ le llamaremos extensión de x.*

Nota 6. *Supongamos $L \in C^2$ y $x_0 \in X(A) \cap C^1 \cap \mathcal{E}$ es no singular. Sea $a \in T$ fija. Sea (x_0, p_0) con $p_0 = L_{\dot{x}}(x_0(t))$ tales que satisfacen el sistema de ecuaciones (3.1.1) para algún $\epsilon > 0$. Por el teorema de encaje de soluciones de ecuaciones diferenciales (véase apéndice Teorema 5.2.8) existen $\rho > 0$ y F_0 vecindad de $\{(x_0(t), p_0(t)) \mid t \in T\}$ tales que para cada $(\alpha, \beta) \in F_0$ pasa una única solución $(x(\cdot; \alpha, \beta), p(\cdot; \alpha, \beta))\colon [a-\rho, b+\rho] \to \mathbb{R}^{2n}$ la cual cumple*

$$(x(t; x_0(a), p_0(a)), p(t; x_0(a), p_0(a))) = (x_0(t), p_0(t)),$$

para toda $t \in T$. Entonces tenemos que x_0 es un miembro de una familia $2n$-paramétrica de extremales $\{x(\cdot; \alpha, \beta)\}$ para valores $(\alpha, \beta) = (\alpha_0, \beta_0)$; además las funciones $x(t; \alpha, \beta)$, $p(t; \alpha, \beta) = L_{\dot{x}}(x(t; \alpha, \beta))$ y sus derivadas $\dot{x}(t; \alpha, \beta)$, $\dot{p}(t; \alpha, \beta)$ son de clase C^{m-1} si $L \in C^m$. También el determinante

$$\begin{vmatrix} x_\alpha(t; \alpha, \beta) & x_\beta(t; \alpha, \beta) \\ \dot{x}_\alpha(t; \alpha, \beta) & \dot{x}_\alpha(t; \alpha, \beta) \end{vmatrix}$$

es diferente de cero en x_0. Si tomamos $\alpha_0 = x_0(a)$ fijo, nos resulta:

Teorema 3.1.5. *Supongamos* $L \in C^2$ *y* $x_0 \in X(A) \cap \mathcal{E} \cap C^1$ *es no singular. Entonces* x_0 *es un miembro (para* $\beta = \beta_0$*) de una familia* n*-paramétrica de extremales* $\{x(\,\cdot\,;\beta)\}$ *que pasan a través de* x_0 *(o una extensión de este). Las funciones* $x(t;\beta)$ *y* $\dot{x}(t;\beta)$ *son de clase* C^{m-1} *si* $L \in C^m$ *(para* $m \geq 2$*) y la matriz*

$$\begin{pmatrix} x_\beta(t;\beta) \\ \dot{x}_\beta(t;\beta) \end{pmatrix}$$

tiene rango n *en cada punto de* x_0.

Nota 7. *Supongamos que* $x_0 \in \mathcal{E}$ *es miembro para* $\epsilon = 0$ *de una familia uniparamétrica de extremales* $\{x(\cdot;\epsilon)\}$ *con* $x(t;\epsilon)$ *y* $\dot{x}(t;\epsilon)$ *de clase* C^1*. Para toda* $t \in T$ *sea* $y(t) := x_\epsilon(t,0)$ *y nótese que, diferenciando la relación*

$$\frac{d}{dt}\left(L_{\dot{x}}(t, x(t,\epsilon), \dot{x}(t,\epsilon))\right) = L_x(t, x(t,\epsilon), \dot{x}(t,\epsilon))$$

con respecto a ϵ *y evaluando en* $\epsilon = 0$ *obtenemos*

$$\frac{d}{dt}\left(L_{\dot{x}x}(x_0(t))y(t) + L_{\dot{x}\dot{x}}(x_0(t))\dot{y}(t)\right) = L_{xx}(x_0(t))y(t) + L_{x\dot{x}}(x_0(t))\dot{y}(t),$$

es decir, $y \in E_x$.

Lema 3.1.6. *Sea* $x \in X(A) \cap C^1 \cap \mathcal{E}$ *y* $\overline{x}$ *extensión de* x*. Si* $x \in \mathcal{J}'$ *entonces existe* $a_0 < a$ *tal que no hay puntos conjugados a* a_0 *de* $\overline{x}$.

Demostración. Supongamos que $\overline{x}$ está definido en $[a-\epsilon, b+\epsilon]$. Sean $y_1, \ldots, y_{2n}$ en $E_{\overline{x}}$ linealmente independientes. Por el Lema 2.4.9 existe $\delta > 0$ tal que $D(s,t) \neq 0$ para $s,t \in [a-\epsilon, b+\epsilon]$ que cumplen $0 < |s-t| < \delta$. Sea $\delta < 2\epsilon$. Como $x \in \mathcal{J}'$ aplicando el Lema 2.4.9 tenemos $D(t,a) \neq 0$ para $t \in (a,b]$. Por lo tanto $D(t,a) \neq 0$ para $t \in [a+\delta/2, b]$. Por continuidad existe $a_0 \in [a-\delta/2, b)$ tal que $D(t,a_0) \neq 0$ para $t \in [a+\delta/2, b]$. Como $[a_0, a+\delta/2] \subset [a-\epsilon, b+\epsilon]$ y $t \in (a_0, a+\delta/2]$ implica que $0 < |t-a_0| < \delta$, también se sigue que $D(t,a_0) \neq 0$ para $t \in (a_0, a+\delta/2]$. El resultado se sigue del Lema 2.4.9. □

Teorema 3.1.7. *Si* $L \in C^2$ *y* $x \in X(A) \cap C^1 \cap \mathcal{E} \cap \mathcal{L}'$ *entonces son equivalentes:*

a. $I''(x,y) > 0\ \forall y \in X_0$.

b. x *no tiene puntos conjugados en* $(a,b]$.

c. *Existe* (Y,Q) *solución matricial del sistema (2.4.2) que satisface* $|Y(t)| \neq 0$ *y* $Y^T(t)Q(t) = Q^T(t)Y(t)$ *para toda* $t \in T$.

Demostración.

(a) $\Rightarrow$ (b). Por el Teorema 2.4.6 x satisface la condición de Jacobi. Por contradicción. Supongamos que (b) es falso. Entonces existe $y \in X_0 \cap \mathcal{E}_x$ no cero. Para y tenemos $J_x(y) = 0 = I'(x_0, y)$ lo cual contradice (a).

(b) $\Rightarrow$ (c). Sea $\overline{x}$ extensión de x. Por el Lema 3.1.6 existe $a_0 < a$ tal que no existen puntos conjugados a a_0 de $\overline{x}$. Sean $y_1, \ldots y_n \in E_{\overline{x}}$ linealmente independientes con $y_i(a_0) = 0$. Sea $Y := (y_1, \ldots, y_n)$ y $Q := (q_1, \ldots, q_n)$ donde $q_i(t) = \Omega_{\dot{y}}(y_i(t))$. Por el Lema 2.4.11 $Y(t)Q^T(t) - Q(t)Y^T(t)$ tiene los mismos valores para toda t y se anula en $t = a_0$, por lo tanto la expresión es igual a cero. Por el Lema 2.4.10 $|Y(t)| = 0$ solo si $t = a_0$ o t es conjugado de a_0. Por lo tanto $|Y(t)| \neq 0$ en T.

(c) $\Rightarrow$ (a). Sean $y \in X_0$, $w(t) := Y^{-1}(t)y(t)$ y $z(t) := Y(t)\dot{w}(t)$ por lo que $\dot{y}(t) = \dot{Y}(t)w(t) + z(t)$. Sean

$$\begin{aligned} Q(t) &= L_{\dot{x}x}(x(t))Y(t) + L_{\dot{x}\dot{x}}(x(t))\dot{Y}(t) \\ \dot{Q}(t) &= L_{xx}(x(t))Y(t) + L_{x\dot{x}}(x(t))\dot{Y}(t) \end{aligned}$$

de manera que

$$\begin{aligned} \Omega_y(y(t)) &= L_{xx}(x(t))y(t) + L_{x\dot{x}}(x(t))\dot{y}(t) = \\ & L_{xx}(x(t))Y(t)w(t) + L_{x\dot{x}}(x(t))(\dot{Y}(t)w(t) + z(t)) = \\ & \dot{Q}(t)w(t) + L_{\dot{x}x}(x(t))z(t), \end{aligned}$$

$$\begin{aligned} \Omega_{\dot{y}}(y(t)) &= L_{\dot{x}\dot{x}}(x(t))\dot{y}(t) + L_{\dot{x}x}(x(t))y(t) = \\ & L_{\dot{x}x}(x(t))Y(t)w(t) + L_{\dot{x}\dot{x}}(x(t))(\dot{Y}(t)w(t) + z(t)) = \\ & Q(t)w(t) + L_{\dot{x}\dot{x}}(x(t))z(t). \end{aligned}$$

De lo anterior obtenemos

$$\begin{aligned} 2\Omega(y(t)) &= \langle y(t), \Omega_y(y(t))\rangle + \langle \dot{y}(t), \Omega_{\dot{y}}(y(t))\rangle \\ &= \langle Y(t)w(t), \dot{Q}(t)w(t) + L_{x\dot{x}}(x(t))z(t)\rangle \\ &\qquad + \langle \dot{Y}(t)w(t) + z(t), Q(t)w(t) + L_{\dot{x}\dot{x}}(x(t))z(t)\rangle \\ &= \langle w(t), (\dot{Q}^T(t)Y(t) + Q^T(t)\dot{Y}(t))w(t)\rangle \\ &\qquad + \langle w(t), Q^T(t)Y(t)\dot{w}(t)\rangle + \langle z(t), L_{\dot{x}\dot{x}}(x(t))z(t)\rangle \\ &\qquad + \langle w(t), (Y(t)L_{x\dot{x}}(x(t)) + \dot{Y}(t)L_{\dot{x}\dot{x}}(x(t)))Y(t)\dot{w}(t)\rangle \end{aligned}$$

Ya que (Y, Q) satisfacen el sistema (2.4.2) usando la proposción 2.4.12 nos resulta

$$\begin{aligned} 2\Omega(y(t)) &= \langle w(t), (\dot{Q}^T(t)Y(t) + Q^T(t)\dot{Y}(t))w(t)\rangle \\ &\qquad + 2\langle w(t), Q^T(t)Y(t)\dot{w}(t)\rangle + \langle z(t), L_{\dot{x}\dot{x}}(x(t))z(t)\rangle \\ &= \frac{d}{dt}\langle w(t), Q^T(t)Y(t)w(t)\rangle + \langle z(t), L_{\dot{x}\dot{x}}(x(t))z(t)\rangle. \end{aligned}$$

Como $w(a) = w(b) = 0$, esto implica que

$$I''(x, y) = \int_a^b \langle z(t), L_{\dot{x}\dot{x}}(x(t))z(t)\rangle dt.$$

Por lo tanto $I''(x, y) > 0$ a menos que $z(t) = Y(t)\dot{w}(t) = 0$. Sin embargo, $|Y(t)| \neq 0$ y $w(a) = 0$ por lo que $I''(x, y) = 0$ implicaría que $w \equiv 0$ lo cual a su vez implicaría que $y \equiv 0$. $\square$

3.2. Campos de Mayer

Para la demostración del Teorema 3.1.1, el conjunto clásico de condiciones suficientes para un mínimo fuerte, utilizaremos el concepto de campos de Mayer.

Definición 3.2.1. *A una pareja* (Γ, M) *se le llama campo de Mayer sobre* A *si*

a. M *es una región en* $\mathbb{R} \times \mathbb{R}^n$.

b. $\Gamma\colon M \to \mathbb{R}^n$ *es de clase* C^1 *y* $(t, x, \Gamma(t, x)) \in A \ \forall (t, x) \in M$.

c. *La integral de línea (llamada integral de Hilbert):*

$$I^* := \int P(t, x)dx + Q(t, x)dt$$

es independiente de la trayectoria en M*, es decir,* $Pdx + Qdt$ *es una diferencial total, donde*

$$P(t, x) := L_{\dot{x}}(t, x, \Gamma(t, x)), \quad Q(t, x) := L(t, x, \Gamma(t, x)) - \langle \Gamma(t, x), P(t, x) \rangle.$$

Si (Γ, M) *es un campo de Mayer sobre* A*, a la solución de la ecuación diferencial*

$$\dot{x}(t) = \Gamma(t, x(t))$$

se le llama extremal del campo (Γ, M).

Nota 8. *Dado* (Γ, M) *un campo de Mayer sobre* A *y* x *una trayectoria en* M*, si definimos*

$$L^*(t, x, \dot{x}) := \langle P(t, x), \dot{x} \rangle + Q(t, x),$$

entonces la integral

$$I^*(x) = \int_a^b L^*(t, x(t), \dot{x}(t))dt$$

es minimizada por x *en la clase de trayectorias que unen sus puntos inicial y final. Suponiendo que* $L \in C^2$*, entonces* $L^* \in C^1(M \times \mathbb{R}^n)$ *y podemos aplicar la ecuación de Euler-Lagrange para el integrando* L^**, es decir, existe una constante* $c \in \mathbb{R}^n$ *tal que*

$$c = L^*_{\dot{x}} - \int_a^t L^*_x ds = L_{\dot{x}}(t, x(t), \Gamma(t, x(t))) -$$

$$\int_a^t \{L_x(s, x(s), \Gamma(s, x(s))) + P_x(s, x(s))(\dot{x}(s) - \Gamma(s, x(s)))\} ds.$$

Lo anterior implica que, si x *es un extremal del campo* (Γ, M)*, entonces* x *es un extremal de* L *y* $I^*(x) = I(x)$.

Proposición 3.2.2. *Supongamos que* $L \in C^2$ *y* (Γ, M) *es un campo de Mayer sobre* A *tal que*

$$E(t, x, \Gamma(t, x), \dot{x}) > 0$$

para toda $(t, x) \in M$ *y* $(t, x, \dot{x}) \in A$ *con* $\dot{x} \neq \Gamma(t, x)$*. Si* $x_0 \in X_e(A)$ *es un extremal de* (Γ, M) *entonces la desigualdad* $I(x_0) < I(x)$ *se satisface para toda* $x \neq x_0$ *con* $x \in X_e(A)$ *y* $(t, x(t)) \in M$.

Demostración. Si x está en M,

$$I(x) = I^*(x) + \int_a^b E(t, x(t), \Gamma(t, x(t)), \dot{x}(t))dt$$

ya que

$$E(t, x(t), \Gamma(t, x), \dot{x}) = L(t, x, \dot{x}) - L^*(t, x, \dot{x}).$$

Por lo tanto, si $x \in X_e(A)$ con $x \neq x_0$ y está en M, entonces

$$I^*(x) = I^*(x_0) = I(x_0).$$

Esto implica que $I(x) \geq I(x_0)$ y la igualdad solo se da si $\dot{x}(t) = \Gamma(t, x(t))$. En este caso x es un extremal del campo y coincide con x_0 ya que tienen el mismo punto inicial. □

Teorema 3.2.3. *Sean $L \in C^2$ y $x_0 \in X(A) \cap \mathcal{E} \cap \mathcal{L}' \cap \mathcal{J}' \cap C^1$. Entonces existe un campo de Mayer sobre A del cual x_0 es un extremal.*

Demostración. Por el Lema 3.1.6 existen $a_0 < a$ y $\overline{x}_0$ extensión de x_0 tales que no hay puntos conjugados a a_0 en $\overline{x}_0$. Por el Teorema 3.1.5 existe una familia n-paramétrica de extremales $\{x(\cdot, \beta)\}$ que pasan a través de a_0 ($x(a_0, \beta) = \overline{x}_0(a_0)$) y contienen a $\overline{x}_0$ para $\beta = \beta_0$ ($x(t, \beta_0) = \overline{x}_0(t)$). La primera de estas afirmaciones implica que, si $z_i(t) := x_{\beta_i}(t, \beta_0)$ entonces $z_i(a_0) = 0$. Como $z_1, \ldots, z_n$ son n extremales auxiliares linealmente independientes con respecto a $\overline{x}_0$ (por nota 7), por el Lema 2.4.10 tenemos que $|x_\beta(t, \beta_0)| \neq 0$ para toda $t \in T$.

Por el teorema de la función implícita existen $\epsilon, \mu > 0$ y $\lambda\colon T_0(x_0, \epsilon) \to \mathbb{R}^n$ tales que

$$\lambda(t, x_0(t)) = \beta_0, \quad x(t, \lambda(t, y)) = y \quad \forall (t, y) \in T_0(x_0, \epsilon),$$

y las relaciones

$$x(t, \beta) = y, \quad |\beta - \lambda(t, y)| < \mu, \quad (t, y) \in T_0(x_0, \epsilon)$$

se satisfacen solo si $\beta = \lambda(t, y)$. Además si $x(t, \lambda) \in C^m$ entonces $\lambda \in C^m$. Definamos $M := T_0(x_0, \epsilon)$ y $\Gamma(t, y) := \dot{x}(t, \lambda(t, y))$ para toda $(t, y) \in M$. Entonces $\Gamma \in C^1(M)$ y $(t, y, \Gamma(t, y)) \in A$ para todo $(t, y) \in M$. Por construcción x_0 es un extremal de (Γ, M). Falta ver que I^* es independiente de la trayectoria en M.

Sea $C \subset M$ cualquier trayectoria de clase C^1 parametrizada por

$$C = \{(r(s), w(s)) \mid s \in [s_0, s_1]\}.$$

Definamos $\forall s \in [s_0, s_1]$ y $t \in [a_0, r(s)]$,

$$u(t, s) := x(t, \lambda(r(s), w(s))),$$

$$F(s) := \int_{a_0}^{r(s)} L(t, u(t, s), \dot{u}(t, s))dt.$$

Notemos que

$$\frac{d}{dt}L_{\dot{x}}(t,u(t,s),\dot{u}(t,s)) = L_x(t,u(t,s),\dot{u}(t,s))$$

y $u_s(a_0,s) = 0$. También tenemos

$$w(s) = u(r(s),s), \quad \Gamma(r(s),w(s)) = \dot{u}(r(s),w(s))$$

y $\dot{w}(s) = \dot{u}(r(s),s)\dot{r}(s) + u_s(r(s),s)$. Por lo tanto,

$$\begin{aligned} F'(s) = \int_{a_0}^{r(s)} & \{\langle L_x(t,u(t,s),\dot{u}(t,s)), u_s(t,s)\rangle \\ & \qquad + \langle L_{\dot{x}}(t,u(t,s),\dot{u}(t,s)), \dot{u}_s(t,s)\rangle\}dt \\ & + L(r(s),u(r(s),s),\dot{u}(r(s),s))\dot{r}(s) \\ = \ & \langle L_{\dot{x}}(r(s),u(r(s),s),\dot{u}(r(s),s)), u_s(r(s),s)\rangle \\ & + L(r(s),u(r(s),s),\dot{u}(r(s),s))\dot{r}(s) \\ = \ & \langle P(r(s),w(s)),\dot{w}(s)\rangle + Q(r(s),w(s))\dot{r}(s). \end{aligned}$$

Por lo tanto,

$$I^*(C) = \int_{s_0}^{s_1} F'(s)ds = F(s_1) - F(s_0)$$

y concluimos que I^* es independiente de C. $\square$

Con este resultado tenemos lo necesario para demostrar el Teorema 3.1.1, el cual recordaremos y demostraremos a continuación.

Teorema 3.2.4. *Si $L \in C^2$ y $x_0 \in X_e(A) \cap C^1 \cap \mathcal{E} \cap \mathcal{J}' \cap \mathcal{L}' \cap \mathcal{W}(A,\epsilon)$ entonces x_0 es un mínimo fuerte.*

Demostración. Las hipótesis implican, por el Teorema 3.2.3, que x_0 es un extremal para algun campo de Mayer (Γ, M) en A. Por la Proposición 2.3.13, ϵ puede tomarse lo suficientemente pequeña para que la desigualdad en la definición de $\mathcal{W}(A,\epsilon)$ se vuelva estricta. Si es necesario también se puede disminuir M de tal manera que, para toda $(t,y) \in M$, $(t,y,\Gamma(t,y)) \in T_1(x,\epsilon)$. Por la Proposición 3.2.2, $I(y) > I(x_0)$ $\forall y \neq x_0$ con $y \in X_e(A)$ y $(t,y(t)) \in M$. $\square$

En resumen, al utilizar los resultados de los Teoremas 2.4.6 y 3.2.4 deducimos el siguiente teorema que nos proporciona condiciones suficientes para que un extremal sea un mínimo fuerte estricto o un mínimo débil estricto.

Teorema 3.2.5. *Supongamos que $L \in C^2$ y $x \in X_e(A) \cap C^1 \cap \mathcal{E} \cap \mathcal{J}' \cap \mathcal{L}'$. Entonces,*

a. *x es un mínimo débil estricto.*

b. *Si además $\exists \epsilon > 0$ tal que $x \in \mathcal{W}(A;\epsilon)$ entonces x es un mínimo fuerte estricto.*

Capítulo 4

Problemas con restricciones

Hasta aquí conocemos cómo encontrar mínimos tanto fuertes como débiles para el funcional

$$I(x) = \int_a^b L(t, x(t), \dot{x}(t))dt,$$

por lo que tenemos un resultado análogo al Teorema 1.1.1. En adelante nuestro problema estará sujeto a restricciones y obtendremos el teorema análogo en dimensión infinita del Teorema 1.1.2. Es importante mencionar que existen varios tipos de restricciones, mismas que se desarrollarán en las siguientes secciones.

4.1. Problemas isoperimétricos

A pesar de que nuestro interés es encontrar mínimos para el funcional I bajo restricciones del tipo (1.1.1) (problema clásico de Lagrange), primero se solucionará el problema con restricciones isoperimétricas con la finalidad de mostrar los diferentes tipos de problemas con los que nos podemos encontrar.

El propósito será minimizar el mismo funcional

$$I(x) = \int_a^b L(t, x(t), \dot{x}(t))dt$$

con $x \in X_e(A)$ que deben satisfacer

$$I_s(x) = B_s + \int_a^b L_s(t, x(t), \dot{x}(t))dt = 0, \qquad s = 1, \ldots, p < n$$

donde $L_s \in C^2$ y B_s son constantes. A este tipo de restricciones se les conoce como isoperimétricas.

Introduciremos el funcional

$$J(x) = \sum_{s=1}^{p} \lambda_s I_s(x)$$

donde λ_s son constantes no todas cero.

Definición 4.1.1. *Decimos que x_0 es extremaloide para J si y solo si su primera variación a lo largo de x_0 es cero en X_0, es decir, $\forall y \in X_0$ tenemos*

$$J'(x_0, y) = \sum_{s=1}^{p} \lambda_s I_s'(x_0, y) = 0.$$

Definición 4.1.2. *Sea $x_0 \in X$ diremos que es normal si no es un extremaloide de J.*

Por álgebra lineal obtenemos el siguiente resultado.

Lema 4.1.3. *Sea $I_s'(x_0, y)$ la primera variación de I_s a lo largo de x_0. Entonces x_0 es normal si y solo si los funcionales lineales $I_s'(x_0, y)$, $s = 1, \ldots, p$ son linealmente independientes en X_0, es decir, x_0 es normal si y solo si existen $y_1, \ldots y_p \in X_0$ tales que*

$$|I_s'(x_0, y_q)| \neq 0$$

para $s, q = 1, \ldots p$.

Lema 4.1.4. *Sea $y \in X_0$ y x_0 normal tales que satisfacen*

$$I_s'(x_0, y) = 0$$

para $s = 1, \ldots, p$. Entonces existe una familia uniparamétrica

$$x(\epsilon) : x(t, \epsilon) \qquad t \in [a, b], |\epsilon| < \epsilon_0$$

en $X_e(A)$ tal que contiene a x_0 para $\epsilon = 0$. Además $x_\epsilon(t, \epsilon), x_{\epsilon\epsilon}(t, \epsilon)$ son continuas y definen variaciones en X_0,

$$y(t) = x_\epsilon(t, 0),$$

es decir, y es la variación de la familia $x(\epsilon)$ a lo largo de x_0.

Demostración. Sean $y_1, \ldots, y_p \in X_0$ tales que cumplen el Lema 4.1.3, $y \in X_0$ tal que $I_s'(x_0, y) = 0$. Sean

$$I_s(x_0 + \sum_{q=1}^{p} c_q y_q + \epsilon y) = 0 \qquad q, s = 1, \ldots, p \tag{4.1.1}$$

con solución inicial $(c, \epsilon) = (0, 0)$. Usando el Lema 4.1.3 y el teorema de la función implícita tenemos que las ecuaciones (4.1.1) tienen soluciones de la forma $c_q(\epsilon)$ para $q = 1, \ldots, p$ y $|\epsilon| < \epsilon_0$, las cuales son de clase C^2 y satisfacen $c_q(0) = 0$. Al derivar (4.1.1) respecto a ϵ y evaluando en $\epsilon = 0$ tenemos

$$I_s'(x_0, y_q)c_q'(0) + I_s'(x_0, y) = I_s'(x_0, y_q)c_q'(0) = 0.$$

Por el Lema 4.1.3 $c_q'(0) = 0$, por lo tanto $x(\epsilon) = x_0 + \sum_{q=1}^{p} c_q y_q + \epsilon y$ tiene las propiedades requeridas. □

Teorema 4.1.5. *Supongamos que x_0 es solución del problema y es normal. Entonces existe un único conjunto de multiplicadores $\lambda_1, \ldots, \lambda_p$ tales que para*

$$J = I + \lambda_1 I_1 + \cdots + \lambda_p I_p$$

se satisface $J'(x_0, y) = 0$ $\forall y \in X_0$. Además $J''(x_0, y) \geq 0$ para $y \in X_0$ que cumplen

$$I_s'(x_0, y) = 0 \qquad s = 1, \ldots, p. \tag{4.1.2}$$

Demostración. Sean $y \in X_0$ tal que cumple (4.1.2), $x(\epsilon)$ dado por el Lema 4.1.4 y $W(\epsilon) = I(x(\epsilon))$. Entonces $W(\epsilon)$ tiene un mínimo local en $\epsilon = 0$, por lo tanto

$$W'(0) = I'(x_0, y) = 0.$$

Por el Teorema 1.1.2 existen $\lambda_1, \ldots, \lambda_p$ tales que

$$I'(x_0, y) + \lambda_1 I_1'(x_0, y) + \cdots + \lambda_p I_p'(x_0, y) = 0$$

en X_0. Por lo tanto

$$J'(x_0, y) = I'(x_0, y) + \sum_{s=1}^{p} \lambda_s I_s'(x_0, y) = 0$$

en X_0. Como $I_s(x(\epsilon)) = 0$, obtenemos

$$W(\epsilon) = J(x(\epsilon)) = I(x(\epsilon)) + \sum_{s=1}^{p} \lambda_s I_s(x(\epsilon)).$$

Ya que $\epsilon = 0$ minimiza localmente a $W(\epsilon)$ tenemos

$$W'(\epsilon) = J'(x(\epsilon), x_\epsilon(\epsilon))$$

$$0 \leq W''(0) = J''(x_0, y) + J'(x_0, x_{\epsilon\epsilon}) = J''(x_0, y),$$

ya que $x_{\epsilon\epsilon} \in X_0$. □

Teorema 4.1.6. *Si x_0 es solución del problema y es normal entonces existen $\lambda_1, \ldots, \lambda_p$ únicos tales que satisfacen*

$$F_{\dot{x}} = \int_a^t F_x ds + c \tag{4.1.3}$$

a lo largo de x_0, donde c es constante y $F = L + \sum_{i=1}^{p} \lambda_i L_i$. Además

$$E(t, x_0(t), \dot{x}_0(t), u) \geq 0$$

para $(t, x_0(t), u) \in A$, donde E es la función de Weierstrass de F.

Demostración. Como x_0 minimiza a J tenemos $J'(x_0, y) = 0$, por lo tanto satisface las ecuaciones de Euler-Lagrange por lo que se cumple (4.1.3).

A continuación probaremos la condición de Weierstrass. Por continuidad basta demostrarlo para $\bar{t}$ cualquier punto que no defina un punto esquina o un extremo de x_0; sean $\overline{x} = x_0(\bar{t})$ y $\overline{u} = \dot{x}_0(\bar{t})$. Definimos a $y(\epsilon) = y(t, \epsilon) \in X_0$ como

$$y(t,\epsilon) = \begin{cases} 0 & t \in [a, \bar{t}] \\ (u - \overline{u})(t - \bar{t}) & t \in [\bar{t}, \bar{t} + \epsilon] \\ \dfrac{\epsilon(u - \overline{u})(b - t)}{b - \bar{t} - \epsilon} & t \in [\bar{t} + \epsilon, b] \end{cases}$$

Sean $y_1, \ldots, y_p \in X_0$ tales que satisfacen el Lema 4.1.3. Por este mismo lema tenemos que las ecuaciones

$$I_s\left(x_0 + \sum_{q=1}^{p} c_q y_q + y(\epsilon)\right) = 0$$

tienen soluciones $c_q(\epsilon) \in C^1$ en un intervalo $[0, \epsilon_1]$, tales que $c_q(0) = 0$ y

$$x(\epsilon) = x_0 + \sum_{q=1}^{p} c_q y_q + y(\epsilon)$$

es admisible. La función $W(\epsilon) = J(x(\epsilon))$ tiene un mínimo en $\epsilon = 0$. Como $\epsilon \geq 0$ tenemos que $W'(0) \geq 0$. Nótese que

$$W'(0) = J'(x_0, y_q)c_q'(0) + G'(0) = G'(0)$$

donde

$$\begin{aligned} G(\epsilon) &= J(x_0 + y(\epsilon)) \\ &= \text{constante } + \int_{\bar{t}}^{\bar{t}+\epsilon} F(t, x_0(t) + (u - \overline{u})(t - \bar{t}), \dot{x}_0(t) + u - \overline{u})dt \\ &\quad + \int_{\bar{t}+\epsilon}^{b} F(t, x_0(t) + \phi(\epsilon)z(t), \dot{x}_0(t) + \phi(\epsilon)\dot{z}(t))dt \end{aligned}$$

donde

$$\phi(\epsilon) = \frac{\epsilon(b - \bar{t})}{b - \bar{t} + \epsilon}, \qquad z(t) = \frac{(u - \overline{u})(b - t)}{b - \bar{t}}.$$

Por la definición resulta que $\phi(0) = 0$, $\phi'(0) = 1$, por lo tanto

$$G'(0) = F(\bar{t}, \overline{x}, u) - F(\bar{t}, \overline{x}, \overline{u}) + \int_{\bar{t}}^{b} (F_x z + F_{\dot{x}} \dot{z})dt$$

Por (4.1.3), el integrando es la derivada de $F_{\dot{x}} z$ respecto de t a lo largo de x_0. Como $z(\bar{t}) = u - \overline{u}$ y $z(b) = 0$ entonces

$$G'(0) = F(\bar{t}, \overline{x}, u) - F(\bar{t}, \overline{x}, \overline{u}) + F_{\dot{x}} z|_{\bar{t}}^{b} = E(\bar{t}, \overline{x}, \overline{u}, u).$$

Además como $W'(0) = G'(0) \geq 0$ obtenemos que $E \geq 0$. $\square$

Ejemplo 4.1.7. *Sea Z una variable aleatoria y $\rho(x)$ su función de densidad, $\rho : \mathbb{R} \to (0, \infty)$. Supongamos que conocemos el segundo momento*

$$\sigma^2 = \int_{\mathbb{R}} x^2 \rho(x) \mathrm{d}x.$$

Buscamos obtener la función de densidad con menor sesgo $\rho(x)$. Dicho problema se puede ver como el problema de encontrar los máximos de la función de entropía

$$S(\rho) = -\int_{\mathbb{R}} \rho(x) \log \rho(x) \mathrm{d}x,$$

sujeta a las restricciones

$$\begin{aligned} \int_{\mathbb{R}} \rho(x) \mathrm{d}x &= 1\,, \\ \int_{\mathbb{R}} x^2 \rho(x) \mathrm{d}x &= \sigma^2\,. \end{aligned} \tag{4.1.4}$$

Notemos que el Lagrangiano está definido sobre $U = \mathbb{R} \times (0, +\infty) \times \mathbb{R}$. Tenemos tres funcionales:

$$\begin{aligned} S(\rho) &= -\int_{\mathbb{R}} \rho(x) \log \rho(x) \mathrm{d}x, \\ K_1(\rho) &= \int_{\mathbb{R}} \rho(x) \mathrm{d}x, \\ K_2(\rho) &= \int_{\mathbb{R}} x^2 \rho(x) \mathrm{d}x. \end{aligned}$$

Sus primeras variaciones son:

$$\begin{aligned} S'(\rho, h) &= -\int_{\mathbb{R}} h(x)(1 + \log \rho(x)) \mathrm{d}x, \\ K_1'(\rho, h) &= \int_{\mathbb{R}} h(x) \mathrm{d}x, \\ K_2'(\rho, h) &= \int_{\mathbb{R}} x^2 h(x) \mathrm{d}x. \end{aligned}$$

Cualquier trayectoria es normal ya que si suponemos

$$\det \begin{pmatrix} \int_{\mathbb{R}} h_1(x) \mathrm{d}x & \int_{\mathbb{R}} h_2(x) \mathrm{d}x \\ \int_{\mathbb{R}} x^2 h_1(x) \mathrm{d}x & \int_{\mathbb{R}} x^2 h_2(x) \mathrm{d}x \end{pmatrix} = 0$$

entonces para $h_1, h_2 \in X$ arbitrarios tenemos

$$\frac{\int_{\mathbb{R}} x^2 h_1(x) \mathrm{d}x}{\int_{\mathbb{R}} h_1(x)} = \frac{\int_{\mathbb{R}} x^2 h_2(x) \mathrm{d}x}{\int_{\mathbb{R}} h_2(x)},$$

lo cual es absurdo. Por el Teorema 4.1.5 podemos introducir multiplicadores de Lagrange (λ_1, λ_2) y consideramos:

$$L(x, \rho, \rho') = -\rho(x) \log \rho(x) + \lambda_1 \rho(x) + \lambda_2 x^2 \rho(x).$$

Las ecuaciones de Euler-Lagrange son

$$-\log \rho(x) - 1 + \lambda_1 + \lambda_2 x^2 = 0.$$

con solución

$$\rho(x) = e^{-1+\lambda_1+\lambda_2 x^2}$$

y si la sustituimos en las restricciones (4.1.4) *se puede ver que los multiplicadores de Lagrange son*

$$\lambda_1 = 1 + \log \frac{1}{\sqrt{2\pi}\sigma}, \;\; \lambda_2 = -1/(2\sigma^2),$$

respectivamente. Entonces, la solución a la ecuación de Euler-Lagrange es

$$\rho(x) = (\sqrt{2\pi}\,\sigma)^{-1} \exp\left(-x^2/(2\sigma^2)\right)$$

la cual es la distribución normal de probabilidad con varianza σ^2. El Lagrangiano

$$L_0(x, \rho, \rho') = -\rho(x) \log \rho(x)$$

puede verse como una función real de una variable sobre $(0, +\infty)$. *Además L_0 es cóncava. Al aplicar la Proposición 2.3.11 a $-L_0$ tenemos que $\rho(x)$ es un mínimo global para $-L_0$ por lo tanto la solución obtenida es un máximo global en* $(0, +\infty)$.

4.2. Problema con restricciones tipo Lagrange

En esta sección se desarrollará parcialmente la teoría de control óptimo tomando en cuenta tres tipos de restricciones y se enunciará el principio del máximo (uno de los pilares del control óptimo) a partir del cual se deducirán resultados necesarios para establecer la teoría de aumentabilidad en el caso de dimensión infinita. El concepto de aumentabilidad en el cual vamos a trabajar es el que tiene restricciones de tipo Lagrange.

Tenemos nuestro funcional

$$I(x) = \int_a^b L(t, x(t), \dot{x}(t))dt$$

con $x \in X_e(A)$ sujeto a restricciones

$$\varphi_i(t, x, \dot{x}) = 0, \;\; i = 1, \ldots, m < n \tag{4.2.1}$$

$\varphi\colon T \times \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ de clase C^2. Nuestro interés es encontrar a x que minimice el funcional I sujeto a las restricciones (4.2.1). A este tipo de restricciones se les conoce como restricciones de tipo Lagrange.

Si x resuelve el problema entonces tiene que pertenecer al conjunto

$$D := \{x \in X_e(A) \mid \varphi(t, x(t), \dot{x}(t)) = 0\}. \tag{4.2.2}$$

Supondremos que la matriz

$$\left(\frac{\partial \varphi_i}{\partial \dot{x}_j}\right)$$

evaluada en x_0 tiene rango máximo en D.

Hay dos formulaciones para abordar el problema: la hamiltoniana y la lagrangiana. De la primera, la cual se derivará del principio máximo, se deduce la lagrangiana. Antes de abordar el principio del máximo necesitaremos introducir nueva herramienta de trabajo.

4.3. Conos tangentes, conos derivados, conjuntos derivados y resultados técnicos

Definición 4.3.1. *Sean V un subconjunto de $\mathbb{R}^p$ y $v_0 \in V$. Si K_0 consta de vectores $k \in V$ tales que existen sucesiones $(v_n)_{n\in\mathbb{N}} \subset V$ y $(c_n)_{n\in\mathbb{N}} \subset \mathbb{R}^+$ que satisfacen*

$$\lim_{n\to\infty} c_n = 0, \qquad k = \lim_{n\to\infty} \frac{v_n - v_0}{c_n}$$

entonces K_0 es el cono tangente de V en v_0.

Definición 4.3.2. *Recordemos que un conjunto de N vectores $k_1, \ldots, k_N$ genera un cono de vectores k si*

$$k = \sum_{i=1}^{N} \alpha_i k_i$$

para $\alpha_i \geq 0$, $i = 1, \ldots, N$. Un cono generado de esa manera es un cono diferencial de V en v_0 si existen $\delta > 0$ y una función continua

$$v(\epsilon) = v_0 + \sum_{i=1}^{N} \epsilon_i k_i + r(\epsilon)$$

en V con $0 \leq \epsilon_i \leq \delta$ para $i = 1, \ldots, N$, tal que

$$\lim_{\epsilon\to 0} \frac{r(\epsilon)}{|\epsilon|} = 0$$

donde $|\epsilon| = \epsilon_1 + \cdots + \epsilon_N$. Nótese que $\sum_{i=1}^{N} k_i d\epsilon_i$ es la diferencial de v en $\epsilon = 0$.

Definición 4.3.3. *K es un conjunto derivado de V en v_0 si es una clase de vectores tales que para cada conjunto finito de vectores $k_1, \ldots, k_N \in K$ genera un cono diferencial de V en v_0.*

Definición 4.3.4. *Un cono derivado K^* es un conjunto derivado de V que es un cono convexo.*

El cono K^* generado por un conjunto derivado K es el conjunto de combinaciones $k = \sum_{i=1}^{N} \alpha_i k_i$ con $\alpha_i \geq 0$ donde $k_i \in K$. Por definición K^* es convexo.

Sean $J_i(x)$ $(i = 0, \ldots, p)$ funcionales que satisfacen

$$\begin{aligned} J_i(x) &\leq 0 \qquad i = 1, \ldots, p' \\ J_i(x) &= 0 \qquad i = p' + 1, \ldots, p. \end{aligned} \tag{4.3.1}$$

Sea S el conjunto de vectores de la base canónica $e_1, \ldots, e_{p'}$ y de los vectores $-e_i$ tales que $J_i(x_0) < 0$, para $1 \leq i \leq p'$.

Los siguientes teoremas se utilizarán en secciones posteriores para demostrar resultados que nos interesan; para su demostración véase [8].

Teorema 4.3.5. *Sea x_0 que cumple (4.3.1), K un conjunto derivado de J_p en x_0 y K^+ el cono convexo generado por los vectores en K y los vectores en S para x_0. Supongamos que x_0 minimiza J_0 sujeto a las restricciones (4.3.1). Entonces el cono dual positivo Λ de K^+ es no nulo, por lo que existen multiplicadores $\lambda_0, \lambda_1, \ldots, \lambda_p$, no todos ceros, tal que la desigualdad*

$$L(k) = \sum_{i=1}^{p} \lambda_i k_i \geq 0$$

se cumple para todos los vectores $k \in K$ y por consiguiente para todos los vectores en la cerradura $\overline{K}^*$ del cono convexo K^* generado por K. Además $\lambda_i \geq 0$ para $i = 0, \ldots, p'$ y $\lambda_i = 0$ si $i \geq 1$ y $J_i(x_0) < 0$. Se puede tomar $\lambda_0 = 1$ si y solo si $(-1, 0, \ldots, 0) \notin \overline{K^+}$.

Para el siguiente resultado se trabajará con el funcional

$$J(u) = \int_a^b F(t, u(t))dt.$$

Teorema 4.3.6. *Sea $u_0(t) \in PWC([a, b])$, R_δ la clase de elementos (t, u) de la forma $(t, u_0(t + h))$ con $t \in [a, b]$, $t + h \in [a, b]$ y $|h| < \delta$. Sea $F(t, u)$ continua en R_δ tal que $F_t(t, u)$ también lo es. Si*

$$F(t, u) \geq F(t, u_0(t)) \tag{4.3.2}$$

para $(t, u) \in R_\delta$ entonces

$$F(t, u_0(t)) = \int_a^t F_t(s, u_0(s))ds + c$$

en $[a, b]$ donce c es una constante.

Demostración. Por definición $(t+h, u_0(t)), (t, u_0(t+h)) \in R_\delta$. Por la desigualdad (4.3.2) tenemos

$$F(t, u_0(t + h)) \geq F(t, u_0(t)), \qquad F(t + h, u_0(t)) \geq F(t + h, u_0(t + h)).$$

Tomando límites derechos e izquierdos respecto a h

$$F(t, u_0(t+0)) \geq F(t, u_0(t)) \geq F(t, u_0(t+0)),$$

$$F(t, u_0(t-0)) \geq F(t, u_0(t)) \geq F(t, u_0(t-0)),$$

por lo cual F es continua en T. Sean t y $t+h$ puntos en un intervalo de continuidad de u_0, definamos

$$\phi(h) = F(t+h, u_0(t+h)) - F(t, u_0(t)),$$

entonces

$$F(t+h, u_0(t+h)) - F(t, u_0(t+h)) \leq \phi(h) \leq F(t+h, u_0(t)) - F(t, u_0(t)).$$

Por el teorema del valor medio

$$F_t(t+\delta_1 h, u_0(t+h)) \leq \frac{\phi(h)}{h} \leq F_t(t+\delta_2 h, u_0(t)),$$

con $\delta_1, \delta_2 \in (0,1)$. Haciendo $h \to 0$ tenemos

$$\frac{d}{dt}F(t, u_0(t)) = \lim_{h\to 0} \frac{\phi(h)}{h} = F_t(t, u_0(t))$$

por lo cual $F(t, u_0(t))$ tiene derivada continua a pedazos y como $F(t, u_0(t))$ es continua obtenemos el resultado para $c = F(a, u_0(a))$. $\square$

Sean $f, \varphi_1, \ldots, \varphi_n$ funciones de clase C^1 en una región R de puntos $(t,u) \in \mathbb{R} \times \mathbb{R}^q$. Sea $R_0 \subset R$ el conjunto de puntos donde se satisface

$$\varphi_s(t,u) \leq 0 \ (s = 1, \ldots, m'), \quad \varphi_s(t,u) = 0 \ (s = m'+1, \ldots, m)$$

y supongamos que la matriz

$$\begin{pmatrix} \dfrac{\partial \varphi_s}{\partial u_k} & \delta_{s\beta}\varphi_\beta \end{pmatrix}$$

(donde δ es la delta de Kronecker y $\beta = 1, \ldots, m'$; $s = 1, \ldots, m$; $k = 1, \ldots, q$) es de rango m en R. Decimos que un elemento (t,u) es admisible si pertenece a R_0. Una función u es admisible si es continua a pedazos en T y $(t, u(t))$ es admisible.

Teorema 4.3.7. *Sea u_0 una función admisible tal que*

$$f(t,u) \geq f(t, u_0(t))$$

para (t,u) admisible. Entonces existen $\mu_s(t)$ $(s = 1, \ldots, m)$ multiplicadores únicos tales que para

$$F(t,u,\mu) = f + \sum_{s=1}^{m} \mu_s \varphi_s$$

satisfacen

$$F_{u_k}(t, u_0(t), \mu(t)) = 0.$$

Además para $s \leq m'$, $\mu_s(t) \geq 0$ *en* $[a, b]$ *y* $\mu_s(t) = 0$ *cuando* $\varphi_s(t, u_0(t)) < 0$. *Los multiplicadores* $\mu_s(t) \in PWC([a, b])$ *son continuos donde* $u_0(t)$ *también lo es y existe* c *constante tal que*

$$F(t, u_0(t), \mu(t)) = \int_a^t F_t(r, u_0(r), \mu(r))dr + c. \tag{4.3.3}$$

Si f, φ_s *son de clase* C^{p+1} *entonces* $\mu_s(t) \in C^p$ *en donde* $u_0(t) \in C^p$.

Para el siguiente resultado, sea $R \subset \mathbb{R}^n \times \mathbb{R}^q$ una región. Sean $\varphi_s(x, u)$ con $s = 1, \dots, m$ de clase C^1 en R. Sea R_0 el conjunto de (x, u) donde se satisfacen $\varphi_s(x, u) = 0$. Supondremos que la matriz

$$\left(\frac{\partial \varphi_s(x, u)}{\partial u_k} \right)$$

con $s = 1, \dots, m$ y $k = 1, \dots, q$, es de rango m en R_0.

Lema 4.3.8. *Existen* q *funciones continuas* $U_k(x, u)$ *en una vecindad* $R_1 \subset R_0$ *tales que* $(x, U(x, u)) \in R_0$ *para todo* $(x, u) \in R_1$, *tales que* $U_k(x, u) = u_k$ *cuando* $(x, u) \in R_0$. *Si las funciones* $\varphi_s \in C^p(R)$ *entonces* U *se puede escoger de manera que* $U \in C^p$ *en* R_1.

4.4. Principio del máximo

Para resolver el problema de minimizar a I bajo las restricciones del tipo (4.2.1) pasaremos a trabajar en el área del control óptimo.

El principio del máximo nos muestra cómo optimizar problemas de control óptimo. Se estudiarán tres tipos de restricciones: el problema de control de punto fijo, el problema de control isoperimétrico y el problema de control tipo Lagrange.

Antes de describir cada uno de estos problemas se introducirán elementos necesarios para su estudio.

Sean $x \in X(A)$ y $u(t) \in \mathbb{R}^q$, tal que $u \in PWC([a, b])$ con $(t, x, u) \in A$. A u le llamaremos variable de control y a x variable de estado (en ciertas ocasiones al tiempo t también se le considera como variable de estado). Para cada problema queremos minimizar el funcional

$$I(x, u) = \int_a^b L(t, x(t), u(t))dt \tag{4.4.1}$$

sobre las parejas (x, u) con $x(a) = \xi_1$ y $x(b) = \xi_2$. Notemos que si cambiamos a u por $\dot{x}$ el problema se transforma en uno de cálculo variacional. Los diferentes tipos de problema son:

Definición 4.4.1. *(Problema de control de puntos fijos) Al problema de minimizar al funcional (4.4.1) bajo las restricciones*

$$\dot{x}(t) = f(t, x(t), u(t)) \tag{4.4.2}$$

con $f\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}$ *(de clase* C^2*) se le conoce como problema de control de puntos fijos.*

Definición 4.4.2. *(Problema de control isoperimétrico) Al problema de minimizar el funcional (4.4.1), tal que* x *cumpla*

$$I_i(x) \leq 0 \ (i = 1, \dots, p'), \quad I_i(x) = 0 \ (i = p' + 1, \dots, p) \tag{4.4.3}$$

con

$$I_i(x) = g_i + \int_a^b L_i(t, x, u)dt$$

donde g_i *son constantes (y* L_i *son de clase* C^2*) se le conoce como problema de control isoperimétrico.*

Definición 4.4.3. *(Problema de Lagrange) Al problema de minimizar el funcional (4.4.1) tal que* x *satisfaga*

$$\varphi_i(t, x, u) \leq 0 \ (i = 1, \dots, m'), \quad \varphi_i(t, x, u) = 0 \ (i = m' + 1, \dots, m) \tag{4.4.4}$$

donde $\varphi_i\colon \mathbb{R} \times \mathbb{R}^n \times \mathbb{R}^q \to \mathbb{R}$ *se le conoce como problema de tipo Lagrange.*

De la definición anterior podemos observar que el caso que nos interesa es cuando $m' = 0$.

En ocasiones se pueden resolver las ecuaciones (4.4.4) y expresarlas como restricciones del tipo (4.4.2); inversamente podemos resolver las restricciones del problema clásico de Lagrange (restricciones con igualdades de (4.4.4)) si eliminamos u_i en (4.4.2).

Con el fin de desarrollar nuestro objetivo se requiere que las variaciones admisibles $(t, x(t), u(t))$ se encuentren en el conjunto $R_0 \subset A$ que tiene la propiedad de que, para cada elemento $(t, x, u) \in R_0$, existe $\delta > 0$ tal que si (t_1, x_1, u_1) se encuentra en una δ-vecindad de (t, x, u) entonces $(t', x', u_1) \in R_0$ $\forall (t', x')$ en la δ-vecindad de (t, x). También pediremos que u sea acotada en R_0.

Primero desarrollaremos el principio del máximo para problemas de control isoperimétrico y de puntos fijos.

Teorema 4.4.4. *(Principio del máximo) Sea nuestro funcional*

$$I(x, u) = \int_a^b L(t, x(t), u(t))dt$$

Sean x_0, u_0 *tales que minimizan a* I *sobre el conjunto de variaciones admisibles que satisfacen las condiciones (4.4.2) y (4.4.3). Entonces existen multiplicadores*

$$\lambda_0 \geq 0, \ \lambda_i, \ p_s(t) \quad i = 1, \dots, p; \ s = 1, \dots, n$$

no todos son cero en $t \in [a,b]$ tales que para

$$H(t,x,u,p) = \sum_{i=1}^{n} p_i f_i - \lambda_0 L - \sum_{i=1}^{p} \lambda_i L_i \tag{4.4.5}$$

tenemos:

a. *$\lambda_i \geq 0$ para $i = 1, \ldots, p'$, y $\lambda_i = 0$ si $I_i(x_0) < 0$.*

b. *Los multiplicadores $p_i(t)$ son continuos en $[a,b]$ y en x_0 satisfacen*

$$\begin{cases} \dot{x}_i = \quad H_{p_i} = f_i \\ \dot{p}_i = -H_{x_i} \end{cases} \tag{4.4.6}$$

en cada intervalo donde $u_0(t)$ es continua.

c. *Se cumple*

$$H(t, x_0(t), u, p(t)) \leq H(t, x_0(t), u_0(t), p(t)) \tag{4.4.7}$$

para $t \in [a,b]$ y u tales que $(t, x_0, u) \in R_0$.

d. *La función $H(t, x_0(t), u_0(t), p(t))$ es continua en $[a,b]$ y en x_0 satisface*

$$\frac{dH}{dt} = H_t \tag{4.4.8}$$

en cada intervalo de continuidad de $u_0(t)$.

e. *Si R_0 es abierto entonces para $t \in T$ se satisface*

$$H_u(t, x_0(t), u_0(t), p(t)) = 0. \tag{4.4.9}$$

Demostración. Para probar (a), (b) y (c) supongamos que para cada $(\overline{t}, \overline{x}, \overline{u}) \in R_0$ existe $\delta > 0$ y u continua en $[\overline{t} - \delta, \overline{t} + \delta]$ tal que $u(\overline{t}) = \overline{u}$ y $(t, x, u(t)) \in R_0$ para $|t - \overline{t}| \leq \delta$ y $|x - \overline{x}| \leq \delta$. Además f_i, L, L_i y sus derivadas con respecto a x_i son continuas. Sea $x_0 : (x_0(t), u_0(t))$ en T solución. Definamos $I_0 = I$, $g_0 = 0$, $L_0 = L$,

$$A_j^i(t) = \frac{\partial f_i}{\partial x_j}, \ \ M_j^k(t) = \frac{\partial L_k}{\partial x_j} \tag{4.4.10}$$

evaluadas en (x_0, u_0), con $i, j = 1, \ldots, n$, $k = 0, \ldots, p$. Escojamos soluciones $q_j^k(t)$ y $Z_j^i(t)$ de los sistemas

$$\begin{aligned} \dot{q}_j^k + \sum_{i=1}^{n} q_i^k A_j^i = M_j^k; \quad q_j^k(b) = 0 \\ \dot{Z}_j^i + \sum_{h=1}^{n} Z_h^i A_j^h = 0; \quad Z_j^i(b) = \delta_{ij} \end{aligned} \tag{4.4.11}$$

para $k = 0, \ldots, p$; $h, i, j = 1, \ldots, n$. Sean

$$
\begin{aligned}
F_k(t,x,u) &= L_k - \sum_{j=1}^{n} \left(q_j^k f_j - \dot{q}_j^k x_j \right), \\
F_{p+i}(t,x,u) &= \sum_{j=1}^{n} \left(Z_j^i f_j + \dot{Z}_j^i x_j \right), \\
G_k &= g_k - \sum_{j=1}^{n} q_j^k(a) \xi_1^j, \quad g_0 = 0, \\
G_{p+i} &= \sum_{j=1}^{n} Z_j^i(a) \xi_1^j - \xi_2^i
\end{aligned}
\tag{4.4.12}
$$

donde $G_0, \ldots, G_{p+n}$ son constantes. Sean

$$
J_l(x) = G_l + \int_a^b F_l(t, x(t), u(t)) dt \tag{4.4.13}
$$

para $l = 0, \ldots, p+n$. Dada la forma como se definió F_l nos resulta:

$$
\frac{\partial F_l}{\partial x_j} = 0 \tag{4.4.14}
$$

a lo largo de x_0. Sean $(x(t), u(t))$ soluciones a las ecuaciones

$$
\dot{x}_i = f_i(t, x, u). \tag{4.4.15}
$$

Entonces por (4.4.11), (4.4.12), (4.4.13) tenemos

$$
J_l(x) = I_l(x) + \sum_{j=1}^{n} q_j^l(a)(x_j(a) - \xi_1^j), \qquad (l = 0, \ldots, p)
$$

$$
J_{l+i}(x) = x_i(b) - \xi_2^i - \sum_{j=1}^{n} Z_j^i(a)(x_j(a) - \xi_1^j).
$$

Por consiguiente para cada x admisible que satisface $x_i(a) = \xi_1^i$ se obtiene

$$
\begin{aligned}
J_k(x) &= I_k(x), \\
J_{l+i}(x) &= x_i(b) - \xi_2^i
\end{aligned}
\tag{4.4.16}
$$

para $k = 0, \ldots, p$; $i = 1, \ldots, n$.

Para x admisible tal que satisface las condiciones iniciales $x_i(a) = \xi_i^1$, las restricciones $x_i(b) = \xi_2^i$ son equivalentes a pedir que se cumpla la restricción $J_{l+i}(x) = 0$.

Sea Λ la clase de trayectorias admisibles que satisfacen las condiciones iniciales $x(a) = \xi_1$, donde por trayectorias admisibles nos referimos a los $x(t), u(t)$ soluciones de (4.4.15) tales que $(t, x(t), u(t)) \in R_0$.

Antes de continuar probaremos que si (x_0, u_0) minimiza a I sobre las variaciones admisibles que satisfacen las condiciones (4.4.3) si y solo si (x_0, u_0) minimiza a J_0 en Λ sujeto a las restricciones

$$J_l(x) \leq 0 \ (l = 1, \ldots, p'), \quad J_l(x) = 0 \ (l = p' + 1, \ldots, p + n). \tag{4.4.17}$$

Λ es el conjunto de variaciones admisibles que satisfacen $x_i(a) = \xi_i^1$ y (4.4.16). Para probar esto, sea K la clase de vectores k de la forma

$$k_l = F_l(t, x_0(t), u) - F_l(t, x_0(t), u_0(t))$$

con $l = 0, \ldots, p + n$, $(t, x_0(t), u) \in R_0$ y $t \in [a, b]$ distinto de los puntos de discontinuidad de $u_0(t)$. Entonces K es un conjunto derivado de J_l en $x_0(t)$ sobre Λ (más adelante se comprobará dicha afirmación).

Por el Teorema 4.3.5 existen multiplicadores $\lambda_0 \geq 0$, $\lambda_1, \ldots, \lambda_{p+n}$ tales que $\lambda_i \geq 0$ si $i = 1, \ldots, p'$ y $\lambda_i = 0$ si $J_i(x_0) = I_i(x_0) < 0$, $\sum_{l=1}^{p} \lambda_l k_l \geq 0$ en K. Por lo tanto si

$$F(t, x, u) = \sum_{i=1}^{p} \lambda_i F_i(t, x, u)$$

entonces

$$F(t, x_0(t), u) \geq F(t, x_0(t), u_0(t)) \tag{4.4.18}$$

para $(t, x_0(t), u) \in R_0$, excepto por un número finito de puntos. Por continuidad esta desigualdad se mantiene en estos puntos. Sea

$$p_j(t) = \sum_{k=1}^{n} \lambda_k q_j^k - \sum_{i=1}^{n} \lambda_{p+i} Z_j^i,$$

por (4.4.12) tenemos

$$F = \sum_{i=1}^{n} \lambda_i L_i - \sum_{s=1}^{p} p_s f_s - \sum_{j=1}^{n} \dot{p}_j x_j.$$

Si definimos

$$H(t, x, u, p) = \sum_{s=1}^{p} p_s f_s - \sum_{i=1}^{n} \lambda_i L_i = -F(t, x, u) - \sum_{j=1}^{n} \dot{p}_j x_j$$

entonces la desigualdad

$$H(t, x_0(t), u, p(t)) \leq H(t, x_0(t), u_0(t), p(t))$$

es equivalente a (4.4.18). Por construcción resulta que:

$$\dot{p}_j = -\sum_{j=1}^{n} \left(p_i A_j^i + \lambda_i M_i^j \right) = -H_{x_j} \tag{4.4.19}$$

en x_0. Si $\lambda_i = 0$, y $p_j(\bar{t}) = -\lambda_{p+i} Z_i^j(\bar{t}) = 0$ entonces tenemos $\lambda_{p+i} = 0$ ya que $det(Z_j^i) \neq 0$. De modo que $\lambda_i = 0$ para $i = 0, \dots, p+n$, pero esto no es posible. Por lo tanto obtenemos los incisos (a), (b) y (c).

Enseguida se demostrará (d). Por (b), (c) y el Teorema 4.3.6 se tiene que para la función

$$F(t,u) = -H(t, x_0(t), u, p(t))$$

evaluada en $u_0(t)$

$$F(t, u_0(t)) = \int_a^t F_t(s, u_0(s))ds + c.$$

Usando (b),

$$F_t(t, u_0(t)) = -H_t - \sum_{i=1}^{n} H_{x_i} \dot{x_{0i}} - \sum_{i=1}^{n} H_{p_i} \dot{p}_i = -H_t(t, x_0(t), u_0(t), p(t))$$

en consecuencia se obtiene

$$H = \int_a^t H_t ds + c.$$

Por último, si R_0 es abierto, F tiene un mínimo sin restricciones en $u = u_0(t)$ para cada t en $[a, b]$,

$$F_u = 0, \ \langle \pi, F_{uu} \pi \rangle \geq 0$$

y por lo tanto

$$F_u = H_u(t, x_0(t), u_0(t), p(t)) = 0.$$

□

Es así como se ha obtenido la prueba del principio del máximo. Nos queda por demostrar que K es un conjunto derivado.

Lema 4.4.5. *La clase K es un conjunto derivado para J_l en x_0 sobre Λ.*

Demostración. Sean N vectores

$$k_j^l = F_l(t_j, x_0(t_j), u_j) - F_l(t, x_0(t_j), u_0(t_j)),$$

$j = 1, \dots, N$ en K. Supongamos que $t_1 \leq \cdots \leq t_N$. Sea $\delta > 0$ tal que $t_i + N\delta < t_j$ si $t_i < t_j$ de manera que $t_N + N\delta < b$. Por lo tanto existe $u_j(t)$ continua en $|t - t_j| \leq \delta_0 = N\delta$ tal que $(t, x, u_j(t)) \in R_0$ si (t, x) está en una δ_0-vecindad de $(t_j, x_0(t_j))$. Sea $T_1 = t_1$, $T_j = t_j + \epsilon_1 + \cdots \epsilon_{j-1}$ donde $0 \leq \epsilon_j \leq \delta' < \delta$. Definamos

$$u(t, \epsilon) = \begin{cases} u_j(t) & t \in [T_j, T_j + \epsilon_j] \\ u_0(t) & M(\epsilon) \end{cases}$$

donde $M(\epsilon)$ son los puntos faltantes de $[a, b]$. Si δ' es suficientemente pequeña entonces el sistema

$$\dot{x}_i = f_i(t, x, u(t, \epsilon)), \qquad x_i(a) = \xi_1^i$$

tienen solución $x(\epsilon) : x_i(t,\epsilon)$ tal que pertenecen a Λ y $x_0 = x(t,0)$. Sus derivadas

$$\frac{\partial x_i(t,\epsilon)}{\partial \epsilon_j}$$

son continuas a trozos (véase apéndice Teoremas 5.2.8 y 5.2.10).

El valor de J_l a lo largo de $x(\epsilon)$ se puede escribir de la forma

$$J_l(x(\epsilon)) = J_l(x_0) + \sum_{j=1}^{N} Q_{lj}\epsilon_j + R_l$$

donde

$$Q_{lj} = \frac{1}{\epsilon_j}\int_{T_j}^{T_j+\epsilon_j} (F_l(t, x(t,\epsilon), u_j(t)) - F_l(t, x_0(t), u_0(t)))dt,$$

$$R_l = \int_{M(\epsilon)} S_l(t, x(t,\epsilon))dt,$$

$$S_l(t,x) = F_l(t, x, u_0(t)) - F_l(t, x_0(t), u_0(t)).$$

Obsérvese que

$$\lim_{\epsilon\to 0} Q_{lj} = F_l(t_j, x_j, u_j) - F_l(t_j, x_j, u_0(t_j)) = k_j^l,$$

además

$$\frac{\partial S_l}{\partial x_j} = 0. \tag{4.4.20}$$

en x_0. Sea $\Delta x_i = x_i(t,\epsilon) - x_0^i(t)$, por lo cual

$$S_l(t, x(t,\epsilon)) = \Delta x_i \int_0^1 \frac{\partial S_l}{\partial x_j}(t, x_0(t) + \theta\Delta x)d\theta.$$

Como

$$\frac{\Delta x_i}{\sqrt{\epsilon_1^2 + \cdots + \epsilon_N^2}}$$

es uniformemente acotada, al usar (4.4.20)

$$\lim_{\epsilon\to 0} \frac{S_l(t, x(t,\epsilon)}{\sqrt{\epsilon_1^2 + \cdots + \epsilon_N^2}} = 0$$

uniformemente en $[a,b]$. Por lo tanto

$$\lim_{\epsilon\to 0} \frac{R_l}{\sqrt{\epsilon_1^2 + \cdots + \epsilon_N^2}} = 0.$$

La suma de $k_j^l d\epsilon_j$ es la diferencial de $J_l(x(\epsilon))$ en $\epsilon = 0$. $\square$

Corolario 4.4.6. *Si R_0 es abierto y se satisfacen las hipótesis del teorema del principio del máximo entonces*

$$\langle \pi, H_{uu}(t, x_0(t), u_0(t), p)\pi \rangle \leq 0. \tag{4.4.21}$$

para $\pi \neq 0$.

Definición 4.4.7. *Si R_0 es abierto, al sistema de ecuaciones*

$$\dot{x}_i = H_{p_i}, \quad \dot{p}_i = -H_{x_i}, \quad H_{u_k} = 0 \tag{4.4.22}$$

se le conoce como las ecuaciones de Euler-Lagrange.

Hasta aquí ya se conoce la resolución de problemas de control óptimo sujetos a condiciones de problemas isoperimétricos y de punto fijo. El siguiente paso es resolver el problema al agregarles restricciones de tipo Lagrange solo con igualdades.

La finalidad es minimizar el funcional (4.4.1) bajo las restricciones impuestas por el principio del máximo; además $x(t)$ satisface

$$\varphi_i(t, x, u) = 0, \qquad i = 1, \ldots, m \tag{4.4.23}$$

Supondremos que $\varphi_i \in C^2$ y la matriz

$$\left(\frac{\partial \varphi_i}{\partial x_j} \right)$$

tienen rango m en R_0.

Teorema 4.4.8. *Supongamos que $x_0(t)$ minimiza a I sobre la clase de variaciones admisibles que satisface las condiciones (4.4.2), (4.4.3) y (4.4.23). Entonces existen multiplicadores*

$$\lambda_0 \geq 0, \ \lambda_i, \ p_s(t), \ \mu_r(t) \quad (i = 1, \ldots, p; \ s = 1, \ldots, n; \ r = 1, \ldots, m)$$

no todos cero en T tales que para

$$H(t, x, u, p, \mu) = \sum_{s=1}^{n} p_s f_s - \lambda_0 L - \sum_{i=1}^{p} \lambda_i L_i - \sum_{r=1}^{m} \mu_r \varphi_r \tag{4.4.24}$$

se cumplen:

a. *$\lambda_i \geq 0$ con $i = 1, \ldots, p'$, y $\lambda_i = 0$ si $I_i(x_0) < 0$. Los multiplicadores $\mu_r(t)$ son continuos donde $u_0(t)$ también lo es.*

b. *Los multiplicadores $p_i(t)$ son continuos y satisfacen para $x_0(t)$, $\mu_i(t)$*

$$\dot{x}_i = f_i = H_{p_i}, \qquad \dot{p}_i = -H_{x_i}, \qquad H_{u_k} = 0, \qquad \varphi_r = 0$$

en cada intervalo de continuidad de $u_0(t)$.

c. *Si u satisface*

$$\varphi_r(t, x_0(t), u) = 0,$$

entonces

$$H(t, x_0(t), u, p(t), \mu(t)) \leq H(t, x_0(t), u_0(t), p(t), \mu(t)).$$

d. *La función* $H(t, x_0(t), u_0(t), p(t), \mu(t))$ *es continua en* $[a, b]$ *y satisface*

$$\frac{d}{dt} H = H_t$$

en cada intervalo de continuidad de $u_0(t)$.

Demostración. Por el Lema 4.3.8 existen $U_k(t, x, u)$ de clase C^2 con $k = 1, \ldots q$ en una vecindad $R' \subset R_0$ tales que $U(t, x, u) = u$ en R_0; por otra parte

$$\varphi_r(t, x, U(t, x, u)) = 0.$$

Las funciones

$$\begin{aligned} \overline{f}_i(t, x, u) &= f_i(t, x, U(t, x, u)); \\ \overline{L}_i(t, x, u) &= L_i(t, x, U(t, x, u)); \\ \overline{L}(t, x, u) &= L(t, x, U(t, x, u)); \end{aligned}$$

son de clase C^2 en R'. Sea B' la clase de $(x(t), u(t))$ que satisfacen las ecuaciones

$$\dot{x}_i = \overline{f}_i(t, x, u), \qquad x_i(a) = \xi_i^1, \; x_i(b) = \xi_i^2$$

$$\overline{I}_i(x) \leq 0 \;\; (i = 1, \ldots, p'), \qquad \overline{I}_i(x) = 0 \;\; (i = p' + 1, \ldots, p), \tag{4.4.25}$$

donde

$$\overline{I}_i(x) = g_i + \int_a^b \overline{L}_i(t, x, u) dt.$$

Definamos

$$\overline{I}(x) = \int_a^b \overline{L}(t, x(t), u(t)) dt.$$

Consideremos $\overline{x} = (\overline{x}(t), \overline{u}(t)) \in B'$. Sea $x : (x(t), u(t)) = (\overline{x}(t), U(t, \overline{x}(t), \overline{u}(t)))$ tales que satisfacen (4.4.25). Por como x se ha definido tenemos

$$x_i = f_i(t, x, u), \;\; \varphi_r(t, x, u) = 0, \;\; I_i(x) = \overline{I}_i(\overline{x}), \;\; I(x) = \overline{I}(\overline{x})$$

por lo tanto $x \in B$. Además, si $\overline{x} \in B$ entonces $u(t) = \overline{u}(t)$, por lo que x_0 está en B' y minimiza a $\overline{I}(x)$ en B'. Por el teorema del principio del máximo 4.4.4 existen multiplicadores $\lambda_0, \lambda_i, p_s(t)$ tales que para

$$\overline{H}(t, x, u, p) = \sum_{s=1}^{n} p_s \overline{f}_s - \lambda_0 \overline{L} - \sum_{i=1}^{p} \lambda_i \overline{L}_i$$

se tiene

$$\dot{p}_i = -\overline{H}_{x_i}, \qquad \frac{d}{dt} \overline{H} = \overline{H}_t$$

a lo largo de x_0,

$$\overline{H}(t, x_0(t), u, p(t)) \leq \overline{H}(t, x_0(t), u_0(t), p(t))$$

cuando $(t, x_0(t), u) \in R'$. Sea

$$H^*(t, x, u, p) = \sum_{s=1}^{n} p_s f_s - \lambda_0 L - \sum_{i=1}^{p} \lambda_i L_i$$

tenemos

$$\overline{H}(t, x, u, p) = H^*(t, x, U(t, x, u), p).$$

en R'. Como $U(t, x, u) = u$ en R_0 entonces

$$H^*(t, x_0(t), u, p(t)) \leq H^*(t, x_0(t), u_0(t), p(t))$$

$\forall (t, x_0(t), u) \in R_0$, es decir, para u que cumple $\varphi_i(t, x_0(t), u) = 0$. Por el Teorema 4.3.7 existe un único conjunto de multiplicadores μ_r tales que para

$$H(t, x, u, p, \mu) = H^*(t, x, u, p) - \sum_{r=1}^{m} \mu_r \varphi_r$$

satisfacen

$$H_{u_k} = H^*_{u_k} - \sum_{r=1}^{m} \mu_r \varphi_{r_{u_k}} = 0$$

cuando $x = x_0(t)$, $u = u_0(t)$, $p = p(t)$, $\mu = \mu(t)$. Los multiplicadores $\mu_r(t)$ son continuos donde $u_0(t)$ también lo es. Además

$$\overline{H}(t, x, u, p) = H(t, x, U(t, x, u), p, \mu)$$

en R'. Por todo esto se cumplen

$$\overline{H}_{x_i} = H_{x_i} + H_{u_k} \frac{\partial U_k}{\partial x_i} = H_{x_i}$$

$$\overline{H}_t = H_t + H_{u_k} \frac{\partial U_k}{\partial t} = H_t$$

en $x_0(t)$. □

Corolario 4.4.9. *Si x_0 minimiza a I sujeto a las condiciones del Teorema 4.4.8 entonces*

$$\langle \pi, H_{uu}(t, x_0(t), u_0(t), p, \mu)\pi \rangle \leq 0$$

para $\pi \in \mathbb{R}^n$ que satisfacen

$$\langle \varphi_u(t, x_0(t), u_0(t)), \pi \rangle = 0.$$

Una vez conocido cómo optimizar problemas de control óptimo, aplicaremos el principio del máximo para resolver el problema de cálculo variacional con restricciones de tipo Lagrange.

4.4.1. Formulación Hamiltoniana

Definición 4.4.10. *Consideremos el conjunto de restricciones en A del problema que estamos analizando, dado por*

$$\mathcal{B} := \{(t, x, \dot{x}) \in A \mid \varphi(t, x, \dot{x}) = 0\}.$$

Para $(t, x, \dot{x}, p, \mu, \lambda) \in T \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$ *sea*

$$H(t, x, \dot{x}, p, \mu, \lambda) := \langle p, \dot{x}\rangle - \lambda L(t, x, \dot{x}) - \langle \mu, \varphi(t, x, \dot{x})\rangle. \tag{4.4.26}$$

Si en el Teorema 4.4.8 se toma $f_i(t, x, u) = u_i$ tenemos que $\dot{x}_i = u_i$, además si tomamos como nulas a nuestras condiciones isoperimétricas, el funcional H, definido en el Teorema 4.4.8, se convierte en el funcional H definido por (4.4.26). Razón por la cual el principio del máximo se transforma en:

Teorema 4.4.11. *Si x_0 minimiza a I en D, entonces existen $\lambda_0 \geq 0$, $p \in X$ y $\mu \in PWC$ continua en cada intervalo de continuidad de $\dot{x}_0$, no anulándose simultáneamente, tales que*

a. $\dot{p}(t) = -H_x^T(t, \lambda_0)$ *y* $H_{\dot{x}}(t, \lambda_0) = 0$ *en intervalos de continuidad de* $\dot{x}_0$.

b. $H(t, x_0(t), u, p(t), \mu(t), \lambda_0) \leq H(t, \lambda_0)$ $\forall (t, u) \in T \times \mathbb{R}^n$ *y* $(t, x_0(t), u) \in \mathcal{B}$,

donde $H(t, \lambda_0)$ *denota* $H(t, x_0(t), \dot{x}_0(t), p(t), \mu(t), \lambda_0)$.

Por el mismo argumento el Corolario 4.4.9 se transforma en:

Corolario 4.4.12. *Supongamos que x_0 minimiza a I en D. Sea (p, μ, λ_0) como en el Teorema 4.4.11. Entonces*

$$\langle \pi, H_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t), p(t), \mu(t), \lambda_0)\pi\rangle \leq 0$$

para π que satisface $\varphi_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\pi = 0$.

Definición 4.4.13. *Diremos que una trayectoria $x \in X(\mathcal{B})$ es normal si dados (p, μ) tales que para $t \in T$ satisfacen*

$$\dot{p}(t) = (\varphi_x(x(t)))^T \mu(t) \ [= -(H_x(x(t), p(t), \mu(t), 0))^T],$$

$$0 = p(t) - (\varphi_{\dot{x}}(x(t)))^T \mu(t) \ [= (H_{\dot{x}}(x(t), p(t), \mu(t), 0))^T],$$

entonces $p \equiv 0$.

Nótese que, en este caso, necesariamente $\mu = 0$ ya que

$$(\mu(t))^T = (p(t))^T (\varphi_{\dot{x}}(x(t)))^T (\varphi_{\dot{x}}(x(t))(\varphi_{\dot{x}}(x(t)))^T)^{-1}.$$

Por lo tanto x es normal si no existe una solución diferente de cero de

$$(\dot{p}(t))^T = (p(t))^T (\varphi_{\dot{x}}(x(t)))^T (\varphi_{\dot{x}}(x(t))(\varphi_{\dot{x}}(x(t)))^T)^{-1} \varphi_x(x(t)).$$

Claramente, si x_0 es una solución normal, entonces $\lambda_0 > 0$ en el Teorema 4.4.11 y por lo tanto podemos escoger a los multiplicadores del teorema de manera que $\lambda_0 = 1$. En ese caso los multiplicadores son únicos.

Definición 4.4.14. *Definamos la segunda variación de H a lo largo de x_0 como*

$$K(x, p, \mu; y) = \int_a^b 2\overline{\Omega}(t, y(t), \dot{y}(t))dt$$

donde

$$2\overline{\Omega}(t, y, \dot{y}) := -[\langle y, H_{xx}(t, 1)y\rangle + 2\langle y, H_{x\dot{x}}(t, 1)\dot{y}\rangle + \langle \dot{y}, H_{\dot{x}\dot{x}}(t, 1)\dot{y}\rangle]$$

con $H(t, 1) = H(t, x_0(t), \dot{x}_0(t), p(t), \mu(t), 1)$.

Definición 4.4.15. *Dado $x \in X$, decimos que y es una $\mathcal{B}$-variación admisible a lo largo de x si $y \in X_0$ y satisface para $t \in T$*

$$\varphi_x(t, x(t), \dot{x}(t))y(t) + \varphi_{\dot{x}}(t, x(t), \dot{x}(t))\dot{y}(t) = 0.$$

El siguiente resultado nos da condiciones de primer y segundo orden para una solución normal del problema.

Teorema 4.4.16. *Supongamos que x_0 es una solución normal del problema. Entonces existe un único par $(p, \mu) \in X \times PWC$ tal que:*
a. *$\dot{p}(t) = -H_x^T(t, 1)$ y $H_{\dot{x}}(t, 1) = 0$ en cada intervalo de continuidad de $\dot{x}_0$.*
b. *$H(t, x_0(t), \dot{x}, p(t), \mu(t), 1) \leq H(t, 1)$ para $(t, \dot{x}) \in T \times \mathbb{R}^n$ y $(t, x_0(t), \dot{x}) \in \mathcal{B}$.*
c. *$\langle \pi, H_{\dot{x}\dot{x}}(t, 1)\pi\rangle \leq 0$ para $\pi \in \mathbb{R}^n$ que cumplen $\varphi_{\dot{x}}(x_0(t))\pi = 0$.*
d. *$K(x_0, p, \mu; y) \geq 0$ para toda y $\mathcal{B}$-variación admisible.*

4.4.2. Formulación Lagrangiana

Esta formulación es la más conveniente para trabajar con aumentabilidad.

Definición 4.4.17. *Para $(t, x, \dot{x}, \mu, \lambda) \in T \times \mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}$, sea*

$$F(t, x, \dot{x}, \mu, \lambda) := \lambda L(t, x, \dot{x}) + \langle \mu, \varphi(t, x, \dot{x})\rangle.$$

Notemos que

$$H(t, x, \dot{x}, p, \mu, \lambda) = \langle p, \dot{x}\rangle - F(t, x, \dot{x}, \mu, \lambda).$$

Por el Teorema 4.4.11 obtenemos que, si x_0 es solución del problema, entonces

$$(H_{\dot{x}}(t, \lambda_0))^T = p(t) - (F_{\dot{x}}(t, x_0(t), \dot{x}_0(t), \mu(t), \lambda_0))^T = 0.$$

Como $H_x = -F_x$ y por el principio del máximo tenemos $\dot{p} = -H_x$ resulta

$$\frac{d}{dt} F_{\dot{x}}(x_0(t), \mu(t), \lambda_0) = F_x(x_0(t), \mu(t), \lambda_0).$$

Además

$$\begin{aligned} E_F(t, x_0(t), \dot{x}_0(t), u, \mu(t), \lambda) &= H(t, x_0(t), \dot{x}_0(t), p(t), \mu(t), \lambda) \\ &\quad - H(t, x_0(t), u, p(t), \mu(t), \lambda) \end{aligned}$$

donde E_F es la función de Weierstrass de F. Por esta observación y el Teorema 4.4.11 obtenemos:

Teorema 4.4.18. *Si x_0 es solución del problema entonces existen $\lambda_0 \geq 0$ y $\mu \in PWC$ continua en cada intervalo de continuidad de $\dot{x}_0$ que no se anulan simultáneamente en $[a,b]$, tales que:*
a. $\exists c \in \mathbb{R}^n$ *tal que* $F_{\dot{x}}(x_0(t), \mu(t), \lambda_0) = \int_a^t F_x(x_0(s), \mu(s), \lambda_0)ds + c$.
b. $E_F(t, x_0(t), \dot{x}_0(t), u, \mu(t), \lambda_0) \geq 0 \ \forall (t,u) \in T \times \mathbb{R}^n$ *con* $(t, x_0, u) \in \mathcal{B}$.

Dadas las definiciones de F y H tenemos

$$H_{xx} = -F_{xx}, \quad H_{x\dot{x}} = -F_{x\dot{x}}, \quad H_{\dot{x}\dot{x}} = -F_{\dot{x}\dot{x}}.$$

Por el Corolario 4.4.12, si x_0 es solución del problema y μ y λ_0 poseen las propiedades del teorema anterior entonces

$$\langle \pi, F_{\dot{x}\dot{x}}(x_0(t), \mu(t), \lambda_0)\pi \rangle \geq 0$$

para $\pi \in \mathbb{R}^n$ que cumple $\varphi_{\dot{x}}(x_0(t))\pi = 0$. Definamos

$$J(x, \mu) := \int_a^b F(t, x, \dot{x}, \mu, 1)dt$$

cuya segunda variación está dada por

$$J''(x, \mu; y) = \int_a^b 2\Omega_\mu(t, y, \dot{y})dt$$

donde

$$2\Omega_\mu(t, y, \dot{y}) = \langle y, F_{xx}(x, \mu, 1)y \rangle + 2\langle y, F_{x\dot{x}}(x, \mu, 1)\dot{y} \rangle + \langle \dot{y}, F_{\dot{y}\dot{y}}(x, \mu, 1)\dot{y} \rangle.$$

Nótese que $K(x, p, \mu; y)$ coincide precisamente con esta segunda variación y, por el Teorema 4.4.16,

$$J''(x, \mu; y) \geq 0$$

para $y \in X_0(\mathcal{B})$.

Recordando los conjuntos introducidos en la búsqueda de mínimos fuertes y débiles se introducirán:

$$\begin{aligned}
\mathcal{E}(\mu) &:= \{x \in X \mid \exists c \text{ tal que } F_{\dot{x}}(t) = \int_a^t F_x(s)ds + c\}, \\
\mathcal{H}(\mu) &:= \{x \in X \mid J''(x, \mu; y) \geq 0 \ \forall y \in X_0(\mathcal{B})\}, \\
\mathcal{L}(\mu) &:= \{x \in X \mid \langle \pi, F_{\dot{x}\dot{x}}(t)\pi \rangle \geq 0, \ \forall \pi \in \mathbb{R}^n \text{ con } \varphi_{\dot{x}}(x(t))\pi = 0\}, \\
\mathcal{W}(\mathcal{B}, \mu) &:= \{x \in X(\mathcal{B}) \mid E_F(x(t), u, \mu(t), 1) \geq 0, \ \forall (t,u) \in T \times \mathbb{R}^n, \\
&\qquad \text{con } (t, x(t), u) \in \mathcal{B}\}
\end{aligned}$$

donde $F(t)$ denota $F(t, x(t), \dot{x}(t), \mu(t), 1)$.

Por lo anterior tenemos el siguiente resultado:

Teorema 4.4.19. *Si x_0 es normal y minimiza a I en D entonces existe un único $\mu \in PWC$ tal que $x_0 \in \mathcal{E}(\mu) \cap \mathcal{L}(\mu) \cap \mathcal{H}(\mu) \cap \mathcal{W}(\mathcal{B}, \mu)$.*

También consideremos los conjuntos para las condiciones reforzadas.

$$
\begin{aligned}
\mathcal{H}'(\mu) &:= \{x \in X \mid J''(x,\mu;y) > 0 \ \forall y \in X_0(\mathcal{B}),\ y \neq 0\}, \\
\mathcal{L}'(\mu) &:= \{x \in X \mid \langle \pi, F_{\dot{x}\dot{x}}(t)\pi\rangle > 0,\ \forall \pi \in \mathbb{R}^n \text{ con } \varphi_{\dot{x}}(x(t))\pi = 0\} \\
\mathcal{W}(\mathcal{B},\mu,\epsilon) &:= \{x_0 \in X(\mathcal{B}) \mid E_F(t,x,\dot{x},u,\mu(t),1) \geq 0,\ \forall (t,x,\dot{x},u) \in T \times \mathbb{R}^{3n} \\
&\qquad \text{con } (t,x,\dot{x}) \in T_1(x_0;\epsilon) \cap \mathcal{B},\ (t,x,u) \in \mathcal{B}\}.
\end{aligned}
$$

Capítulo 5

Aumentabilidad en dimensión infinita

5.1. Un nuevo enfoque

Ya se cuenta con lo indispensable para desarrollar el concepto de aumentabilidad en cálculo variacional. La construcción es análoga a la de dimensión finita introduciendo una función no negativa que se anula en el conjunto D (4.2.2). Definamos

$$G(t,x,\dot{x}) := \frac{1}{2}\sum_{i=1}^{m}(\varphi_i(t,x,\dot{x}))^2.$$

Ahora introduzcamos la función aumentada

$$\tilde{F}(t,x,\dot{x},\mu) := L(t,x,\dot{x}) + \langle \mu, \varphi(t,x,\dot{x})\rangle + \sigma(t,x,\dot{x})G(t,x,\dot{x}) \tag{5.1.1}$$

donde $\sigma\colon T\times\mathbb{R}^n\times\mathbb{R}^n \to \mathbb{R}$. A esta función $\tilde{F}$ le asignaremos su funcional acción.

Definición 5.1.1.

$$\tilde{J}(x,\mu) := \int_a^b \tilde{F}(t,x(t),\dot{x}(t),\mu(t))dt. \tag{5.1.2}$$

Definición 5.1.2. *Para $x_0 \in X_e(\mathcal{B})$, diremos que el problema es aumentable en x_0 si existen μ,σ tales que x_0 es un mínimo para el funcional $\tilde{J}$ en $X_e(A)$.*

De la definición podemos observar que si x_0 es un mínimo para $\tilde{J}$ entonces x_0 es una solución del problema original ya que

$$\tilde{J}(x_0,\mu) \le \tilde{J}(x,\mu)$$

y como $x_0, x \in X_e(\mathcal{B})$, entonces

$$I(x_0) = \tilde{J}(x_0,\mu) \le \tilde{J}(x,\mu) = I(x).$$

Al igual que en los casos anteriores definamos los conjuntos asociados a cada condición necesaria para obtener mínimos.

$$\begin{aligned}
\tilde{\mathcal{E}}(\mu,\sigma) &:= \{x \in X \mid \exists c \in \mathbb{R}^n \text{ tal que } \tilde{F}_{\dot{x}}(x(t)) = \int_a^t \tilde{F}_x(x(s))ds + c\},\\
\tilde{\mathcal{H}}(\mu,\sigma) &:= \{x \in X \mid \tilde{J}''(x,\mu;y) \geq 0 \ \forall y \in X_0\},\\
\tilde{\mathcal{L}}(\mu,\sigma) &:= \{x \in X \mid \tilde{F}_{\dot{x}\dot{x}}(x(t)) \geq 0, \ \ t \in T\}\\
\tilde{\mathcal{W}}(A,\mu,\sigma) &:= \{x \in X(A) \mid E_{\tilde{F}}(t,x(t),\dot{x}(t),u,\mu(t)) \geq 0,\\
&\qquad\qquad \forall (t,u) \in T \times \mathbb{R}^n, \ (t,x,u) \in A\}.
\end{aligned}$$

Teorema 5.1.3. *Supongamos que el problema es aumentable en $x_0 \in X_e(\mathcal{B})$. Entonces $\exists \mu \in PWC$ tal que $x_0 \in \mathcal{E}(\mu) \cap \mathcal{H}(\mu) \cap \mathcal{L}(\mu) \cap \mathcal{W}(\mathcal{B},\mu)$. Además x_0 minimiza a I en $X_e(\mathcal{B})$.*

Demostración. Como x_0 es un mínimo para $\tilde{J}$ entonces

$$x_0 \in \tilde{\mathcal{E}}(\mu,\sigma) \cap \tilde{\mathcal{H}}(\mu,\sigma) \cap \tilde{\mathcal{L}}(\mu,\sigma) \cap \tilde{\mathcal{W}}(A,\mu,\sigma).$$

Como $x_0 \in \tilde{\mathcal{E}}(\mu,\sigma)$ tenemos que existe $c \in \mathbb{R}^n$ tal que

$$\tilde{F}_{\dot{x}}(x_0(t),\mu(t)) = \int_a^t \tilde{F}_x(x_0(s),\mu(s))ds + c.$$

Es así que

$$F_{\dot{x}}(x_0(t),\mu(t),1) = \int_a^t F_x(x_0(s),\mu(s),1)ds + c$$

por lo que $x_0 \in \mathcal{E}(\mu)$.

Tenemos

$$\tilde{J}''(x_0,\mu;y) = \int_a^b (\langle y, \tilde{F}_{xx}y\rangle + 2\langle y, \tilde{F}_{x\dot{x}}\dot{y}\rangle + \langle \dot{y}, \tilde{F}_{\dot{x}\dot{x}}\dot{y}\rangle)dt.$$

Entonces

$$\tilde{J}''(x_0,\mu;y) = J''(x_0,\mu;y) + \int_a^b \sigma(x_0(t))|\varphi_x(x_0(t))y + \varphi_{\dot{x}}(x_0(t))\dot{y}|^2 dt.$$

Como $x_0 \in \tilde{\mathcal{H}}(\mu,\sigma)$, se sigue que $J''(x_0,\mu;y) \geq 0 \ \forall y \in X_0(\mathcal{B})$ y por lo tanto $x_0 \in \mathcal{H}(\mu)$.

Para la tercera contención, notemos que

$$\tilde{F}_{\dot{x}\dot{x}}(x_0(t),\mu(t)) = F_{\dot{x}\dot{x}}(x_0(t),\mu(t),1) + \sigma(x_0(t))G_{\dot{x}\dot{x}}(x_0(t)).$$

Por otro lado, para toda $x \in X$,

$$G_{\dot{x}}(x(t)) = \sum_{i=1}^m \varphi_i(x(t))\varphi_{i\dot{x}}(x(t))$$

y por lo tanto

$$\langle \pi, G_{\dot{x}\dot{x}}(x_0(t))\pi\rangle = \pi^T \left(\sum_{i=1}^{m} \varphi_{i\dot{x}}(x_0(t))^T(\varphi_{i\dot{x}}(x_0(t))\right)\pi =$$

$$\sum_{i=1}^{m}(\varphi_{i\dot{x}}(x_0(t))\pi)^2 = |\varphi_{\dot{x}}(x_0)\pi|^2$$

por lo que

$$\langle \pi, F_{\dot{x}\dot{x}}(x_0(t), \mu(t), 1)\pi\rangle + \sigma|\varphi_{\dot{x}}(x_0)\pi|^2 \geq 0$$

lo cual implica que $x_0 \in \mathcal{L}(\mu)$.

Para el último conjunto tenemos

$$E_{\tilde{F}}(t, x, \dot{x}, u, \mu) \geq 0$$

para $(t, x_0(t), u) \in A$, por lo que

$$E_{\tilde{F}} = E_F(t, x_0, \dot{x}_0, u, \mu, 1) + \frac{1}{2}\sigma|\varphi(t, x_0, u)|^2 \geq 0$$

y concluimos que $x_0 \in \mathcal{W}(\mathcal{B}, \mu)$. □

5.2. Suficiencia y aumentabilidad

Para resultados de suficiencia introducimos el concepto de aumentabilidad de acuerdo con el tipo de mínimo local que estudiemos.

Definición 5.2.1. *Para $x_0 \in X_e(\mathcal{B})$, diremos que el problema es fuertemente aumentable en x_0 si existen $\epsilon > 0$, μ y σ tales que x_0 es un mínimo para el problema de minimizar $\tilde{J}$ en $X_e(T_0(x_0; \epsilon) \cap A)$. Claramente, en este caso, x_0 es un mínimo fuerte del problema original. Análogamente, diremos que el problema es débilmente aumentable en x_0 si en la definición anterior sustituimos T_0 por T_1, implicando que x_0 es un mínimo débil del problema original.*

Trabajaremos con $x\colon T \to \mathbb{R}^n$ absolutamente continuas; al conjunto de estos x los denotaremos por X'; además trabajaremos con las variaciones $y \in X_0$ tales que su derivada es cuadrado integrable (llamadas admisibles) y satisfacen

$$\varphi_x(x(t))y(t) + \varphi_{\dot{x}}(x(t))\dot{y}(t) = 0 \tag{5.2.1}$$

para casi todo T, con $x \in X'(\mathcal{B})$.

Definiremos los conjuntos para las condiciones reforzadas

$$\begin{aligned}
\tilde{\mathcal{H}}'(\mu, \sigma) &:= \{x \in X' \mid \tilde{J}''(x, \mu; y) > 0 \ \forall y \in X_0,\ y \neq 0\},\\
\tilde{\mathcal{L}}'(\mu, \sigma) &:= \{x \in X' \mid \langle \pi, \tilde{F}_{\dot{x}\dot{x}}(x(t), \mu(t))\pi\rangle > 0 \ \forall \pi \in \mathbb{R}^n,\ \pi \neq 0\},\\
\tilde{\mathcal{W}}(A, \mu, \sigma; \epsilon) &:= \{x_0 \in X'(A) \mid E_{\tilde{F}}(t, x, \dot{x}, u, \mu(t)) \geq 0 \ \forall (t, x, \dot{x}, u) \in T \times \mathbb{R}^{3n},\\
&\qquad (t, x, \dot{x}) \in T_1(x_0; \epsilon),\ (t, x, u) \in A\}.
\end{aligned}$$

El siguiente lema es necesario para demostrar resultados posteriores. Para su demostración véase [7]. La primera afirmación se probará además al final de esta sección.

Lema 5.2.2. *Si* $x_0 \in \mathcal{L}'(\mu)$ *entonces* $\exists \sigma_0 > 0$ *tal que, si* $\sigma(t, x, \dot{x}) \geq \sigma_0$*, entonces:*

a. $x_0 \in \tilde{\mathcal{L}}'(\mu, \sigma)$.

b. *Si* $(y_n)_{n \in \mathbb{N}} \subset X_0$ *converge uniformemente a* y_0 *en* T *entonces*

$$\liminf_{n \to \infty} \tilde{J}''(x_0, \mu; y_n) \geq \tilde{J}''(x_0, \mu; y_0).$$

Lema 5.2.3. *Si* $x_0 \in \mathcal{L}'(\mu) \cap \mathcal{H}'(\mu)$ *entonces existe* $\sigma_0 > 0$ *tal que, si* $\sigma(t, x, \dot{x}) \geq \sigma_0$*, entonces* $x_0 \in \tilde{\mathcal{H}}'(\mu, \sigma)$.

Demostración. Definamos

$$\Phi(t, y, \dot{y}) = \varphi_x(x_0(t))y + \varphi_{\dot{x}}(x_0(t))\dot{y}, \tag{5.2.2}$$

$$P(y) := J''(x_0, \mu; y),$$

$$Q(y) := \int_a^b |\Phi(t, y(t), \dot{y}(t))|^2 dt.$$

Por las hipótesis $P(y) > 0$ para toda y variación admisible no nula que satisface $\Phi(t, y(t), \dot{y}(t)) = 0$ casi en todo T y $y(a) = y(b) = 0$. Como

$$\tilde{J}''(x_0, \mu; y) = P(y) + \int_a^b \sigma(x_0(t))|\Phi(t, y(t), \dot{y}(t))|^2 dt$$

entonces

$$\tilde{J}''(x_0, \mu; y) \geq P(y) + \sigma_0 Q(y)$$

para $\sigma(x_0(t)) \geq \sigma_0$.

Supongamos que la primera afirmación del lema es falsa. Entonces $\forall n \in \mathbb{N}$ existe y_n variación admisible no nula con $y_n(a) = y_n(b) = 0$ tal que

$$P(y_n) + nQ(y_n) \leq \tilde{J}''(x_0, \mu; y_n) \leq 0 \tag{5.2.3}$$

para $\sigma(t, x, \dot{x}) \geq n$. Como las funciones son homogéneas en y entonces podemos suponer que nuestras variaciones y_n son tales que

$$\int_a^b \left(|y_n(t)|^2 + |\dot{y}_n(t)|^2\right) dt = 1 \tag{5.2.4}$$

y por lo tanto podemos reemplazar a (y_n) por una subsucesión (usando mismos subíndices) tal que esta converja a una variación admisible y_0 uniformemente en T.

Claramente $y_0(a) = y_0(b) = 0$. Por el Lema 5.2.2 existe $\sigma > 0$ tal que

$$\liminf_{n\to\infty}(P(y_n) + \sigma Q(y_n)) \geq P(y_0) + \sigma Q(y_0). \tag{5.2.5}$$

Esta desigualdad, junto con (5.2.3) y el hecho de que $Q(y) \geq 0$ implica que

$$\liminf_{n\to\infty} Q(y_n) \leq 0.$$

Como la condición de Legendre es verdadera para $Q(y)$ (véase [11] pág. 358) tenemos que

$$\liminf_{n\to\infty} Q(y_n) \geq Q(y_0) \geq 0.$$

Por lo tanto $Q(y_0) = 0$, lo cual implica que $\Phi(t, y_0(t), \dot{y}_0(t)) = 0$ casi en todo T. Supongamos que $y_0 \not\equiv 0$. Entonces $P(y_0) > 0$. Sin embargo, por (5.2.5) con $Q(y_0) = 0$, se tiene a partir de cierta n que

$$P(y_n) + \sigma Q(y_n) > 0$$

lo cual contradice la desigualdad en (5.2.3). Por lo tanto $y_0 \equiv 0$.

Completaremos la demostración probando que y_0 no puede ser la variación nula. Supongamos que sí lo es. Sea σ como en el Lema 5.2.2. Por (5.2.5) obtenemos

$$\liminf_{n\to\infty} \tilde{J}''(x_0, \mu; y_n) = \liminf_{n\to\infty}(P(y_n) + \sigma Q(y_n)) \geq 0.$$

Esto último es consecuencia de $P(y_0) = Q(y_0) = 0$. Por (5.2.3), vemos que se da la igualdad. Por lo tanto, como la convergencia es uniforme y $y_0 \equiv 0$, tenemos

$$0 = \liminf_{n\to\infty} \tilde{J}''(x_0, \mu; y_n) = \liminf_{n\to\infty} \int_a^b \langle \dot{y}_n(t), \tilde{F}_{\dot{x}\dot{x}}(x_0(t), \mu(t))\dot{y}_n(t)\rangle dt. \tag{5.2.6}$$

Como, por el Lema 5.2.2, el integrando es una forma positiva definida, existe $c > 0$ tal que

$$\langle \pi, \tilde{F}_{\dot{x}\dot{x}}(x_0(t), \mu(t))\pi\rangle \geq c|\pi|^2.$$

Por lo tanto, la ecuación (5.2.6) implica que

$$\liminf_{n\to\infty} \int_a^b |\dot{y}_n|^2 dt = 0$$

pero por la definición de y_n resulta

$$\lim_{n\to\infty} \int_a^b |\dot{y}_n|^2 = 1.$$

Esta contradicción completa la demostración. □

El siguiente resultado es el principal de este trabajo. Probaremos que las condiciones clásicas de suficiencia para un mínimo local, ya sea débil o fuerte, implican el concepto de aumentabilidad correspondiente. Algunas de las ideas utilizadas en la demostración se basan principalmente en el artículo de Hestenes [7] donde se analiza el problema normal de Bolza.

Teorema 5.2.4. *Supongamos que* $x_0 \in X'_e(\mathcal{B}) \cap C^1$ *y* μ *es absolutamente continua. Entonces se cumplen las siguientes afirmaciones.*

a. *Si* $x_0 \in \mathcal{E}(\mu) \cap \mathcal{H}'(\mu) \cap \mathcal{L}'(\mu)$*, el problema es débilmente aumentable en* x_0*.*

b. *Si además* $x_0 \in \mathcal{W}(\mathcal{B}, \mu; \epsilon)$ *para alguna* $\epsilon > 0$*, el problema es fuertemente aumentable en* x_0*.*

Demostración.

(a): Por el Teorema 3.2.5, la conclusión de que el problema es débilmente aumentable en x_0 se sigue si probamos que, para alguna función $\sigma(t, x, \dot{x})$, x_0 pertenece a

$$\tilde{\mathcal{E}}(\mu, \sigma) \cap \tilde{\mathcal{H}}'(\mu, \sigma) \cap \tilde{\mathcal{L}}'(\mu, \sigma).$$

Como $x_0 \in \mathcal{E}(\mu)$, $\exists c \in \mathbb{R}^n$ tal que para toda $t \in [a, b]$,

$$F_{\dot{x}}(x_0(t), \mu, 1) = \int_a^t F_x(x_0(s), \mu(s), 1) ds + c$$

y, por lo tanto,

$$\tilde{F}_{\dot{x}}(x_0(t), \mu(t)) = \int_a^t \tilde{F}_x(x_0(s), \mu(s)) ds + c.$$

Así pues $x_0 \in \tilde{\mathcal{E}}(\mu, \sigma)$ para cualquier σ.

Para la segunda contención (véase Lema 5.2.2) definamos, $\forall t \in [a, b]$ y $h \in \mathbb{R}^n$,

$$P(t, h) := \langle h, F_{\dot{x}\dot{x}}(x_0(t), \mu(t), 1) h \rangle, \quad Q(t, h) := |\varphi_{\dot{x}}(x_0(t)) h|^2.$$

Como $x_0 \in \mathcal{L}'(\mu)$, $P(t, h) > 0$ $\forall t \in [a, b]$ y $h \neq 0$ con $Q(t, h) = 0$. Afirmamos que existe $c > 0$ tal que $P(t, h) + cQ(t, h) > 0$ $\forall t \in [a, b]$ y $h \neq 0$. Supongamos lo contrario. Entonces $\forall n \in \mathbb{N}$ existe $(t_n, h_n) \in T \times \mathbb{R}^n$ con $h_n \neq 0$ tal que

$$P(t_n, h_n) + nQ(t_n, h_n) \leq 0.$$

Definamos $k_n := h_n / |h_n|$, de manera que $|k_n| = 1$ y

$$P(t_n, k_n) + nQ(t_n, k_n) \leq 0.$$

Por lo tanto existen una subsucesión (sin volver a etiquetar), $t_0 \in [a, b]$ y un vector unitario k_0 tales que $(t_q, k_q) \to (t_0, k_0)$. Por lo tanto $P(t_0, k_0) \leq 0$ y $Q(t_0, k_0) = 0$, contrario a $x_0 \in \mathcal{L}'(\mu)$. Ahora, sea $\sigma(t, x, \dot{x})$ tal que $\sigma(x_0(t)) > c$ para toda $t \in T$. Entonces

$$\langle h, \tilde{F}_{\dot{x}\dot{x}}(x_0(t), \mu(t)) h \rangle = P(t, h) + \sigma(x_0(t)) Q(t, h) > 0$$

para toda $h \in \mathbb{R}^n$, $h \neq 0$, y $t \in T$, lo cual implica que $x_0 \in \tilde{\mathcal{L}}'(\mu, \sigma)$.

Finalmente, la afirmación de que existe $\sigma_0 > 0$ tal que para $\sigma(x_0(t)) \geq \sigma_0$, $x_0 \in \tilde{\mathcal{H}}'(\mu, \sigma)$ se sigue del Lema 5.2.3.

(b): Para este inciso queremos probaremos que, si $x_0 \in \mathcal{W}(\mathcal{B}, \mu; \epsilon)$, entonces existe $\sigma(t, x, \dot{x})$ tal que $x_0 \in \tilde{\mathcal{W}}(A, \mu, \sigma; \epsilon)$. Una vez probada esta afirmación, el resultado se sigue del Teorema 3.2.5.

La idea de la demostración es la siguiente. Sea $E_\psi(\dot{x}, u)$ la función exceso de Weierstrass para $\psi(\dot{x}) := (1 + |\dot{x}|^2)^{1/2}$. Como se prueba en [5], x_0 satisface la condición reforzada de Weierstrass

$$E_F(t, x, \dot{x}, u, \mu(t)) \geq 0$$

para toda $(t, x, \dot{x}) \in T_1(x_0; \epsilon) \cap \mathcal{B}$, $(t, x, u) \in \mathcal{B}$ (es decir, $x_0 \in \mathcal{W}(\mathcal{B}, \mu; \epsilon)$), y la trayectoria es no singular en el sentido de que

$$\begin{vmatrix} F_{\dot{x}\dot{x}} & \varphi_{\dot{x}} \\ \varphi_{\dot{x}}^T & 0 \end{vmatrix} \neq 0 \quad \text{a lo largo de } x_0$$

si y solo si existen una vecindad $\mathcal{B}_0$ de x_0 relativa a $\mathcal{B}$ y $\tau > 0$ tales que

$$E_F(t, x, \dot{x}, u, \mu(t)) \geq \tau E_\psi(\dot{x}, u)$$

siempre que $(t, x, \dot{x}) \in \mathcal{B}_0$, $(t, x, u) \in \mathcal{B}$. Ahora, como $\varphi_{\dot{x}}(x_0(t))$ tiene rango m y

$$\langle h, F_{\dot{x}\dot{x}}(x_0(t))h \rangle \geq 0$$

para toda $h \in S$ donde

$$S = \{h \in \mathbb{R}^n \mid \varphi_{\dot{x}}(x_0(t))h = 0\},$$

la condición de no singularidad es equivalente a la relación

$$\langle h, F_{\dot{x}\dot{x}}(x_0(t))h \rangle > 0 \quad \text{para toda } h \neq 0 \text{ en } S$$

es decir, $x_0 \in \mathcal{L}'(\mu)$. Además, como se puede verificar fácilmente,

$$E_{\tilde{F}}(t, x, \dot{x}, u, \mu(t)) \geq \tau E_\psi(\dot{x}, u)$$

siempre que $(t, x, \dot{x}) \in \mathcal{B}_0$, $(t, x, u) \in \mathcal{B}$. La idea principal de la demostración consiste en probar que existen una función $\sigma(t, x, \dot{x}) \geq \sigma_0$, un número positivo τ y A_0 vecindad de x_0 relativa a A tales que

$$E_{\tilde{F}}(t, x, \dot{x}, u, \mu(t)) \geq \tau E_\psi(\dot{x}, u) \tag{5.2.7}$$

siempre que $(t, x, \dot{x}) \in A_0$, $(t, x, u) \in A$. En vista de la caracterización mencionada arriba, esto implicará que x_0 pertenece a $\tilde{\mathcal{W}}(A, \mu, \sigma; \epsilon)$

Observemos primero que existen $\tau_1 > 0$ y A_1 vecindad de x_0 tales que

$$E_{\tilde{F}}(t, x, \dot{x}, u) \geq \tau_1 E_\psi(\dot{x}, u) \tag{5.2.8}$$

siempre que $(t, x, \dot{x}) \in \mathcal{B} \cap A_1$, $(t, x, u) \in \mathcal{B}$. Por el teorema de Taylor y continuidad se sigue del Lema 5.2.2 que podemos disminuir a τ_1 y A_1 de manera que (5.2.8) se satisfaga para toda $(t, x, \dot{x})$ y $(t, x, u) \in A_1$ (no necesariamente en $\mathcal{B}$).

Afirmamos que existen $\tau > 0$ y A_0 vecindad de x_0 con la cerradura de A_0 contenida en A_1 tal que (5.2.7) se satisface para toda $(t, x, \dot{x}) \in A_0$, $(t, x, u) \in \mathcal{B}$. Para problarlo, seleccionemos primero una vecindad A^* de x_0 con la cerradura de A^* contenida en A_1 y sea $0 < \epsilon < 1$ tal que, si $(t, x, \dot{x}) \in A^*$ y (t, x, u) es exterior a A_1 entonces

$$3\epsilon\psi(u) \leq E_\psi(\dot{x}, u) \leq 2\psi(u). \tag{5.2.9}$$

Ahora, escojamos a A_0 de tal manera que exista $r(t, x, \dot{x})$ definida en A_0 tal que

$$(t, x, r(t, x, \dot{x})) \in A^* \cap \mathcal{B} \quad \text{para toda } (t, x, \dot{x}) \in A_0.$$

Entonces tenemos que

$$E_{\tilde{F}}(t, x, \dot{x}, u) = E_{\tilde{F}}(t, x, r(t, x, \dot{x}), u) + K(t, x, \dot{x}, u)$$

donde

$$\begin{aligned} K(t, x, \dot{x}, u) \quad = \quad & \tilde{F}(t, x, r(t, x, \dot{x})) - \tilde{F}(t, x, \dot{x}) + \tilde{F}_{\dot{x}}(t, x, r(t, x, \dot{x}))u \\ & -\tilde{F}_{\dot{x}}(t, x, \dot{x})u + \tilde{F}_{\dot{x}}(t, x, \dot{x})\dot{x} - \tilde{F}_{\dot{x}}(t, x, r(t, x, \dot{x}))r(t, x, \dot{x}). \end{aligned}$$

Si A_0 se toma suficientemente pequeño tendremos

$$|K(t, x, \dot{x}, u)| < \epsilon\tau_1\psi(u) \quad \text{para toda } (t, x, \dot{x}) \in A_0.$$

Por (5.2.8) y (5.2.9), si $(t, x, \dot{x}) \in A_0$ y $(t, x, u) \in \mathcal{B}$ pero no en A_1 entonces

$$E_{\tilde{F}}(t, x, \dot{x}, u) \geq \tau_1 E_\psi(r(t, x, \dot{x}), u) - \epsilon\tau_1\psi(u) \geq 2\epsilon\tau_1\psi(u) \geq \epsilon\tau_1 E_\psi(\dot{x}, u).$$

Haciendo $\tau = \epsilon\tau_1$ en (5.2.7), se sigue la afirmación.

Seleccionemos ahora conjuntos abiertos $A_2, A_3, \ldots$ cuya unión sea A y tales que la cerradura de A_j esté contenida en A_{j+1} $(j = 1, 2, \ldots)$. Sean $\sigma_j(t, x, \dot{x})$ funciones de clase C^2 tales que

$$\sigma_j = 0 \text{ en } A_{j-1}, \quad \sigma_j \geq 0 \text{ en } A_j, \quad \sigma_j = 1 \text{ en } A - A_j. \tag{5.2.10}$$

Sea $(t, x, \dot{x}) \in A_0$. Si $(t, x, u) \in A_{j+1} - A_j$ $(j \geq 1)$ entonces (5.2.7) se satisface si $(t, x, u) \in \mathcal{B}$ y por lo tanto, por continuidad, si

$$\varphi_\alpha(t, x, u)^2 < \epsilon_j E_\psi(\dot{x}, u) \tag{5.2.11}$$

donde ϵ_j es una constante positiva. Sea $\delta_j > 0$ tal que

$$E_{\tilde{F}}(t, x, \dot{x}, u) > \delta_j \quad \text{si } (t, x, u) \in A_{j+1} - A_j.$$

Sea $d_j > 0$ tal que, en este conjunto,

$$d_j\epsilon_j E_\psi(\dot{x}, u) + \delta_j > \tau_1 E_\psi(\dot{x}, u). \tag{5.2.12}$$

Definamos $\sigma(t, x, \dot{x}) := \sigma_0 + \sum d_j\sigma_j(t, x, \dot{x})$.

Por (5.2.10), tenemos que

$$\sigma - \sigma_0 = 0 \text{ en } A_0, \quad \sigma - \sigma_0 \geq 0 \text{ en } A, \quad \sigma - \sigma_0 \geq d_j \text{ en } A - A_j. \tag{5.2.13}$$

Sea $h(t, x, \dot{x}) := (\sigma(t, x, \dot{x}) - \sigma_0)|\varphi(t, x, \dot{x})|^2$. Nótese que $h \equiv 0$ en A_0 y

$$E_h(t, x, \dot{x}, u) = h(t, x, u) \geq 0$$

(siempre y cuando $(t, x, \dot{x}) \in A_0$ como hemos supuesto). Sea $F^* = \tilde{F} + h$. Entonces, siempre que $(t, x, \dot{x}) \in A_0$,

$$E_{F^*}(t, x, \dot{x}, u) = E_{\tilde{F}}(t, x, \dot{x}, u) + h(t, x, u). \tag{5.2.14}$$

Si $(t, x, u) \in A_1$ entonces (5.2.7) se satisface, de manera que

$$E_{F^*}(t, x, \dot{x}, u) \geq \tau E_\psi(\dot{x}, u). \tag{5.2.15}$$

Esto es válido si $(t, x, u) \in A_{j+1} - A_j$ $(j \geq 1)$ y (5.2.11) se satisface. En caso contrario entonces, por (5.2.13),

$$h(t, x, u) \geq (\sigma(t, x, \dot{x}) - \sigma_0)\epsilon_j E_\psi(\dot{x}, u) \geq d_j \epsilon_j E_\psi(\dot{x}, u).$$

Se sigue de (5.2.12), (5.2.14) que (5.2.15) también se satisface en este caso. □

Conclusiones

Mediante el trabajo realizado mostramos cómo algunos aspectos fundamentales relacionados con la teoría de aumentabilidad se pueden desarrollar para ciertos problemas de optimización definidos en espacios de dimensión infinita. En particular nos concentramos en el problema básico de cálculo de variaciones sujeto a restricciones con igualdades. Es importante mencionar que dicho concepto es ampliamente conocido y utilizado en la literatura para resolver problemas con restricciones en espacios de dimensión finita.

Después de un breve resumen de la teoría básica de aumentabilidad para problemas con restricciones en dimensión finita, derivamos para el problema básico de cálculo de variaciones con puntos fijos las condiciones necesarias y suficientes clásicas de primer y segundo orden. Dicha derivación resulta fundamental en la teoría desarrollada posteriormente al introducir funciones de penalización que remueven las restricciones involucradas. Posteriormente enunciamos el problema con restricciones en forma de igualdades y derivamos las condiciones clásicas de primer orden a través del principio máximo, visto el problema original como un problema de control óptimo. Así mismo, suponiendo que la solución al problema es normal, enunciamos las condiciones clásicas de segundo orden tanto en su formulación Hamiltoniana como Lagrangiana.

Finalmente desarrollamos dos de los aspectos más importantes del concepto de aumentabilidad que lo han convertido en una herramienta de gran utilidad tanto teórica como práctica para problemas de optimización. En primer lugar, derivamos las condiciones necesarias clásicas de primer y segundo orden sin imponer la hipótesis de normalidad, lo cual permite analizar problemas para los cuales es posible encontrar soluciones locales no necesariamente normales al problema. En segundo lugar probamos que, para mínimos locales fuertes o débilies, las condiciones suficientes clásicas implican el tipo de aumentabilidad correspondiente. Este último resultado justifica ampliamente el análisis del problema aumentado sin restricciones y, por otro lado, nos permite desarrollar una teoría alternativa de suficiencia.

Es de interés para un trabajo posterior llevar a cabo generalizaciones de este enfoque tanto en necesidad sin normalidad como en suficiencia, así como un desarrollo detallado, para problemas de control óptimo, de los métodos de multiplicadores utilizados tanto en problemas definidos en espacios de dimensión finita como en análisis convexo.

Apéndice

Teorema 5.2.5. *(Leibniz) Sean $T := [a,b] \in \mathbb{R}$, $\phi : T \to \mathbb{R}$ y $\psi : T \to \mathbb{R}$ tales que $\phi(t) \leq \psi(t)$, $\forall t \in T$. Definimos*

$$K = \{(x,t) \in \mathbb{R}^2 | t \in T, \phi(t) \leq x \leq \psi(t)\}.$$

Sea $D \subset \mathbb{R}^2$ tal que $K \subset D$, $f : D \to \mathbb{R}$:

a. *Si f, ϕ, ψ son continuas entonces existe*

$$F(t) := \int_{\phi(t)}^{\psi(t)} f(x,t)dx$$

y además F es continua.

b. *Si $\phi, \psi \in C^1(T)$, f es continua, existe $\dfrac{\partial f}{\partial t}$ y además es continua entonces F es derivable y su derivada está dada por*

$$F'(t) = \int_{\phi(t)}^{\psi(t)} \frac{\partial}{\partial t} \left(f(x,t)\right) dx + \psi'(t) f(\psi(t),t) - \phi'(t) f(\phi(t),t).$$

Teorema 5.2.6. *(Teorema de la función implícita) Sea $f : S \to \mathbb{R}^n$ con $S \subset \mathbb{R}^m \times \mathbb{R}^n$ abierto. Supongamos f y $f_x(v,x)$ son continuas en S y existe $V_0 \subset \mathbb{R}^m$ compacto y $x_0 : V_0 \to \mathbb{R}^n$ continua tal que para $v \in V_0$ tenemos $(v, x_0(v)) \in S$, $f(v, x_0(t)) = 0$ y $|f_x(v, x_0(t))| \neq 0$. Entonces existe V vecindad de V_0, $\epsilon > 0$ y $x : V \to \mathbb{R}^n$ continuas tales que:*

a. *$x(v) = x_0(v)$ para toda $v \in V_0$.*

b. *$f(v, x(v)) = 0$ para toda $v \in V$.*

c. *Si $v \in V$, $f(v,x) = 0$ y $|x - x(v)| < \epsilon$ entonces $x = x(v)$.*

d. *Si $f \in C^m(s)$ entonces $x \in C^m(V)$.*

Sea $B \subset \mathbb{R} \times \mathbb{R}^n$ una región. Sea $f \colon B \to \mathbb{R}^n$ con f_x continua. Sea

$$\dot{x}(t) = f(t, x(t)). \tag{5.2.16}$$

Teorema 5.2.7. *Dado $(\alpha, \beta) \in B$ existen $\delta > 0$ y $x \colon (\alpha - \delta, \beta + \delta) \to \mathbb{R}^n$ única solución a la ecuación (5.2.16) que satisface $x(\alpha) = \beta$. Si $f \in C^m(B)$ entonces $x \in C^{m+1}(\alpha - \delta, \alpha + \delta)$.*

Teorema 5.2.8. *Sea x_0 solución de (5.2.16). Entonces existen $\rho > 0$ y B_0 vecindad de $\{(t, x_0(t)) \mid t \in T\}$ tales que para cada $(\alpha, \beta) \in B_0$ pasa un único $x(\cdot; \alpha, \beta) \colon [a - \rho, b + \rho] \to \mathbb{R}^n$ tal que es solución a (5.2.16) y satisface*

$$x(t; \alpha, x_0(\alpha)) = x_0(t).$$

Las funciones $x(t; \alpha, \beta)$, $\dot{x}(t; \alpha, \beta)$ son de clase C^1 respecto a β, $\forall (t, \alpha, \beta) \in [a - \rho, b + \rho] \times B_0$ y $|x_\beta(t; \alpha, \beta)| \neq 0$. Si $f \in C^m(B)$ entonces $x, \dot{x} \in C^m$.

Teorema 5.2.9. *Sea $\delta > 0$, supongamos que $x : T \times (-\delta, \delta) \to \mathbb{R}^n$ es una familia uniparamétirca de soluciones a la ecuación (5.2.16) que satisface $x(t, 0) = x_0(t)$. Si x, x_ϵ son continuas entonces $y(t) := x_\epsilon(t, 0)$ satisface*

$$\dot{y}(t) = \langle f_x(t, x_0(t)), y(t) \rangle$$

para $t \in T$.

Sea Λ un subconjunto de un espacio normado, $\lambda_0 \in \Lambda$ fijo. Sea x_0 solución de

$$\dot{x} = f(t, x, \lambda). \tag{5.2.17}$$

Para $\lambda = \lambda_0$ tenemos por el teorema 5.2.8 existe solución $x(\cdot; \alpha, \beta, \lambda_0)$ tal que satisface $x_0(t) = x(t; \alpha, x_0(\alpha), \lambda_0)$.

Teorema 5.2.10. *Existen $\sigma, p > 0$ tal que para cada (α, β) que satisface las relaciones*

$$a - p \leq \alpha \leq b + p, \quad |\beta - x(\alpha)| < \sigma, \quad |\lambda - \lambda_0| < \sigma \tag{5.2.18}$$

pasa una única solución

$$y(t, \alpha, \beta, \lambda), \qquad (a - p \leq t \leq b + p)$$

de (5.2.17) que contiene a x para $t \in [a, b]$, $\lambda = \lambda_0$, $\beta = x(\alpha)$. La función $y(t, \alpha, \beta, \lambda)$ es continua. Existe C constante tal que

$$|y(t, \alpha, \beta, \lambda) - x(t)| < C|\beta - x(\alpha)| + CG(x, \lambda, \lambda_0) \tag{5.2.19}$$

en $[a - p, b + p]$ donde

$$G(x, \lambda, \lambda_0) = \int_{a-p}^{b+p} |f(t, x(t), \lambda) - f(t, x(t), \lambda_0)| dt.$$

Además si $(\alpha', \beta', \lambda')$ satisface (5.2.18) entonces la desigualdad (5.2.19) se sigue cumpliendo si reemplazamos a $x(t)$ por $y(t, \alpha', \beta', \lambda')$ y λ_0 por λ'.

Bibliografía

[1] Bertsekas DP (1976) *On penalty and multiplier methods for constrained minimization*, SIAM Journal on Control and Optimization, **14**: 216-235

[2] Bertsekas DP (1982) *Constrained Optimization and Lagrange Multiplier Methods*, Computer Science and Applied Mathematics, Academic Press, New York

[3] Fortin M, Glowinski R (1983) *Augmented Lagrangian Methods. Applications to the Numerical Solution of Boundary-Value Problems*, North-Holland, Amsterdam

[4] Glowinski R (1984) *Numerical Methods for Nonlinear Variational Problems*, Springer Series in Computational Physics, Springer-Verlag, New York

[5] Hestenes MR (1946) *The Weierstrass E-function in the calculus of variations*, Transactions of the American Mathematical Society, **61**: 51-71

[6] Hestenes MR (1947) *An alternate sufficiency proof for the normal problem of Bolza*, Transactions of the American Mathematical Society, **61**: 256-264

[7] Hestenes MR (1947) *An indirect sufficiency proof for the problem of Bolza in nonparametric form*, Transactions of the American Mathematical Society, **62**: 509-535

[8] Hestenes MR (1966) *Calculus of Variations and Optimal Control Theory*, John Wiley, New York

[9] Hestenes MR (1975) *Optimization Theory. The Finite Dimensional Case*, John Wiley, New York

[10] López E, Molgado A, Vallejo JA (2012) *The principle of stationary action in the calculus of variations*, Communications in Mathematics, **20**: 89-116

[11] McShane EJ (1942) *Sufficient conditions for a weak relative minimum in the problem of Bolza*, Transactions of the American Mathematical Society, **52**: 344-379

[12] Miele A, Moseley PE, Levy AV, Coggins GM (1972) *On the method of multipliers for mathematical programming problems*, Journal of Optimization Theory and Applications, **10**: 1-33

[13] Reid WT (1939) *Isoperimetric problems of Bolza in nonparametric form*, Duke Mathematical Journal, **5**: 675-691

[14] Rockafellar RT (1973) *The multiplier method of Hestenes and Powell applied to convex programming*, Journal of Optimization Theory and Applications, **12**: 555-562

[15] Rosenblueth JF (2013) *Augmented integrals for optimal control problems*, International Journal of Applied Mathematics and Informatics, **7**: 44-54