UNIVERSIDAD NACIONAL AUTONOMA DE MEXICO

FACULTAD DE CIENCIAS

Puntos conjugados en control óptimo

T E S I S

Que para obtener el título de

A c t u a r i o

p r e s e n t a

Gerardo Sánchez Licea

Director de Tesis:
Dr. Javier Fernando Rosenblueth Laguette

MEXICO, D. F. **2000**

VNIVERSIDAD NACIONAL
AVFNºMA DE
MEXICO

MAT. MARGARITA ELVIRA CHÁVEZ CANO
Jefa de la División de Estudios Profesionales de la
Facultad de Ciencias
Presente

Comunicamos a usted que hemos revisado el trabajo de Tesis:

Puntos conjugados en control óptimo.

realizado por GERARDO SANCHEZ LICEA.

con número de cuenta 9140691-0 , pasante de la carrera de Actuaría.

Dicho trabajo cuenta con nuestro voto aprobatorio.

Atentamente

Director de Tesis Propietario	Dr. Javier Fernando Rosenblueth Laguette.	
Propietario	Dr. Ricardo Berlanga Zubiaga	
Propietario	M. en C. María Lourdes Velasco Arregui.	
Suplente	Dr. Luis Bernardo Morales Mendoza	
Suplente	M. en C. Apolinar Calderón Segura	

Consejo Departamental de
Matemáticas
M. en C. José Antonio Flores Díaz

FACULTAD DE CIENCIAS
CONSEJO DEPARTAMENTAL
DE
MATEMATICAS

A mis padres:

RUTH y CARLOS

Por su apoyo, comprensión,

y tolerancia durante mis estudios.

Con afecto al:

Dr Javier F Rosenblueth

Por su atención y su sencillez

A Wendy Nieva

Puntos Conjugados en Control Optimo

Gerardo Sánchez Licea

Contenido

1 Introducción 1

2 Cálculo de variaciones 3

2.1 Resultados Básicos 3

1 Espacio Euclidiano n-dimensional y Espacios Lineales 3
2 Funciones continuas y semicontinuas inferiores 5
3 Funciones de clase C^m, Mínimos locales de funciones 9
4 Teoremas de la Función Implícita 12
5 Teoremas de Existencia e Inclusión para Ecuaciones Diferenciales 13

2.2 Problemas clásicos de punto final fijo 14

1 Introducción 14
2 Terminología y resultados iniciales 15
3 Ejemplos 18
4 Prueba de la condición necesaria de Weierstrass 20
5 La primera y la segunda variaciones de I 21
6 La condición de Jacobi 24
7 Determinación de puntos conjugados 29
8 La positividad de la segunda variación 33

3 Control óptimo 37

1 Declaración del problema 37
2 Soluciones del problema 37
3 La primera y la segunda variaciones de I 38

3.1 Condiciones necesarias 40

1 La ecuación de Euler 41
2 La condición de Jacobi 45

3.2 Condiciones suficientes 47

1 Una prueba indirecta 48
2 Ejemplos 56

3 La condición modificada de Jacobi 57
4 Campos de extremos 60
5 La Ecuación de Hamilton Jacobi 65
6 La desigualdad de Hamilton Jacobi 67

4 La condición de punto conjugado 70

1 Direcciones admisibles y variaciones 72
2 El problema accesorio 74
3 Puntos conjugados y Ecuación de Riccati 75

5 Puntos conjugados generalizados 81

1 Condiciones necesarias de primer orden 81
2 Condiciones necesarias de segundo orden 81
3 Puntos conjugados generalizados 81

6 Comparación entre diferentes trabajos 87

1 Introducción.

Para el problema clásico del cálculo de variaciones, un conjunto de condiciones necesarias para minimizadores consiste en las condiciones de Euler, Weierstrass, Legendre y Jacobi. Bajo ciertas hipótesis, la primera es equivalente a la anulación de la primera variación sobre el conjunto de variaciones admisibles, mientras que la última es equivalente a la no negatividad de la segunda variación. También, el restringir las últimas tres convierte al conjunto en una serie de condiciones suficientes para minimizadores locales. En particular, la condición modificada de Jacobi es equivalente a la positividad de la segunda variación sobre el conjunto de variaciones admisibles no nulos. Un conjunto de condiciones suficientes para un mínimo débil consiste en la anulación de la primera variación, la positividad de la segunda variación, y la condición modificada de Legendre; estas condiciones junto con la condición modificada de Weierstrass son suficientes para obtener un mínimo fuerte. Este método se establece bajo una aplicación directa de la ecuación de Euler y la condición modificada de Jacobi a través de campos de extremos. Las condiciones de Euler, Weierstrass y Legendre se han generalizado, para problemas de control óptimo, a través del principio de Pontryagin, pero una laguna de gran importancia en la teoría, aún sin resolverse satisfactoriamente, ha sido el encontrar el equivalente a la condición de Jacobi, en particular en términos de " puntos conjugados ". Se trata de una condición de segundo orden y se han podido obtener resultados al respecto en control óptimo sólo para ciertos sistemas específicos. Más de una definición de punto conjugado en control óptimo se ha encontrado errónea y el problema sigue siendo discutido.

En la sección 3, para el problema de Lagrange involucrando dinámicas lineales, se introduce el equivalente a la ecuación de Euler, la condición de Legendre, Weierstrass, y Jacobi. Así, como el equivalente a las condiciones modificadas de Legendre, Weierstrass, y Jacobi. Se introduce la noción de punto conjugado [J.F.Rosenblueth] directamente de la ecuación de Euler, y bajo ciertas hipótesis se prueba que, la no existencia de un punto sobre (a, b) conjugado a a es necesaria para un óptimo. Posteriormente, se extiende de manera natural el conjunto clásico de condiciones suficientes del cálculo de variaciones. Estas condiciones, involucran la segunda variación con respecto al Hamiltoniano y son obtenidas directamente sin ninguna referencia a posibles extensiones de puntos conjugados, campos de extremos, soluciones de una cierta ecuación de Riccati, o la teoría de Hamilton-Jacobi. En otras palabras, la prueba de suficiencia es autocontenida en términos de control óptimo. Posteriormente, se caracteriza la positividad de la segunda variación en términos de la condición modificada de Jacobi, y se prueba que las condiciones modificadas de Legendre y Jacobi junto la condición de Euler son suficientes para un mínimo estricto débil; estas condiciones junto con la condición modificada de Weierstrass son suficientes para obtener un mínimo estricto fuerte.

Después, se presentan tres pruebas de suficiencia, basadas en teorías análogas del cálculo de variaciones, exhibiendo campos de extremos, una solución de la ecuación diferencial parcial de Hamilton-Jacobi, y una solución de la desigualdad de Hamilton-Jacobi.

En la sección 4, nuevamente para el problema de Lagrange con dinámicas lineales, se da otra noción de punto conjugado [Zeidan & Zezza]. Ahí bajo la hipótesis de que la solución óptima es " fuertemente normal " sobre cada intervalo de la forma $[a, c]$ y $[c, b]$ de $[a, b]$, se prueba la no existencia de un punto en (a, b) conjugado a a, lo cual es equivalente a la existencia de una solución sobre (a, b) de una cierta ecuación de Riccati. Esta teoría resulta ser equivalente a la teoría dada en la sección 3.

En la sección 5 se suavizan las hipótesis dadas en las secciones 3 y 4 considerablemente. Se define un punto conjugado generalizado [P.D.Lowen & H.Zheng] y se afirma que una condición necesaria para un óptimo es la no existencia de tales puntos. Si la condición de Legendre no se cumple, o si existe un punto conjugado en el sentido de las secciones 3 y 4, entonces habrá un punto conjugado generalizado de acuerdo a esta definición.

En las secciones 3 y 4, la teoría de puntos conjugados consiste en considerar el conocido " problema accesorio ", que consiste en minimizar la segunda variación $I''(y, v)$ sobre todas las parejas solución (y, v) de algún sistema lineal. Si (x, u) resuelve el problema original entonces $I''(y, v)$ es no negativa. Si algún v resuelve el problema accesorio, aplicando condiciones necesarias a este caso produce la no existencia de puntos cunjugados. En esta sección la teoría es completamente diferente. Suponiendo que un punto conjugado generalizado existe, se construye un control directamente para probar que la segunda variación debe ser negativa.

En la sección 6 se lleva a cabo un estudio comparativo, donde se analizan algunas diferencias fundamentales entre las distintas definiciones de " Punto Conjugado " conocidas en la literatura, que nos permiten encontrar deficiencias y posibles generalizaciones.

2 Cálculo de variaciones.

2.1 Resultados Básicos.

1 Espacio Euclidiano n-dimensional y Espacios Lineales.

Aquí estaremos interesados primordialmente en el estudio de funciones definidas sobre un conjunto S en un espacio Euclidiano n-dimensional $\mathcal{E}^n$, y en funciones definidas sobre un conjunto en un espacio lineal $\mathcal{E}$.

Por un espacio lineal $\mathcal{E}$ se entenderá un conjunto de elementos $x, y, z, \ldots$, llamados vectores, para los cuales la adición y la multiplicación por escalares $a, b, c, \ldots$ están definidas sujetas a las reglas usuales del análisis vectorial.

Las reglas básicas de las cuales se pueden obtener las demás son:

1. $x + y = y + x$, $\qquad x + (y + z) = (x + y) + z$.

2. Hay un elemento 0 tal que $x + 0 = x$ para todo x. Para todo x hay un elemento $-x$ tal que $x + (-x) = 0$.

3. $1 \cdot x = x$, $0 \cdot x = 0$, $a(bx) = (ab)x$,
$(a + b)x = ax + bx$, $a(x + y) = ax + ay$.

Aquí estaremos interesados solamente en el caso en que los escalares sean números reales.

Un conjunto de m vectores $x_1, \ldots, x_m$ en $\mathcal{E}$ se dice *linealmente independiente* si una relación de la forma

$$a_1 x_1 + \cdots + a_m x_m = 0$$

se cumple sólo si los escalares $a_1, \ldots, a_m$ son todos cero. Si hay n vectores $x_1, \ldots, x_n$ que forman conjuntos maximales de vectores linealmente independientes en $\mathcal{E}$ entonces se dice que $\mathcal{E}$ es n-dimensional. En este evento cada vector x en $\mathcal{E}$ se expresa de manera única en la forma

$$x = a_1 x_1 + \cdots + a_n x_n.$$

La clase de todas las n-tuplas $x = (x^1, \ldots, x^n)$, $y = (y^1, \ldots, y^n)$,...de números reales forman un espacio lineal n-dimensional con adición vectorial y multiplicación escalar definidas por las fórmulas

$$x + y = (x^1 + y^1, \ldots, x^n + y^n), \qquad ax = (ax^1, \ldots, ax^n).$$

La norma o longitud de un vector x se definirá por la fórmula

$$|x| = [(x^1)^2 + \cdots + (x^n)^2]^{1/2}$$

Si $|x| = 1$ entonces x es un vector unitario. Si $x \neq 0$, entonces $y = x/|x|$ es un vector unitario. La norma $|x|$ de x tiene las siguientes propiedades.

1. $|x| \geq 0$, cumpliéndose la igualdad si y sólo si $x = 0$;
2. $|ax| = |a|\,|x|$;
3. $|x + y| \leq |x| + |y|$ (desigualdad del triángulo).

Un espacio lineal $\mathcal{E}$ se dice un *espacio lineal normado* si para cada x en $\mathcal{E}$ existe un número real $|x|$ que tiene las propiedades que se acaban de describir.

Un espacio Euclidiano $\mathcal{E}^n$ es aquel cuyos puntos pueden ser representados por n-tuplas $x = (x^1, \ldots, x^n)$, $y = (y^1, \ldots, y^n)\ldots$ de números reales tal que la distancia entre dos puntos x y y está dada por la fórmula

$$|x - y| = \left[\sum_i (x^i - y^i)(x^i - y^i)\right]^{1/2}.$$

En consecuencia, $\mathcal{E}^n$ puede ser identificado con un espacio lineal normado n-dimensional. En $\mathcal{E}^n$ el *producto interno* de dos vectores x y y está definido por la fórmula

$$\langle x, y\rangle = \sum_1^n x^i y^i$$

Tenemos las relaciones

1. $\langle x, x\rangle \geq 0$, cumpliéndose la igualdad si y sólo si $x = 0$;
2. $\langle x, y\rangle = \langle y, x\rangle$;
3. $\langle ax + by, z\rangle = a\langle x, z\rangle + b\langle y, z\rangle$.

Además, $|x| = \langle x, x\rangle^{1/2}$, y

4. $|\langle x, y\rangle| \leq |x|\,|y|$.

Un espacio lineal $\mathcal{E}$ se dice un *espacio con producto interno* si para cada par de vectores x, y en $\mathcal{E}$ hay un escalar $\langle x, y\rangle$ llamado el producto interno o escalar de x y y, que tiene las tres propiedades descritas anteriormente.

Dado un conjunto de puntos S en $\mathcal{E}^n$, los puntos que no están en S forman un conjunto llamado el *complemento* de S. Un punto x se dice un *punto frontera* de S si toda bola alrededor de x contiene puntos de S y puntos de su complemento. Un punto x se dice un punto de acumulación de S si toda bola alrededor de x contiene puntos de S diferentes de x. Un *conjunto abierto* es aquel que no contiene a ninguno de sus puntos frontera, y un *conjunto cerrado* es aquel que contiene a todos sus puntos frontera. Por la *cerradura* de S se entiende el conjunto S junto con sus puntos frontera.

Por una *vecindad* de un punto x se entiende un conjunto abierto que contiene a x. Por una *ϵ-vecindad* de x se entiende el interior de una bola de radio ϵ alrededor de x. Por una vecindad de un conjunto S se entiende un conjunto abierto que contiene a S. Por una ϵ-vecindad de S se entiende el conjunto de puntos, cada uno de los cuales cae en una ϵ-vecindad de algún punto de S.

Una colección F de conjuntos abiertos se llama una *cubierta abierta* de un conjunto S si S es un subconjunto de su *unión*. Un conjunto S se dice *compacto* si toda cubierta abierta de S contiene un número finito de conjuntos los cuales cubren a S. Esta propiedad es llamada la propiedad de *Heine – Borel.*

Un conjunto acotado es aquel que está contenido en alguna bola alrededor del origen. Una propiedad fundamental de $\mathcal{E}^n$ es que un conjunto S es compacto si y sólo si éste es cerrado y acotado. En consecuencia, los conjuntos compactos en $\mathcal{E}^n$ pueden ser definidos como los conjuntos cerrados y acotados.

Un conjunto S se dice *convexo* si, dados dos puntos x, y en S, el segmento de línea $z = tx + (1-t)y$ $(0 \leq t \leq 1)$ está en S.

Por un *cono* $\mathcal{C}$ se entiende una colección de vectores x tal que si x está en $\mathcal{C}$ entonces también está ax para todo $a \geq 0$. Un cono $\mathcal{C}$ es convexo si $x + y$ está en $\mathcal{C}$ siempre y cuando x y y están en $\mathcal{C}$. Los múltiplos no negativos de un vector x pueden ser interpretados como mitades de líneas. Un cono es por lo tanto una colección de mitades de líneas.

Por un *subespacio lineal* $\mathcal{U}$ de $\mathcal{E}$ se entiende una clase de vectores que es cerrada bajo la adición vectorial y la multiplicación escalar. El número de vectores en un conjunto maximal de vectores linealmente independientes en $\mathcal{U}$ es llamado la dimensión de $\mathcal{U}$. Es fácil ver que un subespacio lineal es un cono convexo $\mathcal{U}$ tal que si x está en $\mathcal{U}$ entonces también está $-x$.

2 Funciones continuas y semicontinuas inferiores.

Sea f una función real definida sobre un subconjunto S de $\mathbf{R}^n$.

La máxima cota inferior m y la mínima cota superior M de f sobre S será denotada respectivamente por los símbolos

$$m = \inf f(x) = \inf f(x) \ \textit{sobre}\ S,$$

$$M = \sup f(x) = \sup f(x) \ \textit{sobre}\ S.$$

Admitimos los valores $\inf f(x) = -\infty$, $\sup f(x) = +\infty$. Sea x_0 un punto de acumulación de S. Entonces

$$b = \lim_{x \to x_0} f(x) = \lim_{x \to x_0} f(x) \ \text{ sobre } S$$

si para toda vecindad B de b hay una vecindad N de x_0, tal que $f(x)$ está en B para toda $x \neq x_0$ en N que pertenece a S. Si $b = +\infty$, los puntos y de la forma $y > c$, donde c es un número arbitrario, definen una vecindad de b.

Una vecindad de $b = -\infty$ está definida de manera similar. El límite inferior de f en x_0,

$$b = \liminf_{x \to x_0} f(x) = \liminf_{x \to x_0} f(x) \ \text{ sobre } S,$$

está definido por la fórmula

$$b = \lim_{\epsilon \to 0} [\inf f(x) \text{ sobre } N_\epsilon],$$

donde N_ϵ son todos los puntos $x \neq x_0$ de S en la ϵ-vecindad de x_0. De manera similar, el límite superior de f en x_0,

$$b = \limsup_{x \to x_0} f(x) = \limsup_{x \to x_0} f(x) \quad \text{sobre } S$$

está definido por

$$b = \lim_{\epsilon \to 0} [\sup f(x) \text{ sobre } N_\epsilon]$$

Sea x_0 un punto límite de S en S. La función f es *semicontinua inferior* en x_0 si

$$\liminf_{x \to x_0} f(x) \geq f(x_0),$$

y es *semicontinua superior* en x_0 si

$$\limsup_{x \to x_0} f(x) \leq f(x_0),$$

y es *continua* en x_0 si

$$\lim_{x \to x_0} f(x) = f(x_0).$$

Esta es semicontinua inferior, semicontinua superior, y continua sobre S si ésta tiene la propiedad en cada punto límite x_0 de S perteneciente a S. Es claro que si f es semicontinua inferior sobre S, la función $-f$ es semicontinua superior sobre S.

Tenemos lo siguiente

Teorema 2.1 *Sea f semicontinua inferior sobre una vecindad de un conjunto cerrado no nulo S. Si la desigualdad $f(x) > m$ se cumple sobre S, entonces ésta se cumple sobre una vecindad N de S. Si S es compacto, N puede ser seleccionada como una ϵ-vecindad de S.*

Prueba. Por la semicontinuidad inferior, cada punto x de S tiene una vecindad N_x de S sobre la cual $f(x) > m$. La unión N de estas vecindades N_x forma una vecindad de S en la cual esta desigualdad se cumple. Si S es compacto, N contiene una ϵ-vecindad de S para algún $\epsilon > 0$ y N puede ser reemplazada por esta ϵ-vecindad.

Teorema 2.2 *Sea f semicontinua inferior sobre un conjunto compacto no nulo S. Existe un punto x_0 en S tal que la desigualdad*

$$f(x_0) \leq f(x)$$

se cumple sobre S. El valor $f(x_0)$ es el mínimo de f sobre S.

Prueba. Sea $m = \inf f(x)$ sobre S. Sea S_k el conjunto de puntos x sobre S tal que

$$f(x) \leq c_k$$

donde $c_k > c_{k+1} > m$ y $\lim c_k = m$. Sea x en la frontera de S_k. Si x está en el complemento de S_k, entonces $f(x) > c_k$ y como f es semicontinua inferior sobre S, entonces $f(y) > c_k$ para todo $y \in N_\epsilon(x)$, donde $N_\epsilon(x)$ es una $\epsilon -$ *vecindad* de x. Pero como x está en la frontera de S_k existe $z \in S_k \cap N_\epsilon(x)$. Lo cual es una contradicción. Por lo tanto, los conjuntos S_k son subconjuntos cerrados no vacíos de S. Puesto que S_{k+1} es un subconjunto de S_k, y como S es compacto, la intersección de estos conjuntos contiene un punto x_0 en S. Puesto que x_0 está en S_k, tenemos que

$$m \leq f(x_0) \leq c_k \qquad (k = 1, 2, 3, \ldots)$$

y en consecuencia $f(x_0) = m$.

Corolario 2.1 *Sea f una función semicontinua inferior sobre S y sea $m = \inf f(x)$ sobre S. Si existe una constante $c > m$ tal que el conjunto S_c de puntos x sobre S que tienen $f(x) \leq c$ es compacto, entonces hay un punto x_0 tal que $f(x_0) = m$.*

Prueba. Este resultado se sigue del teorema anterior con S reemplazada por S_c.

Teorema 2.3 *Si f es continua sobre un conjunto compacto no nulo S, entonces existen puntos x_0 y x_1 en S tal que la desigualdad*

$$f(x_0) \leq f(x) \leq f(x_1)$$

se cumple sobre S. Los valores $f(x_0)$ y $f(x_1)$ son respectivamente los valores mínimo y máximo de f sobre S.

Prueba. Este resultado se sigue del Teorema 2.2 escogiendo a x_0 relacionado a f y x_1 relacionado a $-f$ como se describió en el Teorema 2.2.

Los Teoremas 2.2 y 2.3 son teoremas fundamentales de existencia para el máximo y el mínimo de funciones.

Definición 2.1 *Una función f se dice positivamente homogénea de grado k sobre un conjunto S, si $x \in S$ y α un número positivo, implica que $\alpha x \in S$ y*

$$f(\alpha x) = \alpha^k f(x) \qquad (\alpha > 0) \tag{1}$$

Si $f(x)$ es positivamente homogénea de grado k sobre S, la función

$$g(x) = f(x)/|x|^m$$

es positivamente homogénea de grado $k - m$ sobre S. Por este hecho nos restringiremos al caso en el que $k > 0$. En este evento podemos suponer que $x = 0$ está en S y que $f(0) = 0$. Se supondrá que S contiene un punto $x \neq 0$.

Teorema 2.4 *Sea f una función positivamente homogénea de grado $k > 0$ cuyo dominio S es un cono cerrado. Supongamos que f es continua sobre S. Existen vectores unitarios x_0 y x_1 en S tal que la desigualdad*

$$f(x_0)|x|^k \leq f(x) \leq f(x_1)|x|^k \tag{2}$$

se cumple sobre S.

Prueba. Sea S_1 el conjunto de vectores unitarios en S. Este conjunto es no vacío y es compacto. En virtud del Teorema 2.3 existen puntos x_0 y x_1 en S_1 tal que

$$f(x_0) \leq f(x) \leq f(x_1)$$

sobre S_1, y en consecuencia tal que

$$f(x_0) \leq f\left(\frac{x}{|x|}\right) \leq f(x_1)$$

para todo $x \neq 0$ en S. Usando (1) con $\alpha = 1/|x|$, obtenemos (2). Puesto que $f(0) = 0$, la relación también se cumple en $x = 0$.

Definición 2.2 *Una forma cuadrática Q es una función de la forma*

$$Q(x) = \langle Ax, x\rangle = \langle x, Ax\rangle,$$

donde A es una matriz simétrica.

Es claro que Q es una función homogénea de grado 2. La función

$$Q(x, y) = \langle Ax, y\rangle = \langle x, Ay\rangle$$

es llamada la *forma bilineal asociada*. Tenemos la identidad

$$Q(\alpha x + \beta y) = \alpha^2 Q(x) + 2\alpha\beta Q(x, y) + \beta^2 Q(y).$$

Los términos *positiva*, *negativa*, *no positiva*, y *no negativa*, son usados para Q o para A de acuerdo a como las relaciones

$$Q(x) > 0, \quad Q(x) < 0, Q(x) \leq 0, Q(x) \geq 0$$

se cumplen para todo $x \neq 0$.

En vista del Teorema 2.4 una forma cuadrática Q tiene vectores unitarios asociados x_0 y x_1 tal que

$$Q(x_0)|x|^2 \leq Q(x) \leq Q(x_1)|x|^2.$$

Se sigue que si Q es positiva, hay un número $m > 0$, a saber $m = Q(x_0)$, tal que

$$Q(x) \geq m|x|^2 \qquad (m > 0)$$

se cumple para toda x. Una extensión de este resultado está dada en el siguiente.

Teorema 2.5 *Sea $A(z)$ una matriz simétrica cuyos elementos son funciones continuas de $z = (z^1, \ldots, z^m)$ sobre una vecindad de un conjunto compacto S. Sea*

$$Q(z; x) = \langle A(z)x, x\rangle.$$

Si $Q(z;x) > 0$ para todo $x \neq 0$ y para todo z en S, hay una vecindad N de S y un número positivo m tal que la desigualdad

$$Q(z; x) \geq m|x|^2 \tag{3}$$

se cumple para todo x y para todo z en N.

Prueba. El conjunto de todos los puntos (x, z) con $|x| = 1$ y z en S es un conjunto compacto sobre el cual $Q(z; x)$ es positivo. Por el Teorema 2.1, $Q(z; x)$ es positiva sobre una vecindad de este conjunto, y en consecuencia sobre un conjunto compacto de la forma

$$|x| = 1, \qquad z \in S_\epsilon \qquad (\epsilon > 0),$$

donde S_ϵ es la cerradura de una ϵ-vecindad de S. El mínimo m de $Q(z; x)$ sobre este conjunto es por lo tanto positivo. En consecuencia

$$Q\left(z, \frac{x}{|x|}\right) \geq m \qquad (x \neq 0, z \in S_\epsilon)$$

y la relación (3) se cumple, con N como la ϵ-vecindad de S.

3 Funciones de clase C^m, Mínimos locales de funciones.

Una función real f es de clase C^m sobre un conjunto abierto S si ésta es continua y todas las derivadas parciales hasta la de orden m son continuas. Se dice que es de clase C^m sobre un conjunto arbitrario S si ésta es de clase C^m sobre una vecindad de S. Usaremos las siguientes notaciones,

$$f_x = \frac{\partial f}{\partial x}, \qquad f_{xx} = \frac{\partial^2 f}{\partial x \partial x}, \ldots,$$

para las derivadas parciales de f. Como se verá después, será conveniente introducir las notaciones

$$f'(x, h) = f^{(1)}(x, h) = \langle f_x, h\rangle$$

$$f''(x, h) = f^{(2)}(x, h) = \langle f_{xx}h, h\rangle$$

para los correspondientes diferenciales de f.

Si f es de clase C^1, sobre S, entonces en un punto x en S la función

$$g(t) = f(x + th),$$

donde h es un vector dado, es de clase C^1 en t sobre un intervalo $-\epsilon \leq t \leq \epsilon$ y

$$g'(t) = f'(x + th, h), \qquad g'(0) = f'(x, h).$$

Si f es de clase C^2, entonces también es g y

$$g''(t) = f''(x + th, h), \qquad g''(0) = f''(x, h).$$

La función $f'(x, h)$ puede ser interpretada como la derivada de f en la dirección h. De manera similar, $f''(x, h)$ puede ser interpretada como la segunda derivada de f en la dirección h. Si f es de clase C^1, tenemos las familiares fórmulas de Taylor

$$f(x + h) = f(x) + f'(x + t_1 h, h) = f(x) + \int_0^1 f'(x + th, h)dt$$

siempre que el segmento de línea $x + th$ $(0 \leq t \leq 1)$ esté en S y t_1 es un punto seleccionado adecuadamente sobre $0 < t < 1$. De manera similar, para las funciones de clase C^2 tenemos

$$\begin{aligned} f(x+h) &= f(x) + f'(x,h) + \frac{1}{2} f''(x + t_2 h, h) \qquad (4) \\ &= f(x) + f'(x,h) + \int_0^1 (1-t) f''(x + th, h)dt, \end{aligned}$$

donde t_2 es algún punto sobre $0 < t < 1$. Otra vez el segmento de línea $x + th$ $(0 \leq t \leq 1)$ debe caer en S.

El vector definido por las derivadas parciales de primer orden de f será llamado el gradiente de f en x y está denotado por los símbolos grad$f$ o $f'(x)$

Por lo tanto

$$\text{grad} f = f'(x) = \left(\frac{\partial f}{\partial x}\right).$$

De acuerdo a esto tenemos en términos de productos internos

$$f'(x, h) = \langle f'(x), h \rangle.$$

Asimismo, denotamos el Hessiano de f por $f''(x)$. En consecuencia $f''(x)$ es la matriz simétrica

$$f''(x) = \left(\frac{\partial^2 f(x)}{\partial x \partial x}\right).$$

En términos de productos internos

$$f''(x, h) = \langle f''(x)h, h \rangle = \langle h, f''(x)h \rangle.$$

Una función f tiene un mínimo local en un punto x_0 en S si existe una vecindad N de x_0 tal que la desigualdad

$$f(x) \geq f(x_0)$$

se cumple para todo x en S que cae en N. Si la igualdad puede ser excluida cuando $x \neq x_0$, entonces x_0 aporta un mínimo local propio a f sobre S.

Teorema 2.6 *Sea f una función de clase C^2 sobre S. Si un punto interior x_0 aporta un mínimo local a f, entonces*

$$f'(x_0) = 0, \qquad f''(x_0) \geq 0. \tag{5}$$

Recíprocamente, si las relaciones

$$f'(x_0) = 0, \qquad f''(x_0) > 0 \tag{6}$$

se cumplen en un punto x_0 en S, entonces existe una vecindad N de x_0 y un número positivo m, tal que

$$f(x) \geq f(x_0) + m|x - x_0|^2 \tag{7}$$

se cumple para todo x en N.

Prueba. Si x_0 minimiza f localmente, la función

$$g(t) = f(x_0 + th)$$

tiene un mínimo local en $t = 0$. En consecuencia

$$g'(0) = f'(x_0, h) = 0, \qquad g''(0) = f''(x_0, h) \geq 0$$

para todo h y (5) se cumple. Recíprocamente, si (6) se cumple, entonces por el Teorema 2.5 existe una vecindad N de x_0 y un número positivo m tal que

$$f''(x, h) \geq 2m|h|^2$$

para todo x en N. Si x está en N se sigue de (4) con x reemplazada por x_0 y con $h = x - x_0$ que (7) se cumple.

Un punto en que $f'(x) = 0$ es llamado un *punto crítico* de f y el valor de f es llamado un *valor crítico*.

Como una aplicación posterior a formas cuadráticas y matrices simétricas tenemos el siguiente

Teorema 2.7 *Una matriz simétrica no negativa A es positiva si y sólo si ésta es no singular.*

Prueba. Sea

$$F(x) = \langle Ax, x \rangle.$$

Si A es singular, hay un vector $x_0 \neq 0$ tal que $Ax_0 = 0$. En consecuencia $F(x_0) = 0$ y A es no positiva. Recíprocamente, si hay un punto $x_0 \neq 0$ tal que $F(x_0) = 0$, entonces x_0 aporta un mínimo a F ya que $F(x) \geq 0$, por hipótesis. Entonces,

$$F'(x_0) = 2Ax_0 = 0$$

y A es singular.

4 Teoremas de la Función Implícita.

Sea $f : \mathbf{R}^{m+n} \to \mathbf{R}^n$ una función continua sobre un conjunto abierto S en el espacio-tx, con derivadas parciales continuas con respecto a $x^1, \ldots, x^n$ sobre S.

Consideremos la ecuación

$$f(t,x) = f(t^1, \ldots, t^m, x^1, \ldots, x^n) = 0,$$

y sea $D(t,x)$ el determinante

$$D(t,x) = \left|\frac{\partial f}{\partial x}\right|.$$

Con estas hipótesis podemos establecer el siguiente

Teorema 2.8 *Supongamos que las relaciones*

$$f(t_0, x_0) = 0, \qquad D(t_0, x_0) \neq 0$$

se cumplen en el punto (t_0, x_0) en S. Existe una función continua

$$x(t) = x(t^1, \ldots, t^m)$$

sobre una vecindad T de t_0 y una constante $\epsilon > 0$ tal que

$$x(t_0) = x_0, \qquad f(t, x(t)) = 0$$

y tal que las relaciones

$$f(t,x) = 0, \qquad |x - x(t)| < \epsilon \qquad (t \in T) \tag{8}$$

se cumplen sólo si $x = x(t)$. Si la función f es de clase C^m sobre S, la función $x(t)$ es de clase C^m sobre T.

Este resultado es un caso especial de el siguiente

Teorema 2.9 *Supongamos que existe una función continua*

$$x_0(t)$$

sobre un conjunto compacto T_0 tal que si t está en T_0, entonces $(t, x_0(t))$ está en S, y

$$f(t, x_0(t)) = 0, \qquad D(t, x_0(t)) \neq 0.$$

Existe una función continua

$$x(t)$$

definida sobre una vecindad T de T_0 y una constante $\epsilon > 0$ tal que

$$x(t) = x_0(t) \ (t \in T_0), \qquad f(t, x(t)) = 0 \ (t \in T),$$

y tal que (8) se cumple sólo en el caso en que $x = x(t)$. Además, si la función f es de clase C^m sobre S, la función $x(t)$ es de clase C^m sobre T.

Ver Hestenes MR(1966) Calculus of Variations and Optimal Control Theory, John Wiley.

5 Teoremas de Existencia e Inclusión para Ecuaciones Diferenciales.

En el problema de minimizar una integral en ocasiones haremos uso de teoremas de la teoría de ecuaciones diferenciales que serán útiles frecuentemente.

Sea

$$f(t,x) = f(t,x^1,\dots,x^n)$$

una función continua sobre una región $\mathcal{F}$ en el espacio-tx la cual posee derivadas parciales continuas con respecto a x^i sobre $\mathcal{F}$. Como un primer teorema de existencia para la ecuación diferencial

$$\dot{x} = f(t,x) \qquad \left(\dot{x} = \frac{dx}{dt}\right) \tag{9}$$

tenemos el siguiente

Teorema 2.10 *Dado un punto (α,β) en $\mathcal{F}$ hay una constante $\delta > 0$ tal que existe una única solución*

$$x(t) \qquad (\alpha-\delta < t < \alpha+\delta)$$

de la ecuación (9) con $x(\alpha) = \beta$. Si la función f es de clase C^m sobre $\mathcal{F}$, entonces la función $x(t)$ es de clase C^{m+1} sobre $\alpha-\delta < t < \alpha+\delta$.

Si

$$x_0 : \qquad x_0(t) \qquad (a \le t \le b)$$

es una solución de la ecuación diferencial (9), entonces ésta puede ser incluida en una familia n-paramétrica de soluciones de esta ecuación. Este resultado puede ser obtenido manteniendo α fija en el resultado dado en el siguiente teorema.

Teorema 2.11 *Existe una constante $\rho > 0$ y una vecindad $\mathcal{F}_0$ de los puntos (t,x) sobre una solución x_0 de la ecuación diferencial (9) tal que a través de cada punto $(\alpha,\beta) = (\alpha,\beta^1,\dots,\beta^n)$ en $\mathcal{F}_0$ pasa una única solución*

$$x(t,\alpha,\beta) \qquad (a-\rho \le t \le b+\rho)$$

de la ecuación (9) tal que

$$x(t,\alpha,x_0(\alpha)) = x_0(t) \qquad (a \le t \le b).$$

Las funciones $x(t,\alpha,\beta)$, $\dot{x}(t,\alpha,\beta)$ son continuas y poseen derivadas continuas con respecto a β_j para todo (t,α,β) con t sobre el rango $a-\rho \le t \le b+\rho$ y (α,β) en $\mathcal{F}_0$. Además, el determinante

$$\left|\frac{\partial x}{\partial \beta}(t,\alpha,\beta)\right|$$

es diferente de cero sobre este conjunto. Si la función $f(t,x)$ es de clase C^m sobre $\mathcal{F}_0$ entonces las funciones $x(t,\alpha,\beta)$, $\dot{x}(t,\alpha,\beta)$ son de clase C^m.

Teorema 2.12 *Para ϵ en el rango $-\delta < \epsilon < \delta$ sea*

$$x(\epsilon): \qquad x(t,\epsilon) \qquad (a \le t \le b)$$

una familia uniparamétrica de soluciones de la ecuación (9) que contiene a x_0 para $\epsilon = 0$. Supongamos que las funciones $x(t,\epsilon)$ y sus derivadas parciales x_ϵ con respecto a ϵ son continuas sobre $a \le t \le b$, $-\delta < \epsilon < \delta$. La variación

$$y: \qquad y(t) = x_\epsilon(t,0) \qquad (a \le t \le b)$$

de la familia $x(\epsilon)$ a lo largo de x_0 satisface la ecuación de variación

$$\dot{y} = \langle f_x(t, x_0(t)), y\rangle \tag{10}$$

de la ecuación (9) a lo largo de x_0.

Prueba. La ecuación (10) se obtiene de la ecuación

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

derivando con respecto a ϵ y evaluando en $\epsilon = 0$.

2.2 Problemas clásicos de punto final fijo.

1 Introducción.

Consideremos el problema de encontrar en una clase de arcos

$$x: \ x(t) \qquad (t_1 \le t \le t_2)$$

que unen dos puntos fijos x_0 y x_1 en el espacio tx uno que minimice una integral de la forma

$$I(x) = \int_{t_1}^{t_2} f(t, x(t), \dot{x}(t))dt$$

Si, por ejemplo, $f = \sqrt{1+\dot{x}^2}$, entonces $I(x)$ es la longitud de x. Si $f = 2\pi x\sqrt{1+\dot{x}^2}$ y $x > 0$, entonces $I(x)$ representa el área de la superficie obtenida girando x alrededor del eje t. En el caso $(n+1)-$*dimensional* el problema es encontrar en una clase de arcos

$$x: \ x^i(t) \qquad (a \le t \le b\ ;\ i = 1, \ldots, n)$$

que unen dos puntos fijos en el *espacio* $-(t, x^1, \ldots, x^n)$ uno que minimice una integral

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

$$donde \qquad \dot{x}(t) = \frac{dx(t)}{dt}$$

Aquí $f(t, x, \dot{x})$ es una función de $2n+1$ variables $(t, x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n)$. El problema que se acaba de definir es llamado el problema de punto final fijo simple. Esta sección es dedicada principalmente al desarrollo de condiciones necesarias de primer orden para un mínimo. Estas primero son dadas en la forma estándar involucrando la ecuación de Euler y la condición de Weierstrass.

2 Terminología y resultados iniciales.

A una función real $u(t)$ se le dirá continua por fragmentos en un intervalo $a \leq t \leq b$ si el intervalo puede ser dividido en un número finito de subintervalos $t_j \leq t \leq t_{j+1}$ $(j = 1 \ldots, N)$ con $t_1 = a$ y $t_{N+1} = b$, sobre cada uno de los cuales $u(t)$ es continua y tiene límites laterales izquierdo y derecho $u(t_j + 0)$ y $u(t_j - 0)$.

Sea $a \leq t \leq b$ un intervalo fijo. Denotemos por $\mathcal{E}$ la clase de todos los arcos continuos

$$x : \ x^i(t) \qquad (a \leq t \leq b \ ; \ i = 1, \ldots, n)$$

que poseen derivadas continuas por fragmentos $\dot{x}^i(t)$ sobre $a \leq t \leq b$. Si x, y son arcos en $\mathcal{E}$, y α, β son números reales, $\alpha x + \beta y$ está también en $\mathcal{E}$. De acuerdo a esto la clase $\mathcal{E}$ es un espacio lineal. La subclase $\mathcal{E}_0$ de todos los arcos y en $\mathcal{E}$ tal que $y(a) = y(b) = 0$ es una subclase lineal de $\mathcal{E}$.

Sea R un conjunto abierto de puntos $(t, x, \dot{x}) = (t, x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n)$ en el espacio $(2n+1)-$*dimensional*. Los elementos de R serán llamados elementos admisibles. Un arco x en $\mathcal{E}$ se le dirá admisible si sus elementos $(t, x(t), \dot{x}(t))$ son admisibles.

Sea $f(t, x, \dot{x})$ una función real de clase C^2 sobre R. La integral

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

está bien definida sobre todo arco admisible x. El problema que se considerará en esta sección es el de encontrar propiedades de un arco admisible x_0 que minimice I en la clase $\mathcal{R}$ de todos los arcos admisibles que unen sus puntos finales. Uno de los principales resultados está dado en el Teorema 2.13. Su demostración se dará en las siguientes subsecciones. Para propósitos de aplicaciones posteriores se debe notar que la demostración dada no requiere que f sea de clase C^2. Las conclusiones del teorema se cumplen si f tiene derivadas parciales continuas con respecto a x y a $\dot{x}$.

Teorema 2.13 (Primera condición necesaria) *Supongamos que x_0 minimiza I en $\mathcal{R}$. Entonces existe un vector constante c tal que la relación*

$$f_{\dot{x}} = \int_a^t f_x ds + c \tag{11}$$

se cumple a lo largo de x_0. Además, la desigualdad

$$E(t, x, \dot{x}, u) \geq 0 \tag{12}$$

se cumple para todo $(t, x, \dot{x}, u)$ tal que $(t, x, \dot{x})$ está sobre x_0 y (t, x, u) está en R, donde E es la función E de Weierstrass

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

La afirmación de que (11) se cumple a lo largo de x_0 quiere decir la relación

$$f_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_a^t f_x(s, x_0(s), \dot{x}_0(s))ds + c$$

Los elementos $(t, x, \dot{x})$ sobre x_0 son los elementos $(t, x_0(t), \dot{x}_0(t))$. En un punto esquina de x_0 existen dos elementos, $(t, x_0(t), \dot{x}_0(t-0))$ y $(t, x_0(t), \dot{x}_0(t+0))$. La ecuación (11) es la forma integral de la ecuación de Euler

$$\frac{d}{dt} f_{\dot{x}} = f_x \tag{13}$$

En un punto esquina de x_0 la derivada d/dt será interpretada como una derivada lateral izquierda y derecha.

Corolario 2.2 (La condición de esquina) *Si x_0 satisface la primera condición necesaria, las funciones*

$$f - \langle \dot{x}, f_{\dot{x}}\rangle, \qquad y \qquad f_{\dot{x}}$$

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

Prueba. Si definimos $p(t) = f_{\dot{x}}(t, x_0(t), \dot{x}_0(t))$ tenemos por (11) la condición de esquina $p(t-0) = p(t+0)$. Esta condición es comúnmente llamada la condición de esquina de $Weierstrass - Erdmann$. La continuidad de $f - \langle \dot{x}, f_{\dot{x}}\rangle$ es una consecuencia de la condición (12) y la continuidad de $p(t)$. Para obtener este resultado definimos

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

Por (12) y la continuidad de $p(t)$ se tiene

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

$$
\begin{aligned}
&= f(t, x_0(t), \dot{x}_0(t+0)) - f(t, x_0(t), \dot{x}_0(t-0)) \\
&\qquad -\langle \dot{x}_0(t+0) - \dot{x}_0(t-0), f_{\dot{x}}(t, x_0(t), \dot{x}_0(t-0))\rangle \\
&= f(t, x_0(t), \dot{x}_0(t+0)) - f(t, x_0(t), \dot{x}_0(t-0)) \\
&\qquad -\langle \dot{x}_0(t+0) - \dot{x}_0(t-0), f_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\rangle \\
&= \{f(t, x_0(t), \dot{x}_0(t+0)) - \langle \dot{x}_0(t+0), f_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\rangle\} \\
&\qquad -\{f(t, x_0(t), \dot{x}_0(t-0)) - \langle \dot{x}_0(t-0), f_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\rangle\} \\
&= F(t, \dot{x}_0(t+0)) - F(t, \dot{x}_0(t-0))
\end{aligned}
$$

De manera similar,

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

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

Así se ha probado el siguiente

Corolario 2.3 *En un punto esquina de un arco admisible x_0 que satisface la primera condición necesaria, se tiene*

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

Corolario 2.4 (La condición de Legendre) *En cada elemento $(t, x, \dot{x})$ de un arco x_0 que satisface la condición de Weierstrass, se tiene que*

$$\langle f_{\dot{x}\dot{x}}\pi, \pi\rangle \geq 0 \tag{14}$$

para todo $\pi = (\pi^1, \ldots, \pi^n)$, i.e., la matriz $f_{\dot{x}\dot{x}}$ es semidefinida positiva.

Prueba. Sea $(t, x, \dot{x})$ un elemento sobre x_0, y definimos $G(u) = E(t, x, \dot{x}, u)$, observamos que $G(\dot{x}) = 0$, $G_u = f_{\dot{x}}(t, x, u) - f_{\dot{x}}(t, x, \dot{x})$, y $G_{uu} = f_{\dot{x}\dot{x}}(t, x, u)$, en virtud de la condición de Weierstrass la función $G(u)$ tiene un mínimo local en $u = \dot{x}$, de esto se sigue que $0 \leq \langle G_{uu}(\dot{x})\pi, \pi\rangle = \langle f_{\dot{x}\dot{x}}(t, x, \dot{x})\pi, \pi\rangle$ para todo π.

Una solución x_0 de (11) será llamada un extremal. Un extremal sin esquinas será llamado un extremo. Un extremal está formado por un número finito de subarcos que son extremos. Un arco admisible x_0 se dice no singular si el determinante

$$|f_{\dot{x}\dot{x}}| \neq 0$$

en cada elemento $(t, x, \dot{x})$ sobre x_0.

Corolario 2.5 (Teorema de diferenciabilidad de Hilbert) . *Un extremo no singular x_0 es de clase C^m si f es de clase C^m $(m \geq 2)$ sobre R. Entre dos esquinas, un arco x_0 no singular que minimiza a I es de clase C^m.*

Prueba. Observamos que la ecuación

$$0 = F(t,u) = f_{\dot{x}}(t, x_0(t), u) - \int_a^t f_x(s, x_0(s), \dot{x}_0(s))ds - c \tag{15}$$

tiene a $u(t) = \dot{x}_0(t)$ como solución. Además

$$\left|\frac{\partial F}{\partial u}\right| = |f_{\dot{x}\dot{x}}| \neq 0$$

a lo largo de esta solución, ya que x_0 es no singular. Sea f de clase C^m ($m \geq 2$) sobre R. Si x_0 no tiene una esquina en $t = \bar{t}$, la función $F(t,u)$ es de clase C^k ($k < m$), donde C^k es la clase de $x_0(t)$ cerca de $t = \bar{t}$. Tenemos que $k \geq 1$. Por el teorema de la función implícita, la solución $u(t) = \dot{x}_0(t)$ de (15) es también de clase C^k. En consecuencia, si $x_0(t)$ es de clase C^k ($k < m$), éste es también de clase $C^{(k+1)}$. De esto se sigue que $x_0(t)$ es de clase C^m.

La integral I se dirá positivamente regular si la región R es convexa en $\dot{x}$ y la matriz $f_{\dot{x}\dot{x}}$ es definida positiva para todo $(t, x, \dot{x})$ en R. En vista de la convexidad de R en $\dot{x}$, tenemos por el Teorema de Taylor aplicado a $f(t, x, \dot{x})$ considerada como una función de $\dot{x}$,

$$E(t, x, \dot{x}, \dot{x} + \pi) = 1/2\langle f_{\dot{x}\dot{x}}(t, x, v)\pi, \pi\rangle$$

donde $v = \dot{x} + \xi\pi$ para algún valor ξ sobre $0 < \xi < 1$.

De este resultado se obtiene la primera afirmación en el siguiente

Corolario 2.6 *Sea I positivamente regular, entonces*

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

para todo $(t, x, \dot{x}, u)$ tal que $(t, x, \dot{x})$ y (t, x, u) están en R y $u \neq \dot{x}$. Además, todo extremal es un extremo no singular.

Prueba. Por el Corolario 2.3 si $E(t, x, \dot{x}, u) > 0$ para todo $(t, x, \dot{x}, u)$ tal que $(t, x, \dot{x})$ y (t, x, u) están en R y $u \neq \dot{x}$, entonces $x(t)$ no tiene esquinas. Además, como $f_{\dot{x}\dot{x}} > 0$, entonces $|f_{\dot{x}\dot{x}}| \neq 0$ por lo tanto $x(t)$ es no singular.

3 Ejemplos.

Consideremos el caso en el que $f = \dot{x}^2 - x^2$. En este evento

$$\frac{d}{dt} f_{\dot{x}} - f_x = 2(\ddot{x} + x) = 0$$

La solución es $x = \alpha cos(t) + \beta sen(t)$. La función *$E$ de Weierstrass* toma la forma

$$E(t, x, \dot{x}, u) \;=\; f(t, x, u) - f(t, x, \dot{x})$$

$$
\begin{aligned}
& -\langle u-\dot{x}, f_{\dot{x}}(t,x,\dot{x})\rangle \\
= \;& u^2 - x^2 - \dot{x}^2 + x^2 - (u-\dot{x})2\dot{x} \\
= \;& u^2 - \dot{x}^2 - 2u\dot{x} + 2\dot{x}^2 \\
= \;& u^2 - 2u\dot{x} + \dot{x}^2 \\
= \;& (u-\dot{x})^2.
\end{aligned}
$$

Por lo tanto la condición de Weierstrass es satisfecha.

El arco

$$x_0 : \; x_0 \equiv 0 \qquad (0 \leq t \leq b)$$

satisface las condiciones descritas en el Teorema 2.13. Sin embargo, si $b > \pi$, el arco x_0 no minimiza a $I(x) = \int_0^b (\dot{x}^2 - x^2)dt$, en la clase de todos los arcos que cumplen con $x(0) = x(b) = 0$. Tomemos c sobre el intervalo $\pi < c \leq b$, $c < 2\pi$, y sea x el arco

$$x(t) = sen(t) \; (0 \leq t \leq \frac{c}{2}), \;\; x(t) = sen(c-t) \; (\frac{c}{2} \leq t \leq c), \;\; x(t) \equiv 0 \; (c \leq t \leq b).$$

El valor de la integral I para esta función es

$$
\begin{aligned}
I(x) &= \int_0^{c/2} (cos^2(t) - sen^2(t))dt + \int_{c/2}^{c} (cos^2(c-t) - sen^2(c-t))dt \\
&= 2\int_0^{c/2} (cos^2(t) - sen^2(t))dt \\
&= 2\int_0^{c/2} cos(2t)dt \\
&= sen(c) < 0
\end{aligned}
$$

ya que $\pi < c < 2\pi$. Como $I(x_0) = 0$, el arco x_0 no minimiza a I sujeto a la restricción $x(0) = x(b) = 0$ si $b > \pi$. Este ejemplo muestra que las condiciones descritas en el Teorema 2.13 no son suficientes para garantizar un mínimo a I. Se debe notar que un problema variacional no siempre tiene solución en la clase dada de arcos bajo consideración. Para ilustrar este hecho consideremos el problema de minimizar la integral

$$I(x) = \int_0^1 t^2 \dot{x}(t)^2 dt$$

en la clase $\mathcal{R}$ de arcos $x : x(t)$ $(0 \leq t \leq 1)$, que cumplen con $x(0) = 0$ y $x(1) = 1$. Los arcos bajo consideración se suponen ser continuos y tener derivadas continuas por fragmentos. Claramente $I(x) \geq 0$ para todo x en $\mathcal{R}$. Para ϵ en $0 < \epsilon \leq 1$ el arco

$$x : \; x(t) = \frac{t}{\epsilon} \; (0 \leq t \leq \epsilon), \qquad x(t) = 1 \; (\epsilon \leq t \leq 1),$$

claramente tiene $x(0) = 0$ y $x(1) = 1$, por lo tanto x está en $\mathcal{R}$ y

$$I(x) = \int_0^1 t^2 \dot{x}(t)^2 dt = \int_0^\epsilon \frac{t^2}{\epsilon^2} dt \; = \frac{\epsilon}{3}$$

En consecuencia $m = 0$ es la máxima cota inferior de $I(x)$ sobre $\mathcal{R}$. Además, las ecuaciones de Euler de I toman la forma

$$\frac{d}{dt}(2t^2\dot{x}) = 0$$

entonces,

$$t^2\dot{x} = c = constante$$

a lo largo de cada extremo. Tomando $t = 0$, se ve que $c = 0$. Consecuentemente, los únicos extremos definidos sobre el intervalo cerrado $(0 \le t \le 1)$ son aquellos para los cuales $x(t) = constante$. Ningún extremo de este tipo tiene $x(0) = 0$ y $x(1) = 1$. Entonces el problema variacional descrito anteriormente no tiene solución. Si se relajan las soluciones y se permite a uno de los valores finales de $x(t)$ ser libre, entonces existe una solución. Notamos que en este problema $f_{\dot{x}\dot{x}} = 2t^2$ y por lo tanto se tiene una singularidad en $t = 0$.

4 Prueba de la condición necesaria de Weierstrass.

Sea x_0 un arco admisible que minimiza I en la clase de arcos admisibles que unen sus puntos finales. La desigualdad $E(t, x, \dot{x}, u) \ge 0$ se cumple para todo $(t, x, \dot{x}, u)$ tal que $(t, x, \dot{x})$ está sobre x_0 y (t, x, u) está en R. Consideremos un punto $\bar{t}$ sobre $a < t < b$ el cual no corresponde a un punto esquina de x_0. Definimos $\bar{x} = x_0(\bar{t})$, $\bar{u} = \dot{x}_0(\bar{t})$, y escogemos u tal que $(\bar{t}, \bar{x}, u)$ sea admisible. Escogemos $\delta_0 > 0$ tal que $\bar{t} + \delta_0 < b$. Para $0 < \delta \le \delta_0$, $0 \le \epsilon < 1$, sea

$$x(\epsilon, \delta) : \qquad x(t, \epsilon, \delta) \qquad (a \le t \le b)$$

el arco definido por las fórmulas

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

Puesto que R es un conjunto abierto podemos seleccionar ϵ_0 lo suficientemente pequeño tal que, siempre y cuando $0 \le \epsilon \le \epsilon_0$, $0 < \delta \le \delta_0$, el arco $x(\epsilon, \delta)$ sea un arco admisible que une los puntos finales de x_0. En consecuencia,

$$0 \le I(x(\epsilon,\delta)) - I(x_0) = F(\epsilon,\delta) + G(\epsilon,\delta) \tag{16}$$

donde

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

$$G(\epsilon,\delta)=\int_{\bar{t}+\epsilon\delta}^{\bar{t}+\delta}\{f(t,x(t,\epsilon,\delta),\dot{x}_0(t)-\eta(u-\bar{u}))-f(t,x_0(t),\dot{x}_0(t))\}dt$$

Aquí, $\eta=\epsilon/(1-\epsilon)$. Tenemos que

$$\begin{aligned}\lim_{\delta\to 0}\frac{F(\epsilon,\delta)}{\epsilon\delta} &= f(\bar{t},\bar{x},u)-f(\bar{t},\bar{x},\bar{u})\\ \lim_{\delta\to 0}\frac{G(\epsilon,\delta)}{\epsilon\delta} &= \frac{f(\bar{t},\bar{x},\bar{u}-\eta(u-\bar{u}))-f(\bar{t},\bar{x},\bar{u})}{\eta}.\end{aligned}$$

En vista de (16) se tiene la desigualdad

$$f(\bar{t},\bar{x},u)-f(\bar{t},\bar{x},\bar{u})\ +\frac{f(\bar{t},\bar{x},\bar{u}-\eta(u-\bar{u}))-f(\bar{t},\bar{x},\bar{u})}{\eta}\geq 0\ (0<\eta\leq\eta_0)\quad(17)$$

donde $\eta_0=\epsilon_0/(1-\epsilon_0)$. Tomando límite cuando $\eta\to 0$, el último miembro en (17) es la derivada parcial de f en la dirección del vector $(\bar{u}-u)=v$, entonces se tiene que

$$\frac{\partial f(\bar{t},\bar{x},\bar{u})}{\partial v}=\langle\bar{u}-u,f_{\dot{x}}(\bar{t},\bar{x},\bar{u})\rangle$$

Por lo tanto, (17) es igual a

$$f(\bar{t},\bar{x},u)-f(\bar{t},\bar{x},\bar{u})\ -\langle u-\bar{u},f_{\dot{x}}(\bar{t},\bar{x},\bar{u})\rangle\ =E(\bar{t},\bar{x},\bar{u},u)\geq 0$$

La condición de Weierstrass por lo tanto se cumple a lo largo de x_0, excepto posiblemente en los puntos finales y en los puntos esquina de x_0. Por la continuidad de f y $f_{\dot{x}}$, se cumple también en estos puntos.

5 La primera y la segunda variaciones de I.

Sea x un arco admisible y sea y un arco en $\mathcal{E}$. Escogemos $\delta>0$ tal que el arco $x+\epsilon y$ sea admisible si ϵ está en el rango $-\delta<\epsilon<\delta$.

La función

$$F(\epsilon)=I(x+\epsilon y)=\int_a^b f(t,x(t)+\epsilon y(t),\dot{x}(t)+\epsilon\dot{y}(t))dt\qquad(18)$$

es de clase C^2 en ϵ sobre $-\delta<\epsilon<\delta$. La derivada $F'(0)$ de F en $\epsilon=0$ se llama la primera variación de I a lo largo de x y es denotada por $I'(x,y)$. Derivando (18) con respecto a ϵ y evaluando en $\epsilon=0$, se obtiene

$$I'(x,y)=\int_a^b\{\langle f_x,y(t)\rangle+\langle f_{\dot{x}},\dot{y}(t)\rangle\}dt\qquad(19)$$

donde los argumentos en las derivadas de f son los elementos $(t,x(t),\dot{x}(t))$ pertenecientes a x. La segunda derivada de $F(\epsilon)$ en $\epsilon=0$, se llama la segunda variación de I a lo largo de x y se denota por el símbolo $I''(x,y)$. Puesto que $F'(\epsilon) = I'(x+\epsilon y, y)$, derivando con respecto a ϵ y evaluando en $\epsilon = 0$ se obtiene

$$\begin{aligned} I''(x,y) &= \int_a^b \{\langle f_{xx}y(t), y(t)\rangle + 2\langle f_{x\dot{x}}\dot{y}(t), y(t)\rangle + \langle f_{\dot{x}\dot{x}}\dot{y}(t), \dot{y}(t)\rangle\} dt \\ &= \int_a^b 2\omega(t, y(t), \dot{y}(t)) dt \end{aligned}$$

donde $2\omega(t, y(t), \dot{y}(t)) = \langle f_{xx}y(t), y(t)\rangle + 2\langle f_{x\dot{x}}\dot{y}(t), y(t)\rangle + \langle f_{\dot{x}\dot{x}}\dot{y}(t), \dot{y}(t)\rangle$, los argumentos en las derivadas de f son los elementos $(t, x(t), \dot{x}(t))$ que definen a x. Notamos que por ser f de clase C^2 las matrices f_{xx}, $f_{x\dot{x}}$, y $f_{\dot{x}\dot{x}}$ son simétricas, y en consecuencia la segunda variación $I''(x,y)$ es una forma cuadrática en y sobre $\mathcal{E}$. Su forma bilineal asociada $I''(y,z)$ está dada por la fórmula

$$I''(x,y,z) = \int_a^b \{\langle \omega_y, z\rangle + \langle \omega_{\dot{y}}, \dot{z}\rangle\} dt.$$

En la sección 4 se mostrará que la primera y la segunda variación de I coinciden con el primer y segundo diferenciales de Fréchet de I con respecto a una cierta norma débil.

Teorema 2.14 *Si un arco admisible x_0 minimiza a I en la clase de arcos admisibles que unen sus puntos finales, entonces*

$$I'(x_0, y) = 0 \qquad y \qquad I''(x_0, y) \geq 0 \tag{20}$$

para todo arco y en la clase $\mathcal{E}_0$ de arcos en $\mathcal{E}$ que cumplen con $y(a) = y(b) = 0$.

La relación

$$I'(x_0, y) = 0 \tag{21}$$

se cumple sobre $\mathcal{E}_0$ si y sólo si la ecuación

$$f_{\dot{x}} = \int_a^t f_x ds + c$$

es satisfecha por x_0 con vector constante c, esto es, si y sólo si x_0 es un extremal de la integral I.

Prueba. Si y está en $\mathcal{E}_0$, entonces $y(a) = y(b) = 0$ y $x_0 + \epsilon y$ une los puntos finales de x_0. La función (18) con $x = x_0$ por lo tanto tiene un mínimo local en $\epsilon = 0$. De esto se sigue que

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

La prueba de la última afirmación se puede hacer con la ayuda del siguiente

Lema 2.1 *Dado un arco admisible x, hay un único arco z en $\mathcal{E}_0$ tal que*

$$I'(x,y) = (y,z) \tag{22}$$

para todo arco y en $\mathcal{E}_0$, donde

$$(y,z) = \int_a^b \langle \dot{y}(t), \dot{z}(t)\rangle dt \tag{23}$$

El arco z está definido por las relaciones

$$z(a) = 0, \qquad \dot{z}(t) = f_{\dot{x}} - \int_a^t f_x ds - c, \tag{24}$$

donde las derivadas de f son evaluadas a lo largo del arco x y el vector constante c se escoge de manera que $z(b) = 0$.

Prueba. Vemos que

$$z(b) = \int_a^b \dot{z}(t)dt = \int_a^b \left\{ f_{\dot{x}} - \int_a^t f_x ds \right\} dt - c(b-a)$$

Por lo tanto, la fórmula para c es

$$c = \frac{1}{b-a}\int_a^b \left\{ f_{\dot{x}} - \int_a^t f_x ds \right\} dt$$

Para establecer (22), sea y un arco en $\mathcal{E}_0$. Entonces $y(a) = y(b) = 0$, y

$$\begin{aligned}
0 &= \int_a^b \frac{d}{dt}\left\langle y, \int_a^t f_x ds + c \right\rangle dt \\
&= \int_a^b \left[\langle y, f_x\rangle + \left\langle \dot{y}, \int_a^t f_x ds + c \right\rangle \right] dt \\
&= \int_a^b [\langle y, f_x\rangle + \langle \dot{y}, f_{\dot{x}} - \dot{z}\rangle] dt \\
&= \int_a^b [\langle y, f_x\rangle + \langle \dot{y}, f_{\dot{x}}\rangle] dt \ - \int_a^b \langle \dot{y}, \dot{z}\rangle dt \\
&= I'(x,y) - (y,z).
\end{aligned}$$

Si hubieran dos funciones z_1 y z_2 en $\mathcal{E}_0$ tal que (22) se cumple con $z = z_1$ y $z = z_2$, entonces

$$0 = (y, z_1) - (y, z_2) = (y, z_1 - z_2)$$

para todo y en $\mathcal{E}_0$. Tomando $y = z_1 - z_2$, se ve que

$$\int_a^b |\dot{y}(t)|^2 dt = 0$$

y en consecuencia que $\dot{y}(t) = 0$ sobre $a \leq t \leq b$. Puesto que $y(a) = 0$, se sigue que $y = 0$ y que $z_1 = z_2$.

Para probar la última afirmación del Teorema 2.14, escogemos z en $\mathcal{E}_0$ relacionado a x_0 como se describió en el Lema 2.1. Por lo tanto,

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

Entonces, la relación

$$I'(x_0, y) = (y, z) = 0$$

se cumple en $\mathcal{E}_0$ si y sólo si $\dot{z}(t) = 0$ sobre $a \leq t \leq b$, esto es, si y sólo si la relación

$$f_{\dot{x}} = \int_a^t f_x ds + c$$

se cumple a lo largo de x_0.

6 La condición de Jacobi.

Recordemos que un extremo fue definido como un arco admisible sin esquinas que satisface la ecuación de Euler

$$\frac{d}{dt} f_{\dot{x}} = f_x$$

En virtud del Teorema de diferenciabilidad de Hilbert un extremo no singular es de clase C^2 si f es de clase C^2 y la regla de la cadena para derivar se puede llevar al cabo en la ecuación de Euler. Para estudiar las propiedades elementales de extremos no singulares es conveniente reemplazar las variables $(t, x, \dot{x})$ por las variables canónicas (t, x, p) obtenidas definiendo

$$p = f_{\dot{x}}(t, x, \dot{x}) \tag{25}$$

Supongamos ahora que x_0 es un extremo no singular. Entonces el determinante

$$|f_{\dot{x}\dot{x}}| \neq 0$$

a lo largo de x_0, y por el Teorema de la función implícita (25) tiene la solución

$$\dot{x} = P(t, x, p) \tag{26}$$

sobre una vecindad $\mathcal{P}$ de los elementos (t, x, p) sobre x_0. La función P es de clase $C^{(m-1)}$ sobre $\mathcal{P}$ si f es de clase C^m sobre R. La función

$$H(t, x, p) = \langle p, P \rangle - f(t, x, P)$$

es llamada el Hamiltoniano correspondiente a f o a la integral I. Derivando H tenemos

$$H_t = -f_t, \qquad H_x = -f_x, \qquad H_p = P$$

En consecuencia, H es de clase C^m sobre $\mathcal{P}$ siempre y cuando f sea de clase C^m sobre R. En vista de (25) y (26), la ecuación de Euler es equivalente a las ecuaciones

$$\dot{x} = H_p, \qquad \dot{p} = -H_x \tag{27}$$

Estas serán llamadas la forma canónica o Hamiltoniana de la ecuación de Euler.

Ahora consideramos un extremo no singular

$$x_0 : x_0(t) \qquad (a \leq t \leq b)$$

y definimos

$$p_0(t) = f_{\dot{x}}(t, x_0(t), \dot{x}_0(t))$$

Las funciones $x_0(t)$, $p_0(t)$ son soluciones de (27). Por el Teorema de inclusión 2.11 existe a través de cada punto $(\bar{t}, \bar{x}, \bar{p})$ en una vecindad $\mathcal{P}_0$ de aquellos puntos sobre x_0 una única solución

$$x(t, \bar{t}, \bar{x}, \bar{p}),\ p(t, \bar{t}, \bar{x}, \bar{p}) \qquad (a - \delta \leq t \leq b + \delta)$$

que cae en una vecindad preescrita $\mathcal{P}_1$ de x_0 en el *espacio*$-$*txp*. Estas funciones y sus derivadas con respecto a t son de clase $C^{(m-1)}$ si f y en consecuencia H son de clase C^m $(m \geq 2)$. Definiendo $\alpha = \bar{x}$, $\beta = \bar{p}$ y manteniendo $\bar{t}$ fija, se obtiene el siguiente

Teorema 2.15 *Un extremo no singular x_0 es miembro de una familia 2n-paramétria*

$$x(t, \alpha, \beta)$$

de extremos para valores $\alpha = \alpha_0$, $\beta = \beta_0$, $a \leq t \leq b$. Las funciones $x(t, \alpha, \beta)$ y

$$p(t, \alpha, \beta) = f_{\dot{x}}(t, x(t, \alpha, \beta), \dot{x}(t, \alpha, \beta))$$

y sus derivadas $\dot{x}(t, \alpha, \beta)$, $\dot{p}(t, \alpha, \beta)$ con respecto a t son de clase $C^{(m-1)}$ si f es de clase C^m $(m \geq 2)$ sobre R. Además el determinante

$$d^*(t, \alpha, \beta) = \begin{vmatrix} x_\alpha(t, \alpha, \beta) & x_\beta(t, \alpha, \beta) \\ p_\alpha(t, \alpha, \beta) & p_\beta(t, \alpha, \beta) \end{vmatrix}$$

es diferente de cero a lo largo de x_0.

De hecho se puede probar que,

$$d^*(t, \alpha, \beta) = |f_{\dot{x}\dot{x}}| d(t, \alpha, \beta)$$

donde

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

Se sigue del Teorema 2.15 que existe una constante $\delta > 0$ tal que

$$x(t, \alpha_0, \beta_0) \qquad (a - \delta \leq t \leq b + \delta)$$

es también un extremo. Este extremo tiene a x_0 como un subarco y será llamado una extensión de x_0.

Si en la familia del Teorema 2.15 los parámetros (α, β) se escogen tal que las relaciones

$$\alpha = x(t_0, \alpha, \beta), \qquad \beta = \dot{x}(t_0, \alpha, \beta)$$

se cumplan, entonces definiendo $\alpha = \alpha_0 = x_0(t_0)$ uno obtiene el siguiente

Teorema 2.16 *Un extremo no singular x_0 es un miembro de una familia n-paramétrica*

$$x(t, \beta_1, \ldots, \beta_n)$$

de extremos para el valor $\beta = \beta_0$, $a \leq t \leq b$, que pasan a través de un punto dado sobre x_0 (o sobre una extensión de x_0). Esta familia tiene propiedades de continuidad y diferenciabilidad descritas en el teorema anterior.

Además, la matriz

$$\left\| \begin{array}{c} x_\beta(t, \beta) \\ \dot{x}_\beta(t, \beta) \end{array} \right\|$$

tiene rango n en cada punto de x_0.

Teorema 2.17 *Sea un extremo x_0 un miembro para $\epsilon = 0$, $a \leq t \leq b$ de una familia uniparamétrica de extremos*

$$x(\epsilon) : \; x(t, \epsilon)$$

tal que las funciones $x(t, \epsilon)$, $\dot{x}(t, \epsilon)$ son de clase C^1. La variación

$$y : \; y(t) = x_\epsilon(t, 0) \qquad (a \leq t \leq b)$$

de esta familia a lo largo de x_0 satisface la ecuación de variación

$$\frac{d}{dt}[\langle f_{\dot{x}x}, y\rangle + \langle f_{\dot{x}\dot{x}}, \dot{y}\rangle] \; = \langle f_{xx}, y\rangle + \langle f_{x\dot{x}}, \dot{y}\rangle \tag{28}$$

a lo largo de x_0 de la ecuación de Euler.

Prueba. Este resultado se obtiene de la relación

$$\frac{d}{dt} f_{\dot{x}}(t, x(t, \epsilon), \dot{x}(t, \epsilon)) \; = f_x(t, x(t, \epsilon), \dot{x}(t, \epsilon))$$

derivando con respecto a ϵ y evaluando en $\epsilon = 0$.

La ecuación (28) es la ecuación de Euler de la segunda variación de I a lo largo de x_0.

Se ha visto que la segunda variación $I''(x_0, y)$ de I a lo largo de un arco admisible x_0 toma la forma

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

donde

$$2\omega = \langle f_{xx}y, y\rangle + 2\langle f_{x\dot{x}}\dot{y}, y\rangle \ + \langle f_{\dot{x}\dot{x}}\dot{y}, \dot{y}\rangle,$$

y los argumentos en las derivadas de f son $(t, x_0(t), \dot{x}_0(t))$. Además, fue probado que si x_0 minimiza I sobre la clase de arcos admisibles que unen sus puntos finales entonces,

$$I''(x_0, y) \geq 0$$

en la subclase $\mathcal{E}_0$ de arcos y en $\mathcal{E}$ que cumplen con $y(a) = y(b) = 0$.

Esta condición es una condición de segundo orden, y será llamada la *segunda condición necesaria para un mínimo.* La desigualdad $I''(x_0, y) \geq 0$ sobre $\mathcal{E}_0$ es equivalente a la afirmación de que $y = 0$ minimiza $I''(x_0, y)$ sobre $\mathcal{E}_0$. El problema de minimizar $I''(x_0, y)$ sobre $\mathcal{E}_0$ será llamado el problema del mínimo accesorio.

Se supondrá que el arco x_0 que minimiza a I es no singular y que no tiene esquinas. Para un arco con esquinas la segunda condición necesaria se debe cumplir también para arcos y con ciertas discontinuidades en los puntos esquina de x_0.

Será conveniente designar la segunda variación $I''(x_0, y)$ por $J(y)$. La integral

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

es una forma cuadrática en y. La correspondiente forma bilineal es

$$J(y, z) = \int_a^b \{\langle\omega_y, z\rangle + \langle\omega_{\dot{y}}, \dot{z}\rangle\}dt \tag{29}$$

Es claro que

$$J(y, z) = J(z, y), \qquad J(y, y) = J(y)$$

Puesto que x_0 es no singular y satisface la condición de Legendre, se tiene que $f_{\dot{x}\dot{x}} > 0$ en tanto que $2\omega_{\dot{y}\dot{y}} = f_{\dot{x}\dot{x}}$, entonces $2\omega_{\dot{y}\dot{y}} > 0$. Por lo tanto se sigue que $J(y)$ es positivamente regular y que la solución de la ecuación de Euler,

$$\omega_{\dot{y}} = \int_a^t \omega_y ds + c$$

no tiene esquinas. De acuerdo a esto, esta ecuación puede ser escrita en la forma

$$\frac{d}{dt}\omega_{\dot{y}} = \omega_y \tag{30}$$

Esta ecuación es llamada la ecuación diferencial accesorio o de Jacobi y sus soluciones serán llamadas extremos accesorio.

Si y es un extremo accesorio, entonces por (30) y (29),

$$J(y,z) = \int_a^b \frac{d}{dt}\langle z, \omega_{\dot{y}}\rangle dt = \langle z, \omega_{\dot{y}} \rangle \Big|_a^b$$

Si z es también un extremo accesorio, el orden de y y z puede ser intercambiado.

Esto da la relación

$$\langle z, \omega_{\dot{y}}\rangle - \langle y, \omega_{\dot{z}}\rangle \Big|_a^b = 0$$

Puesto que b puede ser arbitrario, se obtiene el Lema 2.2. Este Lema es independiente de la suposición de no singularidad de x_0.

Lema 2.2 *Si y y z son extremos accesorio, la expresión*

$$\langle z, \omega_{\dot{y}}\rangle - \langle y, \omega_{\dot{z}}\rangle$$

es constante sobre $a \leq t \leq b$.

En esta expresión los argumentos en las derivadas de ω son $(t, y, \dot{y})$ si y es usada como subíndice y son $(t, z, \dot{z})$ si z es usado.

Puesto que el determinante $|f_{\dot{x}\dot{x}}| \neq 0$ a lo largo de x_0, se sigue que la ecuación

$$q = \omega_{\dot{y}} = f_{\dot{x}x} y + f_{\dot{x}\dot{x}} \dot{y}$$

puede ser resuelta para $\dot{y}$ en términos de y y q. Al hacer esto se ve que la ecuación de Euler (30) es equivalente a un sistema de la forma

$$\dot{y} = L(t, y, q), \qquad \dot{q} = M(t, y, q), \tag{31}$$

donde L y M son lineales en y y en q. Las variables (t, y, q) son variables canónicas. El sistema (31) es lineal. De la teoría de ecuaciones diferenciales lineales aplicada a (31) se obtiene el siguiente

Lema 2.3 *Un extremo accesorio y está determinado de manera única por los valores de $y(t_0)$, $\dot{y}(t_0)$ o equivalentemente por los valores $y(t_0)$, $q(t_0)$ en un punto $t = t_0$. En particular, si $y(t_0) = \dot{y}(t_0) = 0$, entonces $y(t) \equiv 0$ $a \leq t \leq b$.*

Definición 2.3 *Un punto $t = c$ sobre $a < t \leq b$ se dice que define un punto conjugado a $t = a$ sobre x_0 si hay un extremo accesorio y, tal que $y(a) = y(c) = 0$, y $y(t) \not\equiv 0$ sobre $a < t < c$.*

Teorema 2.18 (La Condición de Jacobi.) *Si un arco admisible no singular sin esquinas minimiza a I en la clase de arcos admisibles que unen sus puntos finales, entonces no hay un punto c sobre x_0 conjugado a $t = a$ entre a y b.*

Prueba. Supongamos que hay un punto c conjugado a a sobre $a < t < b$. Sea y un extremo accesorio no nulo con $y(a) = y(c) = 0$. Sea z el arco que cumple con

$$z(t) = y(t) \quad a \leq t \leq c, \qquad z(t) \equiv 0 \quad c \leq t \leq b.$$

El arco z está en $\mathcal{E}_0$ y

$$J(z) = \int_a^c 2\omega(t, z, \dot{z})dt = \langle z, \omega_{\dot{z}} \rangle \Big|_a^c = 0$$

En consecuencia z minimiza $J(z)$ sobre $\mathcal{E}_0$ y es entonces un extremo accesorio. Puesto que $z(t) \equiv 0$ $c \leq t \leq b$, se sigue del Lema 2.3 que $z(t) \equiv 0$ sobre $a \leq t \leq b$, lo cual no es el caso. Esta contradicción prueba el Teorema.

Como un ejemplo de puntos conjugados observamos que si

$$J(y) = \int_a^b (\dot{y}^2 - y^2)dt$$

$2\omega(t, y, \dot{y}) = \dot{y}^2 - y^2$, entonces $\omega_{\dot{y}} = \dot{y}$ y $\omega_y = -y$, por lo tanto la ecuación diferencial accesorio es

$$\ddot{y} + y = 0$$

Los extremos accesorio que se anulan en $t = 0$ son de la forma

$$y = Asen(t)$$

donde A es una constante. Los puntos conjugados de $t = 0$ son los valores $c = n\pi$. En consecuencia si $b > \pi$, el punto $c = \pi$ es conjugado a $t = 0$ sobre $0 < t < b$, y $J(y)$ no puede ser no negativa sobre la clase de arcos que se anulan en $t = 0$ y en $t = b$.

7 Determinación de puntos conjugados.

Lema 2.4 *Un conjunto de r extremos accesorio $y_1, \ldots, y_r$ es linealmente independiente si y sólo si la matriz*

$$\left\| \begin{matrix} y_j(t) \\ \dot{y}_j(t) \end{matrix} \right\|$$

tiene rango r en algún punto $t = t_0$. Si tiene rango r en $t = t_0$, ésta tiene rango r para todos los valores de t sobre $a \leq t \leq b$. Existen $2n$ extremos accesorio linealmente independientes. Si $y_1, \ldots, y_{2n}$ son $2n$ extremos accesorio linealmente independientes, entonces cada extremo accesorio y es expresable en la forma

$$y = \sum_1^{2n} y_j b_j$$

donde $b_1, \ldots, b_{2n}$ son constantes.

Prueba. Supongamos que $y_1, \ldots, y_r$ son linealmente independientes, y que existe un punto t_1 en $a \leq t \leq b$ tal que

$$\sum_1^r \xi_j (y_j(t_1), \dot{y}_j(t_1)) = 0$$

donde las $\xi_j' s$ son constantes no todas cero. El arco $w(t) = \sum_1^r \xi_j y_j(t)$ es un extremo accesorio, con $w(t_1) = \dot{w}(t_1) = 0$. Por el Lema 2.3 $w(t) \equiv 0$, implica que $\xi_j = 0$ $(j = 1, \ldots, r)$ lo cual es una contradicción.

Supongamos que $y_1, \ldots, y_r$ son linealmente dependientes, entonces existe $\alpha = (\alpha_1, \ldots, \alpha_r) \neq 0$ tal que $\sum_1^r \alpha_j y_j(t) = 0$ sobre $a \leq t \leq b$. Entonces $\sum_1^r \alpha_j \dot{y}_j(t) = 0$ sobre $a \leq t \leq b$. Esto implica que $\sum_1^r \alpha_j (y_j(t), \dot{y}_j(t)) = 0$ sobre $a \leq t \leq b$ con $\alpha \neq 0$, entonces la matriz del Lema no tiene rango r en ningún punto de $a \leq t \leq b$.

Si la matriz del Lema tiene rango r en algún punto $t = t_0$, entonces para todo $\beta = (\beta_1, \ldots, \beta_r) \neq 0$, tenemos que

$$\sum_1^r \beta_j (y_j(t_0), \dot{y}_j(t_0)) = (\sum_1^r \beta_j y_j(t_0), \ \sum_1^r \beta_j \dot{y}_j(t_0)) = (z(t_0; \beta), \dot{z}(t_0; \beta)) \neq 0$$

donde $z(t; \beta) = \sum_1^r \beta_j y_j(t)$ es un extremo accesorio.

Por el Lema 2.3 $(z(t; \beta), \dot{z}(t; \beta)) \neq 0$ para todo t sobre $a \leq t \leq b$. Entonces la matriz tiene rango r para todos los valores de t sobre $a \leq t \leq b$.

Por el teorema de inclusión de ecuaciones diferenciales escogiendo $\bar{t}$ fija sobre $a \leq t \leq b$, y por el sistema lineal (31), sabemos que un extremo accesorio es miembro de una familia 2n-paramétrica de extremos accesorio $y(t, \alpha, \beta) = y(t, \alpha_1, \ldots, \alpha_n, \beta_1 \ldots, \beta_n)$. Entonces podemos tomar 2n vectores $(\alpha_i, \beta_i) = (\alpha_i^1, \ldots, \alpha_i^n, \beta_i^1, \ldots, \beta_i^n)$ $(i = 1, \ldots, 2n)$ linealmente independientes tal que a través de cada uno pasa una solución (y_i, q_i) de (31), i.e., existen 2n soluciones del sistema lineal (31) tal que $(y_i(\bar{t}), q_i(\bar{t})) = (\alpha_i, \beta_i)$ y en consecuencia la matriz

$$\left\| \begin{array}{c} y_i(t) \\ q_i(t) \end{array} \right\|$$

tiene rango 2n en $t = \bar{t}$. Además, tenemos que

$$\left\| \begin{array}{c} y_j(t) \\ \dot{y}_j(t) \end{array} \right\| = \left\| \begin{array}{c} y_j(t) \\ L(t, y_j(t), q_j(t)) \end{array} \right\|$$

donde L es lineal en y y en q. En consecuencia, las dos matrices tienen rango $2n$. Por lo tanto, $y_1, \ldots, y_{2n}$ son 2n extremos accesorio linealmente independientes.

Sean $y_1, \ldots, y_{2n}$ 2n extremos accesorio linealmente independientes, y sea y un extremo accesorio. Entonces, para cualquier c en $a \leq t \leq b$ existen constantes $b_i' s$ $(i = 1, \ldots, 2n)$ tal que $(y(c), \dot{y}(c)) = \sum_{i=1}^{2n} b_i (y_i(c), \dot{y}_i(c))$. El arco $z(t) = \sum_{i=1}^{2n} b_i y_i(t)$ es un extremo accesorio que cumple con $z(c) = y(c)$ y $\dot{z}(c) = \dot{y}(c)$. Por el Lema 2.3 $y(t) = \sum_{i=1}^{2n} b_i y_i(t)$ $a \leq t \leq b$.

Como un primer resultado sobre la determinación de puntos conjugados tenemos el siguiente

Teorema 2.19 *Sean $y_1, \ldots, y_n, z_1, \ldots, z_n$ 2n extremos accesorio linealmente independientes y definimos*

$$D(t, t_0) = \begin{vmatrix} y_j(t) & z_j(t) \\ y_j(t_0) & z_j(t_0) \end{vmatrix}$$

Los puntos $t = c$ conjugados a $t = a$ son los ceros $c \neq a$ de $D(t, a)$.

Prueba. Supongamos que c es conjugado a a. Sea y un extremo accesorio no nulo tal que $y(a) = y(c) = 0$. Seleccionamos constantes α_j, β_j tal que

$$y = \sum y_j \alpha_j + z_j \beta_j \tag{32}$$

Puesto que $y \neq 0$, estas constantes no son todas cero. La anulación de $y(t)$ en $t = c$ y $t = a$ produce las ecuaciones

$$\begin{aligned} 0 &= \sum \{y_j(c)\alpha_j + z_j(c)\beta_j\} \\ 0 &= \sum \{y_j(a)\alpha_j + z_j(a)\beta_j\} \end{aligned} \tag{33}$$

En consecuencia, $D(c, a) = 0$. Recíprocamente si $D(c, a) = 0$ $(c \neq a)$, existen constantes α_j, β_j, no todas cero, tal que (33) se cumple. El arco y definido por (32) es un extremo accesorio no nulo que tiene $y(c) = y(a) = 0$. En consecuencia, c es conjugado a $t = a$.

Corolario 2.7 *Sea $x(t, \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n)$ una familia 2n paramétrica que contiene a x_0 para valores $\alpha = \alpha_0$, $\beta = \beta_0$, $a \leq t \leq b$, y que tiene las propiedades descritas en el Teorema 2.15. Un punto c es conjugado a $t = a$ sobre x_0 si y sólo si $c \neq a$ y $D(c, a, \alpha_0, \beta_0) = 0$, donde*

$$D(t, t_0, \alpha, \beta) = \begin{vmatrix} x_\alpha(t, \alpha, \beta) & x_\beta(t, \alpha, \beta) \\ x_\alpha(t_0, \alpha, \beta) & x_\beta(t_0, \alpha, \beta) \end{vmatrix}$$

Prueba. Por el Teorema 2.17 $x_{\alpha_k}(t, \alpha_0, \beta_0)$, x_{β_k} (t, α_0, β_0) $(k = 1, \ldots, n)$ son extremos accesorio y por el Teorema 2.15 la matriz

$$\left\| \begin{matrix} x_\alpha(t, \alpha, \beta) & x_\beta(t, \alpha, \beta) \\ \dot{x}_\alpha(t, \alpha, \beta) & \dot{x}_\beta(t, \alpha, \beta) \end{matrix} \right\|$$

tiene rango 2n a lo largo de x_0. Por lo tanto por el Lema 2.4 $x_{\alpha_1}, \ldots, x_{\alpha_n}$, $x_{\beta_1}, \ldots, x_{\beta_n}$ son 2n extremos accesorio linealmente independientes, donde los $x_{\alpha_k}, x_{\beta_k}$ están evaluados en (t, α_0, β_0). Entonces el resultado se sigue del Teorema 2.19.

Teorema 2.20 *Sean $z_1, \ldots, z_n$ n extremos accesorio linealmente independientes que se anulan en $t = a$, y definimos*

$$\Delta(t) = |z_j(t)|$$

Un punto c es conjugado a $t = a$ si y sólo si $c \neq a$ y $\Delta(c) = 0$.

Prueba. Este resultado se sigue del Teorema 2.19 seleccionando n extremos accesorio $y_1, \ldots, y_n$ tal que $|y_j(a)| = 1$. Ya que entonces tenemos

$$D(t,a) = \begin{vmatrix} y_j(t) & z_j(t) \\ y_j(a) & 0 \end{vmatrix} = (-1)^n \Delta(t)$$

Corolario 2.8 *Sea*

$$x(t, \beta_1, \ldots, \beta_n)$$

una familia n-paramétrica de extremos que pasan a través del punto inicial de x_0, y que contienen a x_0 para el valor $\beta = \beta_0$ $a \leq t \leq b$, y que tienen las propiedades descritas en el Teorema 2.16. Un punto c es conjugado a $t = a$ sobre x_0 si y sólo si $c \neq a$ y $\Delta(c, \beta_0) = 0$, donde

$$\Delta(t, \beta) = |x_\beta(t, \beta)|$$

Prueba. De las relaciones

$$x_0(a) = x(a, \beta)$$

Derivando se sigue que $x_{\beta_j}(a, \beta_0) = 0$. Las funciones

$$z_j(t) = x_{\beta_j}(t, \beta_0)$$

son extremos accesorio que se anulan en $t = a$. El Corolario es entonces una consecuencia del Teorema 2.20.

Teorema 2.21 *Sean $y_1, \ldots, y_{2n}$ extremos accesorio linealmente independientes y sea*

$$D(t, t_0) = \begin{vmatrix} y_j(t) \\ y_j(t_0) \end{vmatrix}$$

Existe una constante $\delta > 0$ tal que $D(t, t_0) \neq 0$, para cada par de puntos distintos t, t_0 sobre $a \leq t \leq b$ que cumplen con $|t - t_0| \leq \delta$.

Prueba. Observamos que por la fórmula de Taylor,

$$y_j^i(t) - y_j^i(t_0) = (t - t_0) A_j^i(t, t_0) \qquad (j = 1, \ldots, 2n)$$

donde

$$A_j^i(t, t_0) = \int_0^1 \dot{y}_j^i(t_0 + \theta(t - t_0)) d\theta$$

tenemos por lo tanto,

$$D(t,t_0) = \left| \begin{array}{c} y_j(t) \\ y_j(t_0) \end{array} \right| = \left| \begin{array}{c} y_j(t) - y_j(t_0) \\ y_j(t_0) \end{array} \right| = (t-t_0)^n \Delta(t,t_0),$$

donde

$$\Delta(t,t_0) = \left| \begin{array}{c} A_j(t,t_0) \\ y_j(t_0) \end{array} \right|$$

Puesto que los arcos $y_1, \ldots, y_{2n}$ son linealmente independientes se sigue que

$$\Delta(t,t) = \left| \begin{array}{c} \dot{y}_j(t) \\ y_j(t) \end{array} \right| \neq 0$$

sobre $a \leq t \leq b$. Puesto que $\Delta(t,t_0)$ es continua en t y t_0, se sigue que hay una constante $\delta > 0$ tal que $\Delta(t,t_0) \neq 0$ para todo t, t_0 sobre $a \leq t \leq b$ que cumple con $|t-t_0| < \delta$. En consecuencia, $D(t,t_0) = (t-t_0)^n \Delta(t,t_0) \neq 0$ si $t \neq t_0$ y $|t-t_0| \leq \delta$.

8 La positividad de la segunda variación.

En esta sección continuamos con la suposición de que x_0 es un extremo no singular que satisface la condición de Legendre.

Puesto que x_0 es no singular, éste se puede extender a la izquierda de a y a la derecha de b para obtener un extremo más grande

$$\bar{x}_0 : \ x_0(t) \qquad (a - \epsilon \leq t \leq b + \epsilon, \quad \epsilon > 0)$$

Este arco será llamado una extensión de x_0.

Lema 2.5 *Si no existe un punto c conjugado a $t = a$ sobre x_0 entre a y b o en b, entonces hay un punto a_0 a la izquierda de a sobre una extensión de x_0 que no tiene un punto conjugado sobre x_0.*

Prueba. Sean $y_1, \ldots, y_{2n}$ 2n extremos accesorio linealmente independientes definidos sobre un intervalo $a - \epsilon \leq t \leq b + \epsilon$ correspondientes a una extensión $\bar{x}_0$ de x_0. Por el Teorema 2.21 existe una constante $\delta > 0$ tal que

$$D(t,t_0) = \left| \begin{array}{c} y_j(t) \\ y_j(t_0) \end{array} \right|$$

es diferente de cero para puntos distintos t, t_0 sobre $a - \epsilon \leq t \leq b + \epsilon$ que satisfacen la relación $|t - t_0| \leq \delta$. Podemos seleccionar $\delta \leq 2\epsilon$. Como no hay un punto conjugado a $t = a$ sobre $a < t \leq b$, tenemos por el Teorema 2.19,

$$D(t,a) \neq 0 \qquad a < t \leq b,$$

y en consecuencia,

$$D(t,a) \neq 0 \qquad \textit{sobre} \qquad (a + \frac{\delta}{2} \leq t \leq b)$$

Por continuidad existe un punto a_0 sobre $a - \delta/2 \leq t < a$ tal que $D(t, a_0) \neq 0$ sobre $a+\delta/2 \leq t \leq b$. Puesto que un punto t sobre $a_0 < t < a+\delta/2$ difiere de a_0 a lo más por δ, tenemos $D(t, a_0) \neq 0$ sobre $a_0 < t \leq a+\delta/2$. Consecuentemente, $D(t, a_0) \neq 0$ sobre $a_0 < t \leq b$. De acuerdo a esto no hay un punto c conjugado a a_0 sobre $a_0 < t \leq b$ y en consecuencia ninguno sobre x_0.

Sea y_j $(j = 1, \ldots, n)$ un conjunto de extremos accesorio linealmente independientes, y sea

$$q_j(t) = \omega_{\dot{y}_j}(t, y_j(t), \dot{y}_j(t))$$

En virtud del Lema 2.2 las expresiones

$$\langle y_j(t), q_k(t)\rangle - \langle q_j(t), y_k(t)\rangle \tag{34}$$

tienen los mismos valores para toda t. Los extremos accesorio $y_1, \ldots, y_n$ se dirá que forman un sistema conjugado si ellos son linealmente independientes y las expresiones (34) son todas cero.

Teorema 2.22 *Si no existe un punto c sobre x_0 conjugado a $t = a$ entre a y b o en b, entonces existe un sistema conjugado $y_1, \ldots, y_n$ de extremos accesorio cuyo determinante*

$$\Delta(t) = |y_j(t)| \qquad (j = 1, \ldots, n)$$

es diferente de cero sobre $a \leq t \leq b$.

Prueba. Sea a_0 relacionado a x_0 como se describió en el Lema 2.5 y sean $y_1, \ldots, y_n$ extremos accesorio linealmente independientes que se anulan en $t = a_0$. Puesto que las expresiones (34) se anulan en $t = a_0$, estos extremos forman un sistema conjugado. El determinante correspondiente $\Delta(t)$ se anula solamente en $t = a_0$ y en los puntos conjugados de a_0. Este determinante por lo tanto no se anula sobre $a \leq t \leq b$.

Teorema 2.23 *La segunda variación $J(y) = I''(x_0, y)$ de I a lo largo de x_0 es positiva para todo $y \neq 0$ en $\mathcal{E}_0$ si y sólo si hay un sistema conjugado $y_1, \ldots, y_n$ de extremos accesorio cuyo determinante $\Delta(t)$ es diferente de cero sobre $a \leq t \leq b$.*

Prueba. Si $J(y)$ es positiva sobre $\mathcal{E}_0$, no puede haber un punto conjugado a $t = a$ sobre $a < t < b$, por el Teorema 2.18. Si b fuera conjugado a $t = a$, existiría un extremo accesorio no nulo que cumple con $y(a) = y(b) = 0$. Para este arco tendríamos

$$J(y) = \int_a^b 2\omega(t, y, \dot{y})dt = \langle y, \omega_{\dot{y}}\rangle\Big|_a^b = 0$$

y $J(y)$ no podría ser positiva sobre $\mathcal{E}_0$. Se sigue que si $J(y)$ es positiva sobre $\mathcal{E}_0$ no hay un punto conjugado c a $t = a$ sobre $a < t \leq b$.

En este evento hay, por el Teorema 2.22, un sistema conjugado cuyo determinante es diferente de cero sobre $a \leq t \leq b$.

Supongamos ahora que hay un sistema conjugado $y_1, \ldots, y_n$ con un determinante diferente de cero $\Delta(t)$ sobre $a \leq t \leq b$. Dado un arco y en $\mathcal{E}_0$, la ecuación

$$y(t) = \sum_j y_j(t) w^j \qquad (a \leq t \leq b) \tag{35}$$

tienen soluciones continuas

$$w: \qquad w^j(t) \qquad (a \leq t \leq b)$$

que se anulan en $t = a$ y en $t = b$ y que tienen derivadas continuas por fragmentos sobre $a \leq t \leq b$. Tenemos que

$$\dot{y} = \sum_j \dot{y}_j w^j + \pi, \qquad (\pi = \sum_j y_j \dot{w}^j) \tag{36}$$

Mostraremos que

$$J(y) = \int_a^b \langle f_{\dot{x}\dot{x}} \pi, \pi \rangle dt. \tag{37}$$

Como primer paso observamos que, como y_j es un extremo accesorio el vector canónico correspondiente $q_j = \omega_{\dot{y}_j}(t, y_j, \dot{y}_j)$ tiene

$$\dot{q}_j = \omega_{y_j}(t, y_j, \dot{y}_j)$$

como su derivada. Entonces para el arco y tenemos que

$$\begin{aligned}
\omega_y(t, y, \dot{y}) &= f_{xx} y + f_{x\dot{x}} \dot{y} \\
&= f_{xx} \sum_j y_j w^j + f_{x\dot{x}} \sum_j \dot{y}_j w^j + f_{x\dot{x}} \pi \\
&= \sum_j \{f_{xx} y_j + f_{x\dot{x}} \dot{y}_j\} w^j + f_{x\dot{x}} \pi \\
&= \sum_j \omega_{y_j} w^j + f_{x\dot{x}} \pi \\
&= \sum_j \dot{q}_j w^j + f_{x\dot{x}} \pi
\end{aligned}$$

Procediendo de manera similar para $\omega_{\dot{y}}(t, y, \dot{y})$ obtenemos las fórmulas

$$\omega_y = \sum_j \dot{q}_j w^j + f_{x\dot{x}} \pi,$$

$$\omega_{\dot{y}} = \sum_j q_j w^j + f_{\dot{x}\dot{x}} \pi.$$

En consecuencia, con la ayuda de (35) y (36), tenemos

$$\begin{aligned} 2\omega(t,y,\dot y) &= \langle y,\omega_y\rangle + \langle \dot y,\omega_{\dot y}\rangle \\ &= \sum_{j,k}(\langle \dot q_j, y_k\rangle \ + \langle q_j,\dot y_k\rangle)w^j w^k + 2\sum_{j,k}\langle q_j, y_k\rangle w^j \dot w^k \ + \langle f_{\dot x\dot x}\pi,\pi\rangle. \end{aligned}$$

Puesto que $y_1,\ldots,y_n$ es un sistema conjugado, tenemos

$$\langle q_j, y_k\rangle = \langle q_k, y_j\rangle.$$

Consecuentemente,

$$2\omega = \frac{d}{dt}\sum_{j,k}\langle q_j, y_k\rangle w^j w^k + \ \langle f_{\dot x\dot x}\pi,\pi\rangle.$$

Usando las relaciones $w(a) = w(b) = 0$ se obtiene la fórmula (37). En virtud de que x_0 es no singular y satisface la condición de Legendre, se sigue de (37) que $J(y) > 0$ al menos que

$$\pi(t) = \sum_j y_j(t)\dot w^j(t) = 0.$$

Puesto que el determinante $|y_j(t)| \neq 0$ $a \leq t \leq b$ y $w(a) = 0$, esto sólo puede ocurrir si $w(t) \equiv 0$, y en consecuencia sólo si $y(t) = 0$ sobre $a \leq t \leq b$. Esto completa la prueba del Teorema 2.23.

En el curso de las demostraciones de los Teoremas 2.22 y 2.23 hemos establecido el siguiente

Teorema 2.24 *La desigualdad $J(y) = I''(x_0, y) > 0$ se cumple para todo $y \neq 0$ en $\mathcal{E}_0$ si y sólo si no existe un punto c conjugado a $t = a$ sobre $a < t \leq b$.*

Más adelante se darán pruebas de suficiencia, que de hecho, serán una extensión del cálculo de variaciones a ciertos problemas de control óptimo.

3 Control óptimo.

1 Declaración del problema.

Consideremos un intervalo $T = [a, b]$ en $\mathbf{R}$, dos puntos ξ_0, ξ_1 en $\mathbf{R}^n$, un conjunto abierto $\mathcal{A}$ en $T \times \mathbf{R}^n \times \mathbf{R}^m$, una función L que mapea $T \times \mathbf{R}^n \times \mathbf{R}^m$ a $\mathbf{R}$, y matrices de funciones continuas $A(t)$ y $B(t)$ de dimensiones $n \times n$ y $n \times m$ respectivamente.

Denotemos por X el espacio vectorial de funciones C^1 por fragmentos que mapean T a $\mathbf{R}^n$, por U el espacio vectorial de funciones continuas por fragmentos que mapean T a $\mathbf{R}^m$, y sea $Z = X \times U$,

$$\begin{aligned} D &= \{(x,u) \in Z \mid \dot{x}(t) = A(t)x(t) + B(t)u(t) \ (t \in T)\}, \\ Z(\mathcal{A}) &= \{(x,u) \in Z \mid (t, x(t), u(t)) \in \mathcal{A} \ (t \in T)\}, \\ Z_e(\mathcal{A}) &= \{(x,u) \in Z(\mathcal{A}) \cap D \mid \ x(a) = \xi_0, \ x(b) = \xi_1\}. \end{aligned}$$

El problema que vamos a considerar, el cual llamamos $P(\mathcal{A})$, es minimizar I sobre $Z_e(\mathcal{A})$, donde

$$I(x,u) = \int_a^b L(t, x(t), u(t))dt \qquad ((x,u) \in Z).$$

El funcional L, será llamado Lagrangiano, los elementos de X trayectorias, de U controles, y de Z procesos. Un proceso será *admisible* si éste pertenece a $Z_e(\mathcal{A})$.

2 Soluciones del problema.

Un proceso (x, u) es una solución (global) de $P(\mathcal{A})$ si éste pertenece a

$$S(\mathcal{A}) = \{(x,u) \in Z_e(\mathcal{A}) \mid \ I(x,u) \leq I(y,v) \text{ para todo } (y,v) \in Z_e(\mathcal{A})\}.$$

Consideremos las siguientes normas:

$$||x||_0 = sup\{|x(t)| : t \in T\} \ (x \in X), \quad ||(x,u)||_1 = ||x||_0 + ||u||_0 \ ((x,u) \in Z).$$

Decimos que (x, u) es un *mínimo fuerte* de $P(\mathcal{A})$ si éste es un mínimo local de I sobre $Z_e(\mathcal{A})$ con respecto a $||\cdot||_0$, la norma fuerte en X, i.e., si existe $\epsilon > 0$ tal que $I(x,u) \leq I(y,v)$ para todo $(y,v) \in Z_e(\mathcal{A}) \cap (N_0(x;\epsilon) \times U)$, donde

$$N_0(x;\epsilon) = \{y \in X : ||x - y||_0 < \epsilon\}.$$

Un *mínimo débil* de $P(\mathcal{A})$ corresponde a reemplazar $N_0(x;\epsilon) \times U$ por

$$N_1((x,u);\epsilon) = \{(y,v) \in Z : ||(x,u) - (y,v)||_1 < \epsilon\}$$

en la definición anterior.

Notamos que, si definimos para todo $(x,u) \in Z$ y $\epsilon > 0$,

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

$$T_1((x,u);\epsilon) = \{(t,y,v) \in T_0(x;\epsilon) \times \mathbf{R}^m : |u(t) - v| < \epsilon\}$$

entonces (x,u) es un mínimo fuerte de $P(\mathcal{A})$ si, para algún $\epsilon > 0$,

$$(x,u) \in S((T_0(x;\epsilon) \times \mathbf{R}^m) \cap \mathcal{A})$$

y (x,u) es un mínimo débil de $P(\mathcal{A})$ si, para algún $\epsilon > 0$,

$$(x,u) \in S(T_1((x,u);\epsilon) \cap \mathcal{A}).$$

3 La primera y la segunda variaciones de I.

Para todo $(x,u) \in Z$ definimos la *primera variación de I a lo largo de (x,u)* por

$$I'((x,u);(y,v)) = \int_a^b \{\langle L_x(t,x(t),u(t)), y(t)\rangle \ + \langle L_u(t,x(t),u(t)), v(t)\rangle\} dt$$

y la *segunda variación de I a lo largo de (x,u)* por

$$I''((x,u);(y,v)) = \int_a^b 2\Omega(t,y(t),v(t))dt \qquad ((y,v) \in Z)$$

donde, para todo $(t,y,v) \in T \times \mathbf{R}^n \times \mathbf{R}^m$,

$$2\Omega(t,y,v) = \langle y, L_{xx}(t,x(t),u(t))y\rangle \ + \langle y, L_{xu}(t,x(t),u(t))v\rangle$$
$$+\langle v, L_{uu}(t,x(t),u(t))v\rangle.$$

En esta sección se prueba que las variaciones de I coinciden con los diferenciales de Gateaux y Fréchet de I (para el último con respecto a la norma débil) bajo las hipótesis de suavidad usuales sobre el Lagrangiano.

Proposición 3.1 *Sea $(x,u) \in Z(\mathcal{A})$, L continua en $\mathcal{A}$. Entonces, para todo $(y,v) \in Z$, $I'((x,u);(y,v))$ y $I''((x,u);(y,v))$ coinciden con el primer y segundo diferenciales de Gateaux de I (restringidos a $Z(\mathcal{A})$) en (x,u) en la dirección (y,v), siempre y cuando las derivadas de L involucradas sean continuas sobre $\mathcal{A}$.*

Prueba. Sea $(y,v) \in Z$. Puesto que $\mathcal{A}$ es abierto, existe $\delta > 0$ tal que $(x,u) + \epsilon(y,v) \in Z(\mathcal{A})$ para todo $|\epsilon| < \delta$. Sea

$$F(t,\epsilon) = L(t, x(t) + \epsilon y(t), u(t) + \epsilon v(t)) \qquad (t \in T,\ |\epsilon| < \delta).$$

Sean $t'_1, \ldots, t'_k$ aquellos puntos donde tanto u como v son discontinuas, $t'_0 = a$, $t'_{k+1} = b$. Por la continuidad de $F(t,\epsilon)$ y $F_\epsilon(t,\epsilon)$ sobre cada intervalo $[t'_i, t'_{i+1}]$ $(i = 0, 1, \ldots, k)$ tenemos que, sumando desde 0 hasta k,

$$\frac{d}{d\epsilon} I((x,u) + \epsilon(y,v))|_{\epsilon=0} = \int_a^b F_\epsilon(t,0)dt = I'((x,u);(y,v)).$$

De manera similar,

$$\frac{d^2}{d\epsilon^2} I((x,u) + \epsilon(y,v))|_{\epsilon=0} = \int_a^b F_{\epsilon\epsilon}(t,0)dt = I''((x,u);(y,v)).$$

Proposición 3.2 *Sea* $(x,u) \in Z(\mathcal{A})$, $L \in C^2(\mathcal{A}; x, u)$. *Entonces, para todo* $(y,v) \in Z$, $I'((x,u);(y,v))$ *y* $I''((x,u);(y,v))$ *coinciden con el primer y segundo diferenciales de Fréchet de* I *(restringidos a* $Z(\mathcal{A})$*) en* (x,u) *en la dirección* (y,v)*, con respecto a la norma débil.*

Prueba. Sea $z_0 = (x_0, u_0) \in Z(\mathcal{A})$. Puesto que $\{(t, x_0(t), u_0(t)) \mid t \in T\}$ es compacto en $T \times \mathbf{R}^n \times \mathbf{R}^m$ y $\mathcal{A}$ es abierto, existe $\rho > 0$ tal que $V = cl(T_1(z_0;\rho)) \subset \mathcal{A}$, donde $cl(C)$ denota la cerradura de C. Sea $z = (x,u) \in N_1(z_0;\rho)$. Entonces para todo $t \in T$, $(t, x(t), u(t)) \in T_1(z_0;\rho) \subset \mathcal{A}$, y entonces $z \in Z(\mathcal{A})$. Esto muestra que $Z(\mathcal{A})$ es un subconjunto abierto de $(Z, \|\cdot\|_1)$.

Ahora, puesto que L y su primera y segunda derivadas parciales con respecto a x y u son continuas sobre $\mathcal{A}$, ellas son uniformemente continuas sobre V. Por lo tanto, si $\|z - z_0\|_1 \to 0$, entonces $L(t, x(t), u(t)) \to L(t, x_0(t), u_0(t))$ uniformemente sobre T, y lo mismo se aplica para las derivadas. Esto implica que para todo $\epsilon > 0$ existe $\delta > 0$ tal que, para todo $w = (y,v) \in Z$ y $z \in N_1(z_0;\delta)$,

$$|I'(z;w) - I'(z_0;w)| < \epsilon\|w\|_1, \qquad |I''(z;w) - I''(z_0;w)| < \epsilon\|w\|_1^2.$$

En consecuencia, si $\|z - z_0\|_1 \to 0$, entonces $I'(z;w) \to I'(z_0;w)$ y $I''(z;w) \to I''(z_0;w)$ uniformemente para $\|w\|_1 \leq 1$.

Por el Teorema de Taylor aplicado a $\lambda \mapsto I(z_0 + \lambda(z - z_0))$ $(\lambda \in [0,1])$, $z \in N_1(z_0;\delta)$,

$$I(z) = I(z_0) + I'(z_0; z - z_0) + P_1(z_0, z - z_0)$$

donde

$$P_1(z_0, w) = \int_0^1 \{I'(z_0 + \lambda w; w) - I'(z_0; w)\}d\lambda = \int_0^1 (1-\lambda)I''(z_0 + \lambda w; w)d\lambda$$

y

$$I(z) = I(z_0) + I'(z_0; z - z_0) + \frac{1}{2}I''(z_0; z - z_0) + P_2(z_0, z - z_0)$$

donde

$$P_2(z_0, w) = \int_0^1 (1-\lambda)\{I''(z_0 + \lambda w; w) - I''(z_0; w)\}d\lambda.$$

El resultado se sigue ya que

$$\lim_{z \to z_0} \frac{P_1(z_0, z - z_0)}{\|z - z_0\|_1} = 0 \qquad y \qquad \lim_{z \to z_0} \frac{P_2(z_0, z - z_0)}{\|z - z_0\|_1{}^2} = 0.$$

3.1 Condiciones necesarias.

Definimos el conjunto de variaciones admisibles por

$$Y = \{(y, v) \in Z \mid \dot{y}(t) = A(t)y(t) + B(t)v(t) \quad (t \in T),\ y(a) = y(b) = 0\}.$$

Consideremos los siguientes conjuntos:

$$\begin{aligned}
E &= \{(x,u) \in Z \mid I'((x,u);(y,v)) = 0 \ \text{ para todo } (y,v) \in Y\} \\
H &= \{(x,u) \in Z \mid I''((x,u);(y,v)) \geq 0 \ \text{ para todo } (y,v) \in Y\} \\
L &= \{(x,u) \in Z \mid L_{uu}(t, x(t), u(t)) \geq 0 \ \text{ para todo } t \in T\} \\
W(\mathcal{A}) &= \{(x,u) \in Z(\mathcal{A}) \mid \mathcal{E}(t, x(t), u(t), v) \geq 0 \ \text{ para todo } (t, x(t), v) \in \mathcal{A}\} \\
H' &= \{(x,u) \in Z \mid I''((x,u);(y,v)) > 0 \ \text{ para todo } (y,v) \in Y,\ (y,v) \neq 0\} \\
L' &= \{(x,u) \in Z \mid L_{uu}(t, x(t), u(t)) > 0 \ \text{ para todo } t \in T\} \\
W(\mathcal{A};\epsilon) &= \{(x_0,u_0) \in Z(\mathcal{A}) \mid \ \mathcal{E}(t, x, u, v) \geq 0 \text{ para todo } (t, x, v) \in \mathcal{A}, \\
&\qquad (t, x, u) \ \in T_1((x_0, u_0); \epsilon)\}
\end{aligned}$$

donde la " función de exceso " de Weierstrass $\mathcal{E}$ está dada por la fórmula

$$\mathcal{E}(t, x, u, v) = L(t, x, v) - L(t, x, u) - \langle v - u, L_u(t, x, u)\rangle.$$

Los elementos de E serán llamados *extremos*, los elementos de L se dice que satisfacen la *condición de Legendre*, y los elementos de $W(\mathcal{A})$ que satisfacen la *condición de Weierstrass*. Mientras que los elementos de L' y $W(\mathcal{A};\epsilon)$, se dice que satisfacen las condiciones *modificadas* de Legendre y Weierstrass respectivamente.

Teorema 3.1 *Sea $(x,u) \in S(\mathcal{A})$, L continua sobre $\mathcal{A}$. Entonces (x,u) está en E, H, siempre y cuando las derivadas de L involucradas (en cada caso) sean continuas. En particular, si $L \in C^2(\mathcal{A}; x, u)$, entonces $S(\mathcal{A}) \subset E \cap H$.*

Prueba. (i): " $L \in C^1(\mathcal{A}; x, u) \Rightarrow S(\mathcal{A}) \subset E$ ". Sea $(x,u) \in S(\mathcal{A})$, $(y,v) \in Y$. Puesto que $(x,u) \in Z_e(\mathcal{A})$, (x,u) es un punto interior de $Z_e(\mathcal{A})$ en la dirección (y,v). Puesto que I tiene un mínimo sobre $Z_e(\mathcal{A})$ en (x,u), se sigue por la Proposición 3.1 que $I'((x,u);(y,v)) = 0$.

(ii): " $L \in C^2(\mathcal{A}; x, u) \Rightarrow S(\mathcal{A}) \subset H$ ". Se sigue como en (i), por la Proposición 3.1.

1 La ecuación de Euler.

El propósito de esta sección es el de caracterizar el conjunto E de extremos en términos de una cierta ecuación diferencial, análoga a la ecuación de Euler en el cálculo de variaciones, y derivar ciertas propiedades de sus elementos.

Lema 3.1 *Sea K una función que mapea $T \times \mathbf{R}^n \times \mathbf{R}^m \to \mathbf{R}$, sea*

$$I(x) = \int_a^b K(t, x(t), \dot{x}(t))dt \qquad (x \in X),$$

y consideremos la primera variación de I a lo largo de $x \in X$:

$$I'(x; y) = \int_a^b \{\langle K_x(t, x(t), \dot{x}(t)), y(t)\rangle + \langle K_{\dot{x}}(t, x(t), \dot{x}(t)), \dot{y}(t)\rangle\}dt \ (y \in X).$$

Para todo $x \in X$, las siguientes condiciones son equivalentes:

a. *$I'(x; y) = 0$ para todo $y \in X$ con $y(a) = y(b) = 0$.*
b. *Existe una constante $c \in \mathbf{R}^n$ tal que*

$$K_{\dot{x}}(t, x(t), \dot{x}(t)) = \int_a^t K_x(s, x(s), \dot{x}(s))ds + c \qquad (t \in T).$$

Este resultado se sigue del Teorema 2.14.

Definición 3.1 *Para todo $t \in T$, $u \in \mathbf{R}^m$, $x, p \in \mathbf{R}^n$, sea*

$$H(t, x, u, p) = \langle p, A(t)x + B(t)u\rangle - L(t, x, u)$$

llamado el Hamiltoniano, y sea $\mathcal{X}$ el conjunto de funciones que mapean T a $\mathbf{R}^n$. Definimos el conjunto

$$R = \{(x, u, p) \in Z \times \mathcal{X} \mid H_u(t, x(t), u(t), p(t)) = 0 \ (t \in T)\}.$$

En lo que sigue supondremos lo siguiente:

$$\left\{ \begin{array}{l} \text{Para todo } y \in X \text{ con } y(a) = y(b) = 0 \\ \text{existe } v \in U \text{ tal que } (y, v) \in D. \end{array} \right. \qquad (\dagger)$$

Teorema 3.2 *Sea $(x_0, u_0, p) \in R$ y consideremos las siguientes afirmaciones:*

a. $(x_0, u_0) \in E$.

b. *Existe una constante $c \in \mathbf{R}^n$ tal que*

$$p(t) = \int_a^t -H_x(s, x_0(s), u_0(s), p(s))ds + c \qquad (t \in T).$$

Entonces $(b) \Rightarrow (a)$ y, si $(\dagger)$ se cumple, entonces $(a) \Rightarrow (b)$.

Prueba. Para todo $t \in T$ y $x, \dot{x} \in \mathbf{R}^n$, sea

$$\begin{aligned} K(t,x,\dot{x}) &= \langle p(t), \dot{x}\rangle - H(t,x,u_0(t),p(t)) \\ &= L(t,x,u_0(t)) + \langle p(t), \dot{x} - A(t)x - B(t)u_0(t)\rangle \end{aligned}$$

y sea

$$J(x) = \int_a^b K(t,x(t),\dot{x}(t))dt \qquad (x \in X).$$

Por el Lema 3.1 sabemos que (b) es equivalente a

c. $J'(x_0;y) = 0$ para todo $y \in X$ con $y(a) = y(b) = 0$.

Ahora, $(x_0,u_0,p) \in R \Leftrightarrow B^T(t)p(t) = L_u(t,x_0(t),u_0(t))$ $(t \in T)$. De esto se sigue que, para todo $(y,v) \in D$

$$\begin{aligned} I'((x_0,u_0);(y,v)) &= \int_a^b \{\langle L_x(t,x_0(t),u_0(t)), y(t)\rangle \\ &\qquad + \langle L_u(t,x_0(t),u_0(t)), v(t)\rangle\}dt \\ &= \int_a^b \{\langle L_x(t,x_0(t),u_0(t)) - A(t)^T p(t), y(t)\rangle \\ &\qquad + \langle p(t), A(t)y(t) + B(t)v(t)\rangle\}dt \\ &= \int_a^b \{\langle K_x(t,x_0(t),\dot{x}_0(t)), y(t)\rangle \\ &\qquad + \langle K_{\dot{x}}(t,x_0(t),\dot{x}_0(t)), \dot{y}(t)\rangle\}dt \\ &= J'(x_0;y). \end{aligned}$$

En consecuencia, $(b) \Leftrightarrow (c) \Rightarrow (a)$. Las desigualdades anteriores claramente implican que, si $(\dagger)$ se cumple, entonces $(a) \Rightarrow (c) \Leftrightarrow (b)$.

Se ha probado que, si $(\dagger)$ se cumple, y $(x,u) \in Z$ y existe $p \in \mathcal{X}$ tal que

$$B(t)^T p(t) = L_u(t,x(t),u(t)) \qquad (t \in T)$$

entonces $(x,u) \in E \Leftrightarrow \dot{p}(t) + A(t)^T p(t) = L_x(t,x(t),u(t))$ $(t \in T)$. Esta ecuación diferencial será llamada la ecuación de *Euler*. Si u tiene una discontinuidad, ésta es satisfecha por límites laterales izquierdo y derecho.

El lagrangiano L se dice *regular* en $\mathcal{A}$ si $\mathcal{A}$ es convexo en u, y para todo $(t,x,u) \in \mathcal{A}$, $L_{uu}(t,x,u) > 0$. Esto implica que, para (t,x) fijo, $L(t,x,\cdot)$ es una función estrictamente convexa, lo cual es equivalente a decir que $L_u(t,x,\cdot)$ es una función estrictamente creciente.

Un proceso $(x,u) \in Z$ es llamado no *singular* si $|L_{uu}(t,x(t),u(t))| \neq 0$ para todo $t \in T$.

Proposición 3.3 *Sea $(x,u,p) \in R$ y supongamos que $(\dagger)$ se cumple. Si $(x,u) \in E$ entonces $L_u(\cdot,x(\cdot),u(\cdot))$ es continua. En particular, si t es un punto de discontinuidad de u, entonces*

$$L_u(t,x(t),u(t-0)) = L_u(t,x(t),u(t+0)).$$

Prueba. Esto es inmediato puesto que p satisface 3.2b y por lo tanto ésta es continua.

Proposición 3.4 *Sea $(x, u) \in W(\mathcal{A})$ y supongamos que $L_u(\cdot, x(\cdot), u(\cdot))$ es continua. Sea*

$$F(t, v) = L(t, x(t), v) - \langle v, L_u(t, x(t), u(t))\rangle \qquad (t \in T,\ v \in \mathbf{R}^m).$$

Entonces $F(\cdot, u(\cdot))$ es continua, y

$$\mathcal{E}(t, x(t), u(t-0), u(t+0)) = \mathcal{E}(t, x(t), u(t+0), u(t-0)) = 0.$$

Prueba. Puesto que $(x, u) \in W(\mathcal{A})$,

$$\begin{aligned} 0 &\le \mathcal{E}(t, x(t), u(t-0), u(t+0)) \\ &= L(t, x(t), u(t+0)) - L(t, x(t), u(t-0)) \\ &\qquad\qquad -\langle u(t+0) - u(t-0), L_u(t, x(t), u(t-0))\rangle \\ &= F(t, u(t+0)) - F(t, u(t-0)). \end{aligned}$$

De manera similar,

$$0 \le \mathcal{E}(t, x(t), u(t+0), u(t-0)) = F(t, u(t-0)) - F(t, u(t+0))$$

y entonces $F(t, u(t+0)) = F(t, u(t-0))$.

Proposición 3.5 *Sea $(x, u) \in Z$ y $p \in \mathcal{X}$, y consideremos las siguientes afirmaciones:*

a. *$(x, u) \in E \cap W(\mathcal{A})$ y $(x, u, p) \in R$.*
b. *(x, u, p) satisface el principio del máximo de Pontryagin con un multiplicador de Lagrange de costo positivo, i.e.,*
1. *$\dot p(t) = -H_x(t, x(t), u(t), p(t))$ para todo $t \in T$.*
2. *$H(t, x(t), u(t), p(t)) \ge H(t, x(t), v, p(t))$ para $v \in \mathbf{R}^m$, $(t, x(t), v) \in \mathcal{A}$.*

Entonces $(b) \Rightarrow (a)$ y, si $(\dagger)$ se cumple, entonces $(a) \Rightarrow (b)$.

Prueba. Observamos que

$$L(t, x, v) - L(t, x, u) - \langle v - u, B(t)^T p\rangle = H(t, x, u, p) - H(t, x, v, p)$$
$$L(t, x, v) - L(t, x, u) - \langle v - u, L_u(t, x, u)\rangle = \mathcal{E}(t, x, u, v).$$

$(b) \Rightarrow (a)$: Por $(b2)$, $B(t)^T p(t) = L_u(t, x(t), u(t))$ para todo $t \in T$. Por lo tanto, $(x, u, p) \in R$ y $(x, u) \in W(\mathcal{A})$. Por $(b1)$ y el Teorema 3.2, $(x, u) \in E$.

$(a) \Rightarrow (b)$: Puesto que $B(t)^T p(t) = L_u(t, x(t), u(t))$ y $(x, u) \in W(\mathcal{A})$, tenemos $(b2)$. Por el Teorema 3.2, si $(\dagger)$ se cumple, $(x, u) \in E \Rightarrow (b1)$.

Proposición 3.6 *Sea $(x, u) \in Z(\mathcal{A})$ no singular con u continua, y supongamos que $p \in \mathcal{X}$ es tal que, para todo $t \in T$,*

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

$$\dot{p}(t) = -H_x(t, x(t), u(t), p(t))$$

$$0 = H_u(t, x(t), u(t), p(t)).$$

Si $A, B \in C^r(T)$ y $L_x, L_u \in C^r(\mathcal{A})$ para algún $r \geq 1$, entonces $x, p \in C^{r+1}(T)$ y $u \in C^r(T)$.

Prueba. Para $t \in T$ y $v \in \mathbf{R}^m$ sea

$$G(t, v) = H_u(t, x(t), v, p(t)) = B(t)^T p(t) - L_u(t, x(t), v).$$

Supongamos que $x, p \in C^i(T)$ para $0 \leq i \leq r$. Entonces $G \in C^i$. Tenemos que $G(t, u(t)) = 0$ y $|G_v(t, u(t))| \neq 0$ para todo $t \in T$. También, u y G_v son continuas. Por el Teorema de la función implícita, $u \in C^i(T)$ y, puesto que

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) \qquad y \qquad \dot{p}(t) = L_x(t, x(t), u(t)) - A(t)^T p(t),$$

$x, p \in C^{i+1}(T)$.

Proposición 3.7 *Sea L regular en $\mathcal{A}$ y $L \in C^2(\mathcal{A}; u)$. Entonces*

a. *$E(t, x, u, v) > 0$ para todo $(t, x, u) \in \mathcal{A}$ y $(t, x, v) \in \mathcal{A}$ con $u \neq v$.*
b. *Si $(x, u) \in E \cap Z(\mathcal{A})$, $(x, u, p) \in R$ y $(\dagger)$ se cumple, entonces u es continua.*

Prueba. (a): Por el Teorema de Taylor, dado (t, x, u) y (t, x, v) en $\mathcal{A}$ con $u \neq v$, existe $\lambda \in (0, 1)$ tal que

$$\begin{aligned}
\mathcal{E}(t, x, u, v) &= L(t, x, v) - L(t, x, u) - \langle v - u, L_u(t, x, u) \rangle \\
&= \frac{1}{2} \langle v - u, L_{uu}(t, x, u + \lambda(v - u))(v - u) \rangle \\
&> 0.
\end{aligned}$$

(b) : Por la proposición 3.3, $L_u(\cdot, x(\cdot), u(\cdot))$ es continua. El resultado se sigue por (a) y la proposición 3.4.

Proposición 3.8 *Supongamos que $L \in C^2(\mathcal{A}; u)$, (x_0, u_0) es un proceso no singular en $W(\mathcal{A}; \epsilon)$ para algún $\epsilon > 0$. Entonces ϵ puede ser reducido de manera que la desigualdad en la definicón de $W(\mathcal{A}; \epsilon)$ llegue a ser estricta, i.e.,*

$$\mathcal{E}(t, x, u, v) > 0 \text{ para todo } (t, x, u) \in T_1((x_0, u_0); \epsilon), \quad (t, x, v) \in \mathcal{A}, \ u \neq v.$$

Prueba. Reducimos ϵ de manera que $|L_{uu}(t,x,u)| \neq 0$ para todo $(t,x,u) \in T_1((x_0,u_0);\epsilon))$. Supongamos que el resultado es falso para ese valor de ϵ, i.e., existe (t,x,u,v), con (t,x,u) en $T_1((x_0,u_0);\epsilon)$, $(t,x,v) \in \mathcal{A}$ y $u \neq v$, tal que $\mathcal{E}(t,x,u,v) = 0$. Sea $f(w) = \mathcal{E}(t,x,w,v)$ para todo $w \in \mathbf{R}^m$. Entonces f tiene un mínimo local en $w = u$ y entonces

$$0 = f'(u) = -L_{uu}(t,x,u)(v-u).$$

Pero esto implica que $u = v$ contradiciendo nuestra suposición.

2 La condición de Jacobi.

Para todo $(x,u) \in Z$ denotemos por $J_{(x,u)}$ la segunda variación de I a lo largo de (x,u), i.e.,

$$J_{(x,u)}(y,v) = I''((x,u);(y,v)) = \int_a^b 2\Omega(t,y(t),v(t))dt \qquad ((y,v) \in Z),$$

y sea $E(x,u)$ el conjunto de procesos que son extremos para el integrando Ω, i.e.,

$$\begin{aligned} E(x,u) &= \{(y,v) \in Z \mid J'_{(x,u)}((y,v);(z,w)) = 0 \text{ para todo } (z,w) \in Y\} \\ &= \{(y,v) \in Z \mid \int_a^b \{\langle \Omega_y(t,y(t),v(t)), z(t)\rangle \\ &\qquad + \langle \Omega_v(t,y(t),v(t)), w(t)\rangle\}dt = 0 \text{ para todo } (z,w) \in Y\}. \end{aligned}$$

Notamos que por el Teorema 3.2, si $(y,v) \in Z$ y $q \in X$ satisfacen

$$B(t)^T q(t) = \Omega_v(t,y(t),v(t)) \qquad (t \in T),$$

entonces $\dot q(t) + A(t)^T q(t) = \Omega_y(t,y(t),v(t))$ para todo $t \in T \Rightarrow (y,v) \in E(x,u)$ y, si (†) se cumple, el recíproco es cierto. Observamos también que, dado $(x,u) \in Z$,

$$\Omega_y(t,y,v) = L_{xx}(t,x(t),u(t))y + L_{xu}(t,x(t),u(t))v$$
$$\Omega_v(t,y,v) = L_{ux}(t,x(t),u(t))y + L_{uu}(t,x(t),u(t))v.$$

Puesto que $\Omega_{vv}(t,y,v) = L_{uu}(t,x(t),u(t))$ se sigue que, si $(x,u) \in L'$, entonces el integrando Ω es regular y, por la proposición 3.7, si $L \in C^2(\mathcal{A};u)$, (†) se cumple, $(y,v) \in E(x,u)$, y existe $q \in X$ tal que $B(t)^T q(t) = \Omega_v(t,y(t),v(t))$ $(t \in T)$, entonces y es C^1 y v es continua.

El resultado siguiente se sigue por las observaciones anteriores y por substitución directa.

Proposición 3.9 *Sea $(x,u) \in L'$, $(y,v) \in Z$, $q \in X$, y supongamos que*

$$B(t)^T q(t) = L_{ux}(\tilde x(t))y(t) + L_{uu}(\tilde x(t))v(t) \qquad (t \in T)$$

donde $\tilde{x}(t)$ denota $(t, x(t), u(t))$. Consideremos las siguientes afirmaciones:

a. $(y, v) \in E(x, u) \cap D$.
b. $(y, v) \in D$ *y* $\dot{q}(t) + A(t)^T q(t) = L_{xx}(\tilde{x}(t))y(t) + L_{xu}(\tilde{x}(t))v(t) \quad (t \in T)$
c. (y, q) *satisface, para todo* $t \in T$, *el sistema lineal*

$$\dot{y}(t) = \alpha(t)y(t) + \beta(t)q(t)$$
$$\dot{q}(t) = \gamma(t)y(t) - \alpha^T(t)q(t)$$

el cual denotamos por (JE), donde

$$\begin{aligned} \alpha(t) &= A(t) - B(t)L_{uu}^{-1}(\tilde{x}(t))L_{ux}(\tilde{x}(t)) \\ \beta(t) &= B(t)L_{uu}^{-1}(\tilde{x}(t))B(t)^T \\ \gamma(t) &= L_{xx}(\tilde{x}(t)) - L_{xu}(\tilde{x}(t))L_{uu}^{-1}(\tilde{x}(t))\ L_{ux}(\tilde{x}(t)). \end{aligned}$$

Entonces $(c) \Leftrightarrow (b) \Rightarrow (a)$ y, si $(\dagger)$ se cumple, las tres afirmaciones son equivalentes.

Definición 3.2 *Dado $(x, u) \in Z$, diremos que un punto $s \in (a, b]$ es conjugado a a sobre (x, u) si existen $(y, v) \in E(x, u) \cap D$ y $q \in X$ que satisfacen*

$$B(t)^T q(t) = L_{ux}(\tilde{x}(t))y(t) + L_{uu}(\tilde{x}(t))v(t) \qquad (t \in T)$$

tal que $y(a) = y(s) = 0$ y $y(t) \not\equiv 0$ sobre (a, s).

Notamos que, si $(x, u) \in L'$, entonces $s \in (a, b]$ es conjugado a a sobre (x, u) si existe $(y, q) \in X \times X$ que satisface (JE) con $y(a) = y(s) = 0$ y $y(t) \not\equiv 0$ sobre (a, s). También, si $(\dagger)$ se cumple, las dos afirmaciones son equivalentes.

Diremos que $(x, u) \in Z$ satisface la *condición de Jacobi* si éste pertenece a

$$J = \{(x, u) \in Z \mid s \in (a, b) \Rightarrow s \text{ no es conjugado a } a \text{ sobre } (x, u)\}.$$

Teorema 3.3 *Supongamos que $L \in C^2(\mathcal{A}; x, u)$, $(\dagger)$ se cumple, $(x, u) \in L' \cap Z(\mathcal{A})$, y u es continua. Si $(x, u) \in H$, entonces $(x, u) \in J$.*

Prueba. Supongamos que $(x, u) \notin J$. Por definición, existe $s \in (a, b)$ conjugado a a sobre (x, u). Sea (y, v) en $E(x, u) \cap D$ que satifaga $y(a) = y(s) = 0$ y $y(t) \not\equiv 0$ sobre (a, s), y sea $q \in X$ que satisfaga

$$B(t)^T q(t) = L_{ux}(\tilde{x}(t))y(t) + L_{uu}(\tilde{x}(t))v(t) \qquad (t \in T).$$

Definimos

$$(z(t), w(t), r(t)) = \begin{cases} (y(t), v(t), q(t)) & t \in [a, s) \\ (0, 0, 0) & t \in [s, b] \end{cases}$$

Entonces

$$\begin{aligned} J_{(x,u)}(z,w) &= \int_a^b 2\Omega(t,z(t),w(t))dt \\ &= \int_a^s \{\langle z(t),\Omega_y(t,z(t),w(t))\rangle + \langle w(t),\Omega_v(t,z(t),w(t))\rangle\}dt \\ &= \int_a^s \{\langle z(t),\dot{q}(t)\rangle + \langle z(t),A(t)^T q(t)\rangle + \langle w(t),B(t)^T q(t)\rangle\}dt \\ &= \int_a^s \frac{d}{dt}\langle z(t),q(t)\rangle dt \\ &= 0. \end{aligned}$$

Puesto que $(x,u) \in H$ y $(z,w) \in Y$, (z,w) minimiza $J_{(x,u)}$ sobre Y. Puesto que x y u son continuas, las funciones Ω, Ω_y, y Ω_v son continuas y entonces, por el Teorema 3.1, $(z,w) \in E(x,u)$. En consecuencia, tenemos que $(z,w) \in E(x,u) \cap D$ y

$$B(t)^T r(t) = L_{ux}(\tilde{x}(t))z(t) + L_{uu}(\tilde{x}(t))w(t) \qquad (t \in T).$$

Como (†) se cumple, se sigue por la proposición 3.9 que (z,r) satisface, para todo $t \in T$, el sistema lineal

$$\dot{z}(t) = \alpha(t)z(t) + \beta(t)r(t)$$

$$\dot{r}(t) = \gamma(t)z(t) - \alpha(t)^T r(t).$$

Puesto que $z(s) = r(s) = 0$, tenemos que $z(t) \equiv r(t) \equiv 0$ sobre T. Esto contradice la no nulidad de y.

Los resultados anteriores y el Teorema 3.1 implican la siguiente condición necesaria para un mínimo.

Teorema 3.4 *Supongamos que $L \in C^2(\mathcal{A};x,u)$, $(x,u) \in S(\mathcal{A})$ con (x,u) no singular,* (†) *se cumple, y u es continua, entonces $(x,u) \in J$.*

3.2 Condiciones suficientes.

El hecho de que I' y I'' son el primer y segundo diferenciales de Fréchet de I restringidos a $Z(\mathcal{A})$ con respecto a $(Z, \|\cdot\|_1)$, tenemos de manera natural, las condiciones suficientes para un mínimo débil.

Teorema 3.5 *Sea $L \in C^2(\mathcal{A};x,u)$, $(x_0,u_0) \in Z_e(\mathcal{A}) \cap E$. Si, para algún $\epsilon > 0$, $I''((x,u);(y,v)) > 0$ para todo $(x,u) \in Z_e(\mathcal{A})$ con $\|(x,u)-(x_0,u_0)\|_1 < \epsilon$ y todo $(y,v) \in Y - \{0\}$, entonces (x_0,u_0) es un mínimo estricto débil para $P(\mathcal{A})$.*

Prueba. Sea $z_0 = (x_0, u_0)$ y disminuimos ϵ como en la prueba de la Proposición 3.2, tal que

$$I(z) = I(z_0) + I'(z_0; z - z_0) + \int_0^1 (1-\lambda) I''(z_0 + \lambda(z - z_0);\ z - z_0) d\lambda$$

para todo $z \in N_1(z_0; \epsilon)$. Se sigue que, si $z \neq z_0$ está en $Z_e(\mathcal{A})$ y $\|z - z_0\|_1 < \epsilon$, entonces $I(z) > I(z_0)$.

El siguiente resultado da condiciones suficientes para un mínimo débil o fuerte de $P(\mathcal{A})$. Este se cumple si extendemos la clase de procesos admisibles a la clase de parejas (x, u) con x absolutamente continua y u medible, cuyos elementos $(t, x(t), u(t))$ están en $\mathcal{A}$ para casi toda $t \in T$, y para los cuales $L(t, x(t), u(t))$ es integrable sobre T.

1 Una prueba indirecta.

Teorema 3.6 *Supongamos que* $L \in C^2(\mathcal{A}; x, u)$ *y* $(x_0, u_0) \in Z_e(\mathcal{A})$ *con* x_0 *de clase* C^1. *Entonces:*

a. $(x_0, u_0) \in E \cap H' \cap L' \Rightarrow (x_0, u_0)$ *es un mínimo estricto débil de* $P(\mathcal{A})$.
b. $(x_0, u_0) \in E \cap H' \cap L' \cap W(\mathcal{A}; \epsilon)$ *para algún* $\epsilon > 0 \Rightarrow (x_0, u_0)$ *es un mínimo estricto fuerte de* $P(\mathcal{A})$.

Prueba. Empezamos probando que (b) implica (a). Sea $(x_0, u_0) \in E \cap H' \cap L'$. Puesto que $L_{uu}(t, x_0(t), u_0(t)) > 0$ para toda $t \in T$, existe $\epsilon > 0$ tal que $L_{uu}(t, x, u) \geq 0$ para toda $(t, x, u) \in T_1((x_0, u_0); \epsilon) \subset \mathcal{A}$. Sea (t, x, u) y (t, x, v) en $T_1((x_0, u_0); \epsilon)$. Puesto que

$$(t, x, \lambda v + (1-\lambda)u) = (t, x, u + \lambda(v - u)) \in T_1((x_0, u_0); \epsilon) \qquad (\lambda \in [0, 1])$$

tenemos que, por el Teorema de Taylor,

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

y entonces $(x_0, u_0) \in W(T_1((x_0, u_0); \epsilon); \epsilon)$. Suponiendo que (b) es verdadero, se sigue que (x_0, u_0) es un mínimo estricto fuerte de $P(T_1((x_0, u_0); \epsilon))$, i.e., para algún $\delta > 0$, $I(x_0, u_0) < I(x, u)$ para todo $(x, u) \neq (x_0, u_0)$ con $(x, u) \in Z_e(T_1((x_0, u_0); \epsilon))$ y $\|x - x_0\|_0 < \delta$. Tomando $\mu = min\{\epsilon, \delta\}$, tenemos que $I(x_0, u_0) < I(x, u)$ para todo $(x, u) \neq (x_0, u_0)$ teniendo que $(x, u) \in Z_e(\mathcal{A})$ y $\|(x, u) - (x_0, u_0)\|_1 < \mu$. Esto implica que (x_0, u_0) es un mínimo estricto débil de $P(\mathcal{A})$ y entonces (a) se cumple.

Ahora, supongamos que $L \in C^2(\mathcal{A}; x, u)$, $(x_0, u_0) \in Z_e(\mathcal{A})$ con x_0 de clase C^1, y $(x_0, u_0) \in E \cap H' \cap L' \cap W(\mathcal{A}; \epsilon)$. Queremos probar que (x_0, u_0) es un mínimo estricto fuerte de $P(\mathcal{A})$.

Denotemos por X' la clase de todas las funciones absolutamente continuas que mapean T a $\mathbf{R}^n$, por U' la clase de todas las funciones medibles que mapean T a $\mathbf{R}^m$, y definimos $Z' = X' \times U'$,

$$Z'(\mathcal{A}) = \{(x,u) \in Z' \mid (t, x(t), u(t)) \in \mathcal{A} \text{ a.e. en } T,\ L(t,x(t),u(t)) \text{ integrable } \}$$

$$Z'_e(\mathcal{A}) = \{(x,u) \in Z'(\mathcal{A}) \mid \dot{x}(t) = A(t)x(t) + B(t)u(t) \text{ a.e. en } T,$$
$$x(a) = \xi_0,\ x(b) = \xi_1\}.$$

Consideremos la norma sobre X':

$$\|x\| = \|x\|_0 = sup\{|x(t)| : t \in T\} \qquad (x \in X').$$

Vamos a probar que existe $\rho, \delta > 0$ tal que, para todo $(y,v) \in Z'_e(\mathcal{A})$ con $\|y - x_0\| < \rho$,

$$I(y,v) \geq I(x_0,u_0) + \delta D(v - u_0)$$

donde, para todo $u \in U'$ y $c \in \mathbf{R}^m$,

$$D(u) = \int_a^b \varphi(u(t))dt, \qquad \varphi(c) = (1 + |c|^2)^{1/2} - 1.$$

La prueba consiste en mostrar que, si suponemos lo contrario, i.e., para todo $\rho, \delta > 0$ existe $(y,v) \in Z'_e(\mathcal{A})$ con $\|y - x_0\| < \rho$, tal que

$$I(y,v) < I(x_0,u_0) + \delta D(v - u_0) \tag{1}$$

entonces $(x_0,u_0) \notin H'$.

Sea $z_0 = (x_0,u_0)$. Observamos primero que

$$I(z) = I(z_0) + I'(z_0; z - z_0) + K(z) + \mathcal{E}^*(z) \qquad (z = (x,u) \in Z'(\mathcal{A}))$$

donde

$$\mathcal{E}^*(x,u) = \int_a^b \mathcal{E}(t, x(t), u_0(t), u(t))dt,$$

$$K(x,u) = \int_a^b \{M(t,x(t)) \ + \langle u(t) - u_0(t), N(t,x(t))\rangle\}dt,$$

$$M(t,y) = L(t,y,u_0(t)) - L(t,x_0(t),u_0(t)) \ - \langle y - x_0(t), L_x(t,x_0(t),u_0(t))\rangle,$$

$$N(t,y) = L_u(t,y,u_0(t)) - L_u(t,x_0(t),u_0(t)).$$

Por el Teorema de Taylor, podemos seleccionar $\mu > 0$ tal que, para todo $(t,y) \in T_0(x_0;\mu)$,

$$M(t,y) = \frac{1}{2}\langle y - x_0(t), P(t,y)(y - x_0(t))\rangle, \qquad N(t,y) = Q(t,y)(y - x_0(t))$$

donde

$$P(t,y) = 2\int_0^1 (1-\lambda)L_{xx}(t, x_0(t) + \lambda(y - x_0(t)), u_0(t))d\lambda,$$

$$Q(t,y) = \int_0^1 L_{ux}(t, x_0(t) + \lambda(y - x_0(t)), u_0(t))d\lambda.$$

Ahora, probaremos que existen $\delta_0, \alpha, h > 0$ tales que

$$\mathcal{E}^*(x,u) \geq hD(u-u_0), \qquad |K(x,u)| \leq \alpha\|x - x_0\|[1 + D(u-u_0)] \tag{2}$$

para todo $(x,u) \in Z'(\mathcal{A})$ que satisface $\|x - x_0\| < \delta_0$. La primera de estas relaciones se verificará una vez que probemos el siguiente resultado auxiliar.

Lema 3.2 *Supongamos que* $L \in C^2(\mathcal{A}; x, u)$, $(x_0, u_0) \in L' \cap W(\mathcal{A}; \epsilon)$ *para algún* $\epsilon > 0$. *Entonces existe* $\delta, h > 0$ *tal que, para todo* $(t,x,u) \in T_1((x_0,u_0);\delta)$ *y* $(t,x,v) \in \mathcal{A}$,

$$\mathcal{E}(t,x,u,v) \geq h[(1 + |v-u|^2)^{1/2} - 1].$$

Prueba. Sea $z_0 = (x_0, u_0)$. Por continuidad de L_{uu} existe ϵ_0 y h_0 positivos tal que

$$\langle c, L_{uu}(t,x,u)c\rangle \geq h_0|c|^2$$

para todo $c \in \mathbf{R}^m$, $(t,x,u) \in T_1(z_0;\epsilon_0)$. Sin pérdida de generalidad supongamos que $\epsilon_0 < \epsilon$.

Por el Teorema de Taylor, si (t,x,u) y (t,x,v) están en $T_1(z_0;\epsilon_0)$, tenemos que

$$\begin{aligned}
\mathcal{E}(t,x,u,v) &= L(t,x,v) - L(t,x,u) - \langle v-u, L_u(t,x,u)\rangle \\
&= \int_0^1 (1-\lambda)\langle v-u, L_{uu}(t,x,u+\lambda(v-u))\ (v-u)\rangle d\lambda \\
&\geq \frac{1}{2}h_0|v-u|^2.
\end{aligned}$$

Sea $\delta > 0$ tal que $cl(T_1(z_0;\delta)) \subset T_1(z_0;\epsilon_0)$, y sea $\rho > 0$ tal que $(t,x,u+c) \in T_1(z_0;\epsilon_0)$ para todo $c \in \mathbf{R}^m$ con $|c| \leq \rho$ y $(t,x,u) \in T_1(z_0;\delta)$.

Ahora, tomamos $(t,x,u) \in T_1(z_0;\delta)$ y $(t,x,v) \in \mathcal{A}$, y observamos que $\varphi(c) \leq |c|^2/2$ y $\varphi(c) \leq |c|$. Si $(t,x,v) \in T_1(z_0;\epsilon_0)$, el resultado se cumple con $h = h_0$ ya que

$$\mathcal{E}(t,x,u,v) \geq \frac{1}{2}h_0|v-u|^2 \geq h_0[(1+|v-u|^2)^{1/2} - 1].$$

Si $(t,x,v) \notin T_1(z_0;\epsilon_0)$, sea $k = |v-u|/\rho$ y $c = (v-u)/k$, y observemos que $|c| = \rho$ y $k > 1$. En consecuencia,

$$\mathcal{E}(t,x,u,v) = \mathcal{E}(t,x,u,u+kc)$$

$$
\begin{aligned}
&= \mathcal{E}(t,x,u+c,u+kc) + k\mathcal{E}(t,x,u,u+c) \\
&\qquad +(k-1)\mathcal{E}(t,x,u+c,u) \\
&\geq k\mathcal{E}(t,x,u,u+c) \\
&\geq \frac{1}{2}kh_0|c|^2 \\
&= \frac{1}{2}h_0|c|\,|kc| \\
&\geq \frac{1}{2}h_0\rho[(1+|kc|^2)^{1/2}-1] \\
&= \frac{1}{2}h_0\rho[(1+|v-u|^2)^{1/2}-1].
\end{aligned}
$$

Por lo tanto el resultado se sigue con $h = min\{h_0, h_0\rho/2\}$.

Regresemos a la demostración del Teorema 3.6. Por el Lema 3.2, podemos seleccionar $\delta_0 = min\{\mu, \delta\}$ y $h > 0$, tales que, para todo $(t,x,u) \in T_1(z_0;\delta_0)$ y $(t,x,v) \in \mathcal{A}$,

$$\mathcal{E}(t,x,u,v) \geq h[(1+|v-u|^2)^{1/2}-1].$$

Sea $(x,u) \in Z'(T_0(x_0;\delta_0) \times \mathbf{R}^m) \cap \mathcal{A})$. Entonces

$$\mathcal{E}^*(x,u) = \int_a^b \mathcal{E}(t,x(t),u_0(t),u(t))dt \geq h \int_a^b \varphi(u(t)-u_0(t))dt = hD(u-u_0).$$

Esto prueba la primera relación de (2). Para la segunda, escogemos $\alpha' > 0$ tal que, para todo $(x,u) \in Z'((T_0(x_0;\delta_0) \times \mathbf{R}^m) \cap \mathcal{A})$ y $t \in T$,

$$|M(t,x(t)) + \langle u(t)-u_0(t), N(t,x(t))\rangle| \leq \alpha'|x(t)-x_0(t)|(1+|u(t)-u_0(t)|^2)^{1/2}$$

y sea $\alpha = máx\{\alpha', \alpha'(b-a)\}$. Entonces

$$|K(x)| \leq \alpha'\|x-x_0\| \int_a^b [1+\varphi(u(t)-u_0(t))]dt \leq \alpha\|x-x_0\|[1+D(u-u_0)].$$

Ahora, por la hipótesis (1), para todo $q \in \mathbf{N}$ existe $(x_q, u_q) \in Z'_e(\mathcal{A})$ tal que

$$\|x_q - x_0\| < min\{\delta_0, 1/q\}, \qquad I(x_q,u_q) - I(x_0,u_0) < D(u_q-u_0)/q. \tag{3}$$

Mostremos que $D(u_q - u_0) \to 0$, $q \to \infty$. Puesto que $I'(z_0; w) = 0$ para todo $w \in Y$, y $(x_q - x_0, u_q - u_0) \in Y$, tenemos por (2):

$$
\begin{aligned}
I(x_q,u_q) - I(x_0,u_0) &= K(x_q,u_q) + \mathcal{E}^*(x_q,u_q) \\
&\geq -\alpha\|x_q-x_0\| + D(u_q-u_0)(h-\alpha\|x_q-x_0\|)
\end{aligned}
$$

y entonces, por (3),

$$D(u_q-u_0)[h-(1+\alpha)/q] < \alpha/q \qquad (q \in \mathbf{N})$$

implicando la convergencia deseada.

Este hecho implica que $u_q \to u_0$ en la norma-L^1 ya que, si definimos

$$w_q(t) = \left[1 + \frac{1}{2}\varphi(u_q(t) - u_0(t))\right]^{1/2}$$

entonces, por la desigualdad de Schwarz, y observando que para todo $u \in U'$

$$D(u) = \int_a^b \frac{|u(t)|^2}{2 + \varphi(u(t))} dt,$$

tenemos que

$$\begin{aligned} \left| \int_a^b |u_q(t) - u_0(t)| dt \right|^2 &\leq \int_a^b \frac{|u_q(t) - u_0(t)|^2}{w_q(t)^2} dt \int_a^b w_q(t)^2 dt \\ &= D(u_q - u_0)[2(b-a) + D(u_q - u_0)]. \end{aligned}$$

Consecuentemente, para alguna subsucesión (denotada otra vez por $\{u_q\}$),

$$\lim_{q \to \infty} u_q(t) = u_0(t) \qquad \text{casi uniformemente sobre } T. \tag{4}$$

Por (3), la cantidad $d_q = [2D(u_q - u_0)]^{1/2}$ es estrictamente positiva para todo $q \in \mathbf{N}$. Sea

$$y_q(t) = \frac{x_q(t) - x_0(t)}{d_q}, \qquad v_q(t) = \frac{u_q(t) - u_0(t)}{d_q} \qquad (t \in T,\ q \in \mathbf{N}).$$

Notamos que,

$$\int_a^b \frac{|v_q(t)|^2}{w_q(t)^2} dt = 1 \qquad (q \in \mathbf{N}). \tag{5}$$

Por lo tanto, $\{v_q/w_q\}$ es una sucesión de funciones integrables al cuadrado, y en vista de (5), existe una subsucesión (que no renombramos) y una función $v_0 \in L^2(T; \mathbf{R}^m)$ tal que $\{v_q/w_q\}$ converge débilmente en $L^2(T; \mathbf{R}^m)$ a v_0, esto es,

$$\lim_{q \to \infty} \int_a^b \left\langle g(t), \frac{v_q(t)}{w_q(t)} \right\rangle dt = \int_a^b \langle g(t), v_0(t) \rangle dt \tag{6}$$

para toda $g \in L^2(T; \mathbf{R}^m)$. Sea y_0 la única solución de la ecuación diferencial

$$\dot{y}(t) = A(t)y(t) + B(t)v_0(t) \qquad (t \in T), \qquad y(a) = 0.$$

El Teorema estará probado si mostramos que $I''((x_0, u_0); (y_0, v_0)) \leq 0$, con $(y_0, v_0) \in Y$ y $(y_0, v_0) \not\equiv 0$.

Para probar que éste es el caso, supongamos por el momento que lo siguiente se cumple:

i. $\{v_q\}$ converge débilmente en $L^1(T, \mathbf{R}^m)$ a v_0.

ii. $\{y_q\}$ converge uniformemente en T a y_0.

iii. $\displaystyle \liminf_{q\to\infty} \frac{\mathcal{E}^*(x_q, u_q)}{d_q^2} \geq \frac{1}{2}\int_a^b \langle v_0(t), L_{uu}(t, x_0(t), u_0(t))v_0(t)\rangle dt.$

El hecho de que $(y_0, v_0) \in Y$ es una simple consecuencia de (ii) (claramente, convergencia puntual es suficiente para este resultado). Ahora, por definición del funcional K, para q lo suficientemente grande tenemos que

$$\frac{K(x_q, u_q)}{d_q^2} = \int_a^b \left\{ \frac{M(t, x_q(t))}{d_q^2} + \left\langle \frac{N(t, x_q(t))}{d_q}, v_q(t) \right\rangle \right\} dt.$$

En vista de (ii),

$$\lim_{q\to\infty} \frac{M(t, x_q(t))}{d_q^2} = \frac{1}{2}\langle y_0(t), L_{xx}(t, x_0(t), u_0(t))y_0(t)\rangle,$$

$$\lim_{q\to\infty} \frac{N(t, x_q(t))}{d_q} = L_{ux}(t, x_0(t), u_0(t))y_0(t)$$

ambos uniformemente sobre T. Este hecho, junto con (i), implica que

$$\begin{aligned} &\frac{1}{2}I''((x_0, u_0); (y_0, v_0)) \\ &= \lim_{q\to\infty} \frac{K(x_q, u_q)}{d_q^2} + \frac{1}{2}\int_a^b \langle v_0(t), L_{uu}(t, x_0(t), u_0(t))v_0(t)\rangle dt \end{aligned}$$

y entonces, por (iii), tenemos

$$\begin{aligned} \frac{1}{2}I''((x_0, u_0); (y_0, v_0)) &\leq \lim_{q\to\infty} \frac{K(x_q, u_q)}{d_q^2} + \liminf_{q\to\infty} \frac{\mathcal{E}^*(x_q, u_q)}{d_q^2} \\ &= \liminf_{q\to\infty} \frac{I(x_q, u_q) - I(x_0, u_0)}{d_q^2} \\ &\leq 0. \end{aligned}$$

Finalmente, si $(y_0, v_0) \equiv 0$, entonces $\lim_{q\to\infty} K(x_q, u_q)/d_q^2 = 0$ y entonces, por (2),

$$\frac{1}{2}h \leq \liminf_{q\to\infty} \frac{\mathcal{E}^*(x_q, u_q)}{d_q^2} \leq 0$$

contradiciendo la positividad de h.

Nos resta probar $(i)-(iii)$. Empezamos notando que, para todo $q \in \mathbf{N}$, tenemos

$$\int_a^b [w_q(t)-1]^2 dt = \int_a^b [w_q(t)^2 - 1]dt - 2\int_a^b [w_q(t)-1]dt$$

y, puesto que $w_q(t)^2 \geq w_q(t) \geq 1$ a.e., también

$$0 \leq \int_a^b [w_q(t)-1]dt \leq \int_a^b [w_q(t)^2-1]dt \leq D(u_q - u_0).$$

Por lo tanto,

$$\lim_{q\to\infty} \int_a^b [w_q(t)^2-1]dt = \lim_{q\to\infty} \int_a^b [w_q(t)-1]dt = \lim_{q\to\infty} \int_a^b [w_q(t)-1]^2 dt = 0. \quad (7)$$

Para probar (i), sea g cualquier función en $L^\infty(T;\mathbf{R}^m)$ y notemos que, para todo $q \in \mathbf{N}$,

$$\left| \int_a^b \left\langle g(t)[w_q(t)-1], \frac{v_q(t)}{w_q(t)} \right\rangle dt \right|^2 \leq \int_a^b |g(t)|^2 [w_q(t)-1]^2 dt$$

y

$$\langle g(t), v_q(t)\rangle = \left\langle g(t), \frac{v_q(t)}{w_q(t)} \right\rangle + \left\langle g(t)[w_q(t)-1], \frac{v_q(t)}{w_q(t)} \right\rangle.$$

Usando (6), (7), y que g es acotada, tenemos

$$\lim_{q\to\infty} \int_a^b \langle g(t), v_q(t)\rangle dt = \lim_{q\to\infty} \int_a^b \left\langle g(t), \frac{v_q(t)}{w_q(t)} \right\rangle dt = \int_a^b \langle g(t), v_0(t)\rangle dt, \quad (8)$$

y así tenemos la convergencia débil. Observamos que esta última se cumple también para toda g en $L^2(T;\mathbf{R}^m)$ cuando el rango de integración es reemplazado por cualquier conjunto medible sobre el cual $w_q(t)$ tiende a 1 uniformemente.

Probemos (ii). Para este fin, observamos primero que, para todo $q \in \mathbf{N}\cup\{0\}$,

$$y_q(t) = \Phi(t)\int_a^t \Phi^{-1}(s)B(s)v_q(s)ds \qquad (t \in T)$$

donde Φ es cualquier solución fundamental del sistema homogéneo $\dot{y}(t) = A(t)y(t)$. La convergencia débil de v_q a v_0 en $L^1(T;\mathbf{R}^m)$ implica entonces la convergencia puntual de y_q a y_0. Para probar que la convergencia es uniforme, sea S cualquier subconjunto medible de T. Por la desigualdad de Schwarz y (5),

$$\left|\int_S v_q(t)dt\right|^2 \leq \int_S \frac{|v_q(t)|^2}{w_q(t)^2}dt \int_S w_q(t)^2 dt \leq m(S) + \int_S [w_q(t)^2-1]dt$$

donde $m(S)$ denota la medida de Lebesgue de S. Dado $\epsilon > 0$, sea $q_0 \in$ **N** tal que

$$q \geq q_0 \Rightarrow \int_a^b [w_q(t)^2 - 1]dt < \frac{\epsilon}{2}$$

y sea $\delta \in (0, \epsilon/2)$ tal que, para todo $q < q_0$,

$$m(S) < \delta \Rightarrow \left| \int_S v_q(t)dt \right| < \epsilon.$$

En consecuencia, si $m(S) < \delta$, entonces $|\int_S v_q(t)dt| < \epsilon$ para todo $q \in$ **N**. Consecuentemente, la sucesión $\{\int_S v_q(t)dt\}$, y en consecuencia también $\{y_q\}$, es equi-absolutamente continua sobre T, y entonces $y_q \to y_0$ uniformemente sobre T.

Finalmente, para probar (iii), observamos que por (4), es suficiente probarlo sobre $S \subset T$ donde $u_q(t) \to u_0(t)$ uniformemente. Por el Teorema de Taylor, existe $q_0 \in$ **N** tal que, para todo $q \geq q_0$ y toda $t \in S$,

$$\frac{1}{d_q^2}\mathcal{E}(t, x_q(t), u_0(t), u_q(t)) = \frac{1}{2}\langle v_q(t), R_q(t)v_q(t)\rangle,$$

donde

$$R_q(t) = 2\int_0^1 (1-\lambda)\{L_{uu}(t, x_q(t), u_0(t) + \lambda[u_q(t) - u_0(t)])\}d\lambda.$$

Sea $R(t) = L_{uu}(t, x_0(t), u_0(t))$ y observemos que, para todo t en S,

$$|\langle v_q(t), (R_q(t) - R(t))v_q(t)\rangle| \leq \|R_q(t) - R(t)\||v_q(t)|^2 \leq M_q \frac{|v_q(t)|^2}{w_q(t)^2}$$

donde

$$M_q = sup\{[\|R_q(t) - R(t)\|^2 w_q(t)^4]^{1/2} \mid t \in S\}.$$

Puesto que $R_q(t) \to R(t)$ y $w_q(t) \to 1$, ambos uniformemente sobre S, se sigue que $M_q \to 0$ cuando $q \to \infty$. Observamos también que

$$\begin{aligned}\langle v_q(t), R(t)v_q(t)\rangle &= \langle v_0(t), R(t)v_0(t)\rangle + 2\langle v_q(t) - v_0(t), R(t)v_0(t)\rangle \\ &\quad + \langle v_q(t) - v_0(t), R(t)[v_q(t) - v_0(t)]\rangle.\end{aligned}$$

Como mencionamos antes, (8) aún se cumple cuando $g(t)$ es reemplazada por $R(t)v_0(t)$ y el rango de integración es reemplazado por S. Consecuentemente,

$$\liminf_{q\to\infty} \int_S \langle v_q(t), R_q(t)v_q(t)\rangle dt = \liminf_{q\to\infty} \int_S \langle v_q(t), R(t)v_q(t)\rangle dt$$

$$= \int_S \langle v_0(t), R(t)v_0(t)\rangle dt + \liminf_{q\to\infty} \int_S \langle v_q(t) - v_0(t), R(t)[v_q(t) - v_0(t)]\rangle dt$$

y ya que $R(t)$ es definida positiva, (iii) se cumple. Esto completa la prueba.

En [8] se da una prueba indirecta, en donde se supone que las dinámicas son no lineales y el conjunto $\mathcal{A}$ es arbitrario.

2 Ejemplos.

Un mínimo fuerte.

Consideremos el problema de minimizar el funcional

$$I(x,u) = \int_0^1 \{u(t)^2 + x(t)^3 + x(t)^2 u(t) c(t)\} dt$$

sobre todos los procesos (x, u) que satisfacen

$$\begin{aligned} &x(0) = x(1) = 0 \\ &\dot{x}(t) = a(t)x(t) + b(t)u(t) \qquad (t \in [0,1]) \\ &(x(t), u(t)) \in (-1,1) \times (-1,1) \qquad (t \in [0,1]) \end{aligned}$$

donde a, b, c son funciones continuas que mapean $[0,1]$ a $\mathbf{R}$.

Sea $(x_0, u_0) = (0,0)$. Por cálculos directos uno obtiene que, para todo (y, v) en Y,

$$I'((x_0,u_0);(y,v)) = 0, \qquad I''((x_0,u_0);(y,v)) = \int_0^1 2v(t)^2 dt$$

y entonces $(x_0, u_0) \in E \cap H'$. También, las condiciones modificadas de Legendre y Weierstrass se cumplen ya que $L_{uu}(t, x_0(t), u_0(t)) = 2$ y $\mathcal{E}(t,x,u,v) = (v-u)^2$. Aplicando el Teorema 3.6 concluimos que (x_0, u_0) es un mínimo estricto fuerte de este problema.

Un mínimo débil el cual no es un mínimo fuerte.

Consideremos el problema de minimizar el funcional

$$I(x,u) = \int_0^1 \{u(t)^2 - 4x(t)u(t)^3 + 2tu(t)^4\} dt$$

sobre todos los procesos (x, u) que satisfacen

$$\begin{aligned} &x(0) = x(1) = 0 \\ &\dot{x}(t) = u(t) \qquad (t \in [0,1]) \\ &(x(t), u(t)) \in (-1,1) \times (-\infty,\infty) \qquad (t \in [0,1]). \end{aligned}$$

Sea $(x_0, u_0) = (0,0)$ y observemos que, como en el ejemplo anterior, tenemos que para todo (y, v) en Y,

$$I'((x_0,u_0);(y,v)) = 0, \qquad I''((x_0,u_0);(y,v)) = \int_0^1 2v(t)^2 dt$$

y $L_{uu}(t, x_0(t), u_0(t)) = 2$ implicando que $(x_0, u_0) \in E \cap H' \cap L'$. Por el Teorema 3.6, (x_0, u_0) es un mínimo estricto débil de este problema.

La condición de Weierstrass es también satisfecha por este proceso ya que

$$\mathcal{E}(t, x_0(t), u_0(t), v) = v^2 + 2tv^4 \geq 0$$

para todo $v \in (-\infty, \infty)$ y $t \in [0, 1]$. Sin embargo, (x_0, u_0) no satisface la correspondiente condición modificada y, de hecho, éste no es un mínimo fuerte. Para probarlo, consideremos la familia de controles

$$u(t; \epsilon, \delta) = \begin{cases} \epsilon/\delta & \text{si } 0 \leq t \leq \delta \\ -\epsilon/(1-\delta) & \text{si } \delta < t \leq 1 \end{cases}$$

y las correspondientes trayectorias que satisfacen $\dot{x}(t) = u(t; \epsilon, \delta)$ y $x(0) = 0$, esto es,

$$x(t; \epsilon, \delta) = \begin{cases} \epsilon t/\delta & \text{si } 0 \leq t \leq \delta \\ \epsilon(1-t)/(1-\delta) & \text{si } \delta \leq t \leq 1. \end{cases}$$

Estas trayectorias satisfacen las restricciones de punto final. También,

$$\begin{aligned} I(x(\cdot; \epsilon, \delta), u(\cdot; \epsilon, \delta)) &= \int_0^\delta \left[\frac{\epsilon^2}{\delta^2} - \frac{2\epsilon^4}{\delta^4} t\right] dt \\ &\quad + \int_\delta^1 \left[\frac{\epsilon^2}{(1-\delta)^2} + \frac{\epsilon^4}{(1-\delta)^4}(4 - 2t)\right] dt \\ &= -\frac{\epsilon^4}{\delta^2} + \frac{\epsilon^2}{\delta} + \frac{\epsilon^2}{1-\delta} + \frac{\epsilon^4}{(1-\delta)^3}(3 - \delta). \end{aligned}$$

Tenemos que $|x(t; \epsilon, \delta)| \leq \epsilon$, $t \in [0, 1]$, y para cada ϵ fijo, $I(x(\cdot; \epsilon, \delta), u(\cdot; \epsilon, \delta)) < 0$ siempre y cuando δ sea suficientemente pequeño. Concluimos que (x_0, u_0) no proporciona un mínimo fuerte a este problema.

3 La Condición modificada de Jacobi.

Supongamos que $(x, u) \in L' \cap Z(\mathcal{A})$ con u continua, $L \in C^2(\mathcal{A}; x, u)$. Entonces las matrices α, β, γ de (JE) son continuas y el Teorema sobre la existencia y unicidad de la solución del problema de Cauchy se cumple para este sistema. Denotemos por $(Y(\cdot, a), Q(\cdot, a))$ la solución (matricial) fundamental de (JE), i.e., la solución (matricial) que satisface las condiciones iniciales

$$Y(a, a) = 0, \qquad Q(a, a) = I.$$

Observamos que un punto $s \in (a, b]$ es conjugado a a sobre (x, u) si $Y(s, a)$ es singular, y si (†) se cumple el recíproco tambin se verifica.

Ahora caracterizamos los elementos de H' en términos de la *condición modificada de Jacobi.*

$$J' = \{(x, u) \in Z \mid \ s \in (a, b] \Rightarrow s \text{ no es conjugado a } a \text{ sobre } (x, u)\}.$$

Teorema 3.7 *Sea $L \in C^2(\mathcal{A}; x, u)$, $(x, u) \in L' \cap Z(\mathcal{A})$, u continua, y consideremos las siguientes afirmaciones:*

a. $(x, u) \in H'$.
b. $(x, u) \in J'$.
c. *existe una solución matricial (Y, Q) de (JE) que satisface $|Y(t)| \neq 0$ y $Y(t)^T Q(t) = Q(t)^T Y(t)$ para todo $t \in T$.*

Entonces $(b) \Rightarrow (c) \Rightarrow (a)$ y si $(\dagger)$ se cumple, las tres afirmaciones son equivalentes.

Prueba. $(a) \Rightarrow (b)$: Supongamos que $(\dagger)$ se cumple. Por el Teorema 3.3, $(x, u) \in J$. Entonces, si (b) fuera falso, b sería conjugado a a sobre (x, u), implicando la existencia de una pareja no nula $(z, w) \in Y$ con $J_{(x,u)}(z, w) = 0$. Esto contradice (a).

$(b) \Rightarrow (c)$: Por (b), $Y(t, a)$ es no singular sobre $(a, b]$. Por el Teorema sobre la dependencia continua de las soluciones de ecuaciones diferenciales sobre las condiciones iniciales, existe $\delta > 0$ tal que $Y(t, a - \delta)$ es no singular sobre el intervalo completo $[a, b]$. La segunda conclusión se sigue del hecho de que la cantidad

$$Y^T(t, a - \delta) Q(t, a - \delta) - Q^T(t, a - \delta) Y(t, a - \delta)$$

tiene derivada cero y valor cero en $t = a - \delta$.

$(c) \Rightarrow (a)$: Sea $(y, v) \in D$ con $y(a) = y(b) = 0$ y sea $w(t) = Y^{-1}(t) y(t)$ y $z(t) = v(t) - V(t) w(t)$ donde

$$V(t) = L_{uu}^{-1}(\tilde{x}(t)) B(t)^T Q(t) - L_{uu}^{-1}(\tilde{x}(t)) \; L_{ux}(\tilde{x}(t)) Y(t).$$

Queremos mostrar que $I''((x, u); (y, v)) > 0$, $(y, v) \not\equiv 0$.

Observamos primero que $B(t)^T Q(t) = L_{ux}(\tilde{x}(t)) Y(t) \; + L_{uu}(\tilde{x}(t)) V(t)$ y, puesto que

$$\dot{Y}(t) = \alpha(t) Y(t) + \beta(t) Q(t)$$
$$\dot{Q}(t) = \gamma(t) Y(t) - \alpha(t)^T Q(t).$$

tenemos que

$$\begin{aligned}
&\dot{Y}(t) = A(t) Y(t) + B(t) V(t) \\
&\dot{Q}(t) + A(t)^T Q(t) = L_{xx}(\tilde{x}(t)) Y(t) + L_{xu}(\tilde{x}(t)) V(t).
\end{aligned}$$

Por lo tanto,

$$\begin{aligned}
\Omega_y(t, y(t), v(t)) &= (\dot{Q}(t) + A(t)^T Q(t)) w(t) + L_{xu}(\tilde{x}(t)) z(t) \\
\Omega_v(t, y(t), v(t)) &= B(t)^T Q(t) w(t) + L_{uu}(\tilde{x}(t)) z(t).
\end{aligned}$$

En consecuencia,

$$\begin{aligned}
2\Omega(t,y(t),v(t)) &= \langle y(t),\Omega_y(t,y(t),v(t))\rangle + \langle v(t),\Omega_v(t,y(t),v(t))\rangle \\
&= \langle Y(t)w(t),(\dot{Q}(t)+A(t)^TQ(t))w(t) + L_{xu}(\tilde{x}(t))z(t)\rangle \\
&\quad + \langle V(t)w(t)+z(t),B(t)^TQ(t)w(t) + L_{uu}(\tilde{x}(t))z(t)\rangle \\
&= \langle w(t),Y^T(t)\dot{Q}(t)w(t)\rangle \\
&\quad + \langle w(t),Q^T(t)(A(t)Y(t) + B(t)V(t))w(t)\rangle \\
&\quad + \langle L_{ux}(\tilde{x}(t))Y(t)w(t) + L_{uu}(\tilde{x}(t))V(t)w(t),z(t)\rangle \\
&\quad + \langle B(t)^TQ(t)w(t),z(t)\rangle + \langle z(t),L_{uu}(\tilde{x}(t))z(t)\rangle \\
&= \langle w(t),Y(t)^T\dot{Q}(t)w(t)\rangle + \langle w(t),Q(t)^T\dot{Y}(t)w(t)\rangle \\
&\qquad +2\langle B(t)^TQ(t)w(t),z(t)\rangle + \langle z(t),L_{uu}(\tilde{x}(t))z(t)\rangle \\
&= \langle w(t),Y(t)^T\dot{Q}(t)w(t)\rangle + \langle w(t),\dot{Y}(t)^TQ(t)w(t)\rangle \\
&\qquad +2\langle B(t)^TQ(t)w(t),z(t)\rangle + \langle z(t),L_{uu}(\tilde{x}(t))z(t)\rangle \\
&= \langle w(t),(Y(t)^T\dot{Q}(t)+\dot{Y}(t)^TQ(t))w(t)\rangle \\
&\quad + 2\langle B(t)^TQ(t)w(t),z(t)\rangle + \langle z(t),L_{uu}(\tilde{x}(t))z(t)\rangle.
\end{aligned}$$

Ahora, observamos que

$$\begin{aligned}
\langle B(t)^TQ(t)w(t),z(t)\rangle &= \langle Q(t)w(t),B(t)v(t)-B(t)V(t)w(t)\rangle \\
&= \langle Q(t)w(t),Y(t)\dot{w}(t)\rangle
\end{aligned}$$

y entonces, puesto que $Y^T(t)Q(t)=Q^T(t)Y(t)$,

$$2\Omega(t,y(t),v(t)) = \frac{d}{dt}\,\langle w(t),Y^T(t)Q(t)w(t)\rangle + \langle z(t),L_{uu}(\tilde{x}(t))z(t)\rangle.$$

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

$$I''((x,u);(y,v)) = \int_a^b \langle z(t),L_{uu}(\tilde{x}(t))z(t)\rangle dt.$$

Consecuentemente, $I''((x,u);(y,v))>0$ excepto cuando $v(t)=V(t)w(t)$. Se sigue que $I''((x,u);(y,v))=0 \Rightarrow B(t)v(t)=B(t)V(t)w(t)$, esto implica que $\dot{y}(t)=\dot{Y}(t)Y^{-1}(t)y(t)$. Puesto que $y(a)=0$, esto implica que $y\equiv 0$. Por lo tanto, $w\equiv 0$ y entonces $v\equiv 0$.

Condiciones suficientes para un mínimo se siguen de este resultado y el Teorema 3.6.

Teorema 3.8 *Si $L\in C^2(\mathcal{A};x,u)$, $(x,u)\in Z_e(\mathcal{A})$, y u es continua, entonces*

a. $(x,u)\in E\cap J'\cap L' \Rightarrow$ (x,u) *es un mínimo estricto débil para* $P(\mathcal{A})$.
b. $(x,u)\in E\cap J'\cap L'\cap W(\mathcal{A};\epsilon)$ *para algún* $\epsilon>0$,
$\Rightarrow (x,u)$ *es un mínimo estricto fuerte para* $P(\mathcal{A})$.

4 Campos de extremos.

En esta sección se probarán condiciones suficientes para un mínimo local débil y un mínimo local fuerte por medio del concepto de un campo de Mayer el cual proporciona una de las pruebas clásicas de suficiencia en el cálculo de variaciones. Aquí este concepto es extendido para el problema de control óptimo $P(\mathcal{A})$, y como en la prueba usual, se prueba que la condición de Jacobi implica la inclusión de un extremo no singular en un campo de Mayer.

Empezamos con un Lema el cual extiende un resultado fundamental en el cálculo de variaciones. Este afirma que una trayectoria no singular (en este contexto, un proceso no singular, i.e., uno para el cual $|L_{uu}(t, x(t), u(t))|$ no se anula) satisface la ecuación de Euler si y sólo si ésta satisface en términos de las *variables canónicas* y la *transformada de Legendre*, una cierta ecuación diferencial de primer orden.

Lema 3.3 *Sea $L \in C^2(\mathcal{A}; x, u)$, $(x_0, u_0) \in Z(\mathcal{A})$ no singular, u_0 continua, y $(x_0, u_0, p_0) \in R$. Definimos*

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

entonces existen $\epsilon, \delta > 0$ y una función $\mathcal{L} : T_2((x_0, p_0); \epsilon) \to \mathbf{R}$ (Transformada de Legendre), tal que, para cualquier $x, p \in X$ que satisface

$$(t, x(t), p(t)) \in T_2((x_0, p_0); \epsilon) \qquad (t \in T),$$

$(b) \Rightarrow (a)$ y, si $(\dagger)$ se cumple, entonces $(a) \Rightarrow (b)$, donde

a. *Existe una única función continua u que mapea T a $\mathbf{R}^m$,*
tal que $(x, u) \in E \cap D$, $(x, u, p) \in R$, y $(t, x(t), u(t)) \in T_1((x_0, u_0); \delta)$ $(t \in T)$.
b. *(x, p) satisface para todo $t \in T$ el sistema el cual denotamos por (LE):*

$$\begin{aligned} \dot{x}(t) &= \mathcal{L}_p(t, x(t), p(t)) \\ \dot{p}(t) &= -\mathcal{L}_x(t, x(t), p(t)). \end{aligned}$$

Prueba. Para todo $(t, x, v) \in \mathcal{A}$ y $p \in \mathbf{R}^n$ sea

$$G(t, x, p, v) = B(t)^T p - L_u(t, x, v).$$

Puesto que $(x_0, u_0, p_0) \in R$, $G(t, x_0(t), p_0(t), u_0(t)) = 0$ $(t \in T)$ y, como (x_0, u_0) es no singular, $|G_v(t, x_0(t), p_0(t), u_0(t))| \neq 0$ $(t \in T)$. También, puesto que L_u y L_{uu} son continuas, también lo son G y G_v. Por el Teorema de la función implícita existe $\lambda, \rho > 0$ y una función continua U que mapea $M = T_2((x_0, p_0); \rho)$ a $\mathbf{R}^m$ tal que

$$U(t, x_0(t), p_0(t)) = u_0(t) \qquad (t \in T),$$

$$B(t)^T p = L_u(t, x, U(t, x, p)) \qquad ((t, x, p) \in M),$$

y las relaciones

$$B(t)^T p = L_u(t,x,v), \qquad |v - U(t,x,p)| < \lambda, \qquad (t,x,p) \in M$$

se cumplen sólo si $v = U(t,x,p)$. También $G \in C^r \Rightarrow U \in C^r$.

Para todo $(t,x,p) \in M$ definimos

$$\mathcal{L}(t,x,p) = \langle p, A(t)x + B(t)U(t,x,p)\rangle - L(t,x,U(t,x,p)).$$

Uno fácilmente verifica que, para todo $(t,x,p) \in M$,

$$\begin{aligned} \mathcal{L}_x(t,x,p) &= A(t)^T p - L_x(t,x,U(t,x,p)) \\ \mathcal{L}_p(t,x,p) &= A(t)x + B(t)U(t,x,p). \end{aligned}$$

Ahora, sea $\delta = \lambda/2$ y sea $\eta > 0$ tal que, para todo $(t,x,p) \in T_2((x_0,p_0);\eta)$,

$$|U(t,x_0(t),p_0(t)) - U(t,x,p)| < \delta.$$

Sea $\epsilon = min\{\rho,\delta,\eta\}$ y $N = T_2((x_0,p_0);\epsilon)$.

$(b) \Rightarrow (a)$: Sean $x,p \in X$ con $(t,x(t),p(t)) \in N$. Puesto que (x,p) satisface (LE),

$$\begin{gathered} \dot{x}(t) = A(t)x(t) + B(t)U(t,x(t),p(t)) \\ \dot{p}(t) + A(t)^T p(t) = L_x(t,x(t),U(t,x(t),p(t)). \end{gathered}$$

Puesto que $(t,x(t),p(t)) \in N \subset M$,

$$B(t)^T p(t) = L_u(t,x(t),U(t,x(t),p(t)).$$

En consecuencia, $u(t) = U(t,x(t),p(t)) \Rightarrow u$ es continua, $(x,u) \in D$, $(x,u,p) \in R$, y

$$(t,x(t),u(t)) \in T_1((x_0,u_0);\delta).$$

El hecho de que $(x,u) \in E$ se sigue por el Teorema 3.2. Para probar la unicidad, supongamos que $v : T \to \mathbf{R}^m$ es tal que $(x,v,p) \in R$ y $(t,x(t),v(t)) \in T_1((x_0,u_0);\delta)$. Entonces

$$|v(t) - U(t,x(t),p(t))| \leq |v(t) - u_0(t)| + |u(t) - u_0(t)| < \lambda$$

y entonces $v(t) = u(t)$.

$(a) \Rightarrow (b)$. Supongamos que $(\dagger)$ se cumple y sea (x,u,p) que satisfaga (a) con $(t,x(t),p(t)) \in N$. Puesto que $(x,u) \in E$, $(x,u,p) \in R$, y $(\dagger)$ se cumple, por el Teorema 3.2 tenemos que

$$\dot{p}(t) + A(t)^T p(t) = L_x(t,x(t),u(t)).$$

También, como $(x,u) \in D$,

$$\dot{x}(t) = A(t)x(t) + B(t)u(t).$$

El resultado se obtendrá si mostramos que $u(t) = U(t,x(t),p(t))$, pero esto es una consecuencia de

$$|u(t) - U(t,x(t),p(t))| \leq |u(t) - u_0(t)| + |u_0(t) - U(t,x(t),p(t))| < \lambda.$$

Definición 3.3 *Dado un proceso* $(x,u) \in Z_e(\mathcal{A})$ *diremos que una pareja* (Γ, M) *es un campo de Mayer sobre* $\mathcal{A}$ *a lo largo de* (x,u) *si*

i. $M = T_0(x;\epsilon)$ *para algún* $\epsilon > 0$.
ii. $\Gamma : M \to \mathbf{R}^m$ *es continua,* $(t,y,\Gamma(t,y)) \in \mathcal{A}$ *para todo* (t,y) *en* M*, y*

$$u(t) = \Gamma(t, x(t)) \qquad \textit{para todo } t \in T$$

iii. Para cualquier (y,v) *en* $Z_e(\mathcal{A})$ *con* $(t,y(t)) \in M$ *para todo* $t \in T$, $I^*(y,v) = I^*(x,u)$, *donde*

$$I^*(y,v) = \int_a^b \{L(t,y(t),\Gamma(t,y(t))) \ + \langle v(t) - \Gamma(t,y(t)), L_u(t,y(t),\Gamma(t,y(t)))\rangle\}dt.$$

Teorema 3.9 *Sea* $L \in C^2(\mathcal{A}; x, u)$, $(x_0,u_0) \in Z_e(\mathcal{A}) \cap E \cap J' \cap L'$, $(x_0,u_0,p_0) \in R$*,* u_0 *continua, y supongamos que* $(\dagger)$ *se cumple. Entonces existe un campo de Mayer sobre* $\mathcal{A}$ *a lo largo de* (x_0,u_0).

Prueba. Por el lema 3.3, (x_0,p_0) satisface (LE) y entonces éste puede ser extendido sobre un intervalo mayor $[a-\rho, b+\rho]$. Para $0 < \delta < \rho$, resolvemos el problema de Cauchy para (LE) con condición inicial

$$x(a-\delta,\lambda) = x_0(a-\delta), \qquad p(a-\delta,\lambda) \ = \lambda + p_0(a-\delta).$$

Por la unicidad del teorema, $x(t,0) = x_0(t)$ y $p(t,0) = p_0(t)$.

Por el lema 3.3, si $(t, x(t,\lambda), p(t,\lambda)) \in T_2((x_0,p_0);\epsilon)$, existe una única función continua $u(\cdot,\lambda) : [a-\rho, b+\rho] \to \mathbf{R}^m$ tal que

$$(x(\cdot,\lambda), u(\cdot,\lambda)) \in E \cap D, \qquad (x(\cdot,\lambda), u(\cdot,\lambda), p(\cdot,\lambda)) \in R.$$

Por la unicidad de $u(\cdot,\lambda)$, tenemos también que $u(t,0) = u_0(t)$. Por lo tanto,

$$\dot{x}(t,\lambda) = A(t)x(t,\lambda) + B(t)u(t,\lambda),$$

$$B(t)^T p(t,\lambda) = L_u(t, x(t,\lambda), u(t,\lambda))$$

y, puesto que $(\dagger)$ se cumple, se sigue que

$$\dot{p}(t,\lambda) + A(t)^T p(t,\lambda) = L_x(t, x(t,\lambda), u(t,\lambda)).$$

Derivando con respecto a λ y evaluando en $\lambda = 0$ obtenemos

$$\dot{Y}(t,a-\delta) = A(t)Y(t,a-\delta) + B(t)V(t,a-\delta),$$

$$\begin{aligned}\dot{Q}(t,a-\delta) + A(t)^T Q(t,a-\delta) &= L_{xx}(t,x_0(t),u_0(t))\ Y(t,a-\delta) \\ &\quad + L_{xu}(t,x_0(t),u_0(t))V(t,a-\delta),\end{aligned}$$

$$B(t)^T Q(t,a-\delta) = L_{ux}(t,x_0(t),u_0(t))Y(t,a-\delta)$$

$$+L_{uu}(t, x_0(t), u_0(t))V(t, a-\delta),$$

donde

$Y(t,a-\delta) = \partial x(t,0)/\partial\lambda, \quad V(t,a-\delta) = \partial u(t,0)/\partial\lambda, \quad Q(t,a-\delta) = \partial p(t,0)/\partial\lambda.$

Por la Proposición 3.9, $(Y(\cdot, a-\delta), Q(\cdot, a-\delta))$ satisface el sistema matricial (JE) y, por las condiciones iniciales impuestas sobre $x(\cdot,\lambda)$ y $p(\cdot,\lambda)$,

$$Y(a-\delta, a-\delta) = 0, \qquad Q(a-\delta, a-\delta) = I.$$

Esto implica que $(Y(\cdot, a-\delta), Q(\cdot, a-\delta))$ es la solución matricial fundamental de (JE). De la dependencia continua de las soluciones sobre la condición inicial, la matriz $Y(t, a-\delta)$ es no singular para alguna $\delta > 0$ sobre el intervalo completo $[a, b]$.

En consecuencia, $|x_\lambda(t,0)| \neq 0$ sobre T y, por el teorema de la función implícita, existe $\mu, \sigma > 0$ y $\lambda : T_0(x_0;\mu) \to \mathbf{R}^n$, tal que

$$\lambda(t, x_0(t)) = 0, \qquad x(t, \lambda(t,y)) = y \qquad \text{para todo } (t,y) \in T_0(x_0;\mu),$$

y las relaciones

$$x(t,\nu) = y, \qquad |\nu - \lambda(t,y)| < \sigma, \qquad (t,y) \in T_0(x_0;\mu)$$

se cumplen sólo si $\nu = \lambda(t,y)$. También, $x(t,\lambda) \in C^r \Rightarrow \lambda \in C^r$.

Sea $M = T_0(x_0;\mu)$ y $\Gamma(t,y) = u(t, \lambda(t,y))$ para todo (t,y) en M. Entonces Γ es continua y, si es necesario, podemos disminuir a μ de manera que $(t, y, \Gamma(t,y)) \in \mathcal{A}$ para todo (t,y) en M. Por construcción, $u_0(t) = \Gamma(t, x_0(t))$ $(t \in T)$.

Ahora, sea (y, v) en $Z_e(\mathcal{A})$ con $(t, y(t)) \in M$ para todo $t \in T$. Para todo $s \in T$ definimos

$$w(t,s) = x(t, \lambda(s, y(s))), \qquad z(t,s) = u(t, \lambda(s, y(s))) \qquad (t \in [a_0, s])$$

y

$$F(s) = \int_{a_0}^{s} L(t, w(t,s), z(t,s))dt$$

donde $a_0 = a - \delta$. Puesto que

$$\begin{aligned}
&\dot{w}(t,s) = A(t)w(t,s) + B(t)z(t,s),\\
&\dot{p}(t, \lambda(s,y(s))) + A(t)^T p(t, \lambda(s, y(s))) = L_x(t, w(t,s), z(t,s)),\\
&B(t)^T p(t, \lambda(s, y(s))) = L_u(t, w(t,s), z(t,s)),
\end{aligned}$$

y $w_s(a_0, s) = 0$, tenemos que

$$\begin{aligned}
F'(s) &= \int_{a_0}^{s} \{\langle L_x(t, w(t,s), z(t,s)), w_s(t,s)\rangle \\
&\qquad + \langle L_u(t, w(t,s), z(t,s)), z_s(t,s)\rangle\}dt + L(s, w(s,s), z(s,s)) \\
&= \int_{a_0}^{s} \frac{d}{dt}\langle p(t, \lambda(s, y(s))), w_s(t,s)\rangle dt + L(s, y(s), \Gamma(s, y(s))) \\
&= \langle p(s, \lambda(s, y(s))), w_s(s,s)\rangle + L(s, y(s), \Gamma(s, y(s))).
\end{aligned}$$

Ahora, en vista de las igualdades

$$w_s(s,s) = \dot{y}(s) - \dot{x}(s,\lambda(s,y(s))) = B(s)(v(s) - \Gamma(s,y(s))).$$

se sigue que

$$F'(s) = L(s,y(s),\Gamma(s,y(s))) \ + \langle v(s) - \Gamma(s,y(s)), L_u(s,y(s),\Gamma(s,y(s)))\rangle.$$

En consecuencia,

$$I^*(y,v) = \int_a^b F'(s)ds = F(b) - F(a) = I^*(x_0,u_0).$$

Concluimos que (Γ, M) es un campo de Mayer sobre $\mathcal{A}$ a lo largo de (x_0,u_0).

Teorema 3.10 *Si $L \in C^2(\mathcal{A};x,u)$, $(x,u) \in Z_e(\mathcal{A})$, u continua, (†) se cumple y, para algún $p \in X$, $(x,u,p) \in R$, entonces*

a. $(x,u) \in E \cap J' \cap L' \Rightarrow (x,u)$ *es un mínimo estricto débil para* $P(\mathcal{A})$.
b. $(x,u) \in E \cap J' \cap L' \cap W(\mathcal{A};\epsilon)$ *para algún* $\epsilon > 0$,
$\Rightarrow (x,u)$ *es un mínimo estricto fuerte para* $P(\mathcal{A})$.

Prueba. Por el Teorema 3.6 es suficiente con probar (b). Por el Teorema 3.9, existe un campo de Mayer (Γ, M) sobre $\mathcal{A}$ a lo largo de (x,u). Sea $\mu > 0$ tal que $M = T_0(x;\mu)$ y, si es necesario, disminuimos ϵ tal que

$$\mathcal{E}(t,y,z,v) > 0 \text{ para todo } (t,y,z) \in T_1((x,u);\epsilon), \quad (t,y,v) \in \mathcal{A}, \ v \neq z.$$

El hecho de que esto se puede hacer se probó en la Proposición 3.8. También, μ puede ser disminuido de manera que $(t,y,\Gamma(t,y)) \in T_1((x,u);\epsilon)$ para todo $(t,y) \in M$.

Ahora, sea $(y,v) \in Z_e(\mathcal{A})$ con $(t,y(t)) \in M$ para todo $t \in T$. Puesto que

$$I^*(y,v) = I^*(x,u) = I(x,u)$$

y

$$I(y,v) = I^*(y,v) + \int_a^b \mathcal{E}(t,y(t),\Gamma(t,y(t)),v(t))dt$$

se sigue que $I(y,v) \geq I(x,u)$, cupliéndose la igualdad sólo si $v(t) = \Gamma(t,y(t))$ $(t \in T)$. En este evento tenemos que x y y satisfacen la ecuación $\dot{x}(t) = F(t,x(t))$ $(t \in T)$ con la misma condición $x(a) = \xi_0$, donde

$$F(t,w) = A(t)w + B(t)\Gamma(t,w).$$

En consecuencia, x y y coinciden y entonces $u(t) = \Gamma(t,y(t)) = v(t)$. Esto muestra que (x,u) es un mínimo estricto fuerte para $P(\mathcal{A})$.

5 La Ecuación de Hamilton Jacobi.

Suficiencia a través de la teoría de Hamilton Jacobi está basada en la existencia de una función que satisface alguna desigualdad o alguna ecuación diferencial parcial. En ésta y en la siguiente sección se mostrará cómo, para ambos casos, la existencia de estas funciones de " verificación " es implicada por las condiciones de suficiencia para un mínimo dadas en el Teorema 3.10.

La primera demostración está basada en la ecuación diferencial de Hamilton Jacobi, y aunque se procede directamente, ésta es equivalente a la de la sección anterior.

Para todo $(t, x, p) \in T \times \mathbf{R}^{2n}$, sea

$$\mathcal{F}(t, x, p) = sup\{\langle p, A(t)x + B(t)u\rangle - L(t, x, u) \mid (t, x, u) \in \mathcal{A}\}$$

llamada la *transformada de Young − Fenchel* (relativa a $\mathcal{A}$). El teorema de verificación que consideraremos en esta sección afirma lo siguiente:

Lema 3.4 *Supongamos que $(x, u) \in Z_e(\mathcal{A})$ y que existe $\epsilon > 0$ y una función C^1 $W : T_0(x;\epsilon) \to \mathbf{R}$ tal que:*

a. $W_t(t, y) + \mathcal{F}(t, y, W_y(t, y)) = 0$ *para todo (t, y) en $T_0(x;\epsilon)$ (ecuación diferencial parcial de Hamilton Jacobi).*
b. $\mathcal{F}(t, x(t), W_y(t, x(t))) = \langle W_y(t, x(t)), \dot{x}(t)\rangle - L(t, x(t), u(t))$ $(t \in T)$.
Entonces (x, u) es un mínimo fuerte de $P(\mathcal{A})$.

Prueba. Sea $(y, v) \in Z_e(\mathcal{A})$ con $(t, y(t)) \in T_0(x;\epsilon)$ para todo $t \in T$. Por (a) y la definición de $\mathcal{F}$,

$$W_t(t, y(t)) + \langle W_y(t, y(t)), \dot{y}(t)\rangle - L(t, y(t), v(t)) \leq 0$$

y, por (a) y (b),

$$W_t(t, x(t)) + \langle W_y(t, x(t)), \dot{x}(t)\rangle - L(t, x(t), u(t)) = 0.$$

Por lo tanto, $I(y, v) \geq W(b, \xi_1) - W(a, \xi_0) = I(x, u)$.

Prueba del Teorema 3.10: Supongamos que $L \in C^2(\mathcal{A}; x, u)$, (†) se cumple, $(x_0, u_0) \in Z_e(\mathcal{A}) \cap E \cap J' \cap L'$, u_0 continua y, para algún $p_0 \in X$, $(x_0, u_0, p_0) \in R$.

Sea (Γ, M) el campo de Mayer sobre $\mathcal{A}$ a lo largo de (x_0, u_0) construido en la prueba del Teorema 3.9, y definimos

$$W(t, y) = \int_{a_0}^{t} L(s, x(s, \lambda(t, y)), u(s, \lambda(t, y)))ds.$$

Calculamos las derivadas parciales de $W(t, y)$. Para y fija, definimos

$$w(s, t) = x(s, \lambda(t, y)), \qquad z(s, t) = u(s, \lambda(t, y)).$$

Entonces

$$\begin{aligned}
W_t(t,y) &= \int_{a_0}^{t}\{\langle L_x(s,w(s,t),z(s,t)),w_t(s,t)\rangle \\
&\quad +\langle L_u(s,w(s,t),z(s,t)),z_t(s,t)\rangle\}ds \ + L(t,w(t,t),z(t,t)) \\
&= \int_{a_0}^{t}\{\langle p_s(s,\lambda(t,y)) \ + A(s)^T p(s,\lambda(t,y)),w_t(s,t)\rangle \\
&\quad +\langle B(s)^T p(s,\lambda(t,y)),z_t(s,t)\rangle\}ds \ + L(t,y,\Gamma(t,y)) \\
&= \int_{a_0}^{t}\frac{d}{ds}\langle p(s,\lambda(t,y)),w_t(s,t)\rangle ds \ + L(t,y,\Gamma(t,y)) \\
&= \langle p(t,\lambda(t,y)),w_t(t,t)\rangle + L(t,y,\Gamma(t,y)) \\
&= L(t,y,\Gamma(t,y)) - \langle p(t,\lambda(t,y)), A(t)y + B(t)\Gamma(t,y)\ \rangle.
\end{aligned}$$

De manera similar, $W_y(t,y) = p(t,\lambda(t,y))$.

Ahora, la transformada de Legendre está dada por

$$\mathcal{L}(t,x,p) = \langle p, A(t)x + B(t)U(t,x,p)\rangle - L(t,x,U(t,x,p))$$

donde, para algún $\rho,\delta > 0$, la función continua $U : T_2((x_0,p_0);\rho) \to \mathbf{R}^m$ es tal que

$$U(t,x_0(t),p_0(t)) = u_0(t) \qquad (t \in T),$$

$$B(t)^T p = L_u(t,x,U(t,x,p)) \qquad ((t,x,p) \in T_2((x_0,p_0);\rho)),$$

y las relaciones

$$B(t)^T p = L_u(t,x,v), \qquad |v - U(t,x,p)| < \delta, \qquad (t,x,p) \in T_2((x_0,p_0);\rho)$$

se cumplen sólo si $v = U(t,x,p)$.

Observamos que podemos disminuir M tal que, para todo $(t,y) \in M$,

$$\Gamma(t,y) = U(t,y,p(t,\lambda(t,y))).$$

Por lo tanto,

a. $W_t(t,y) + \mathcal{L}(t,y,W_y(t,y)) = 0$ para todo (t,y) en M
b. $\mathcal{L}(t,x_0(t),W_y(t,x_0(t))) = \langle W_y(t,x_0(t)),\dot{x}_0(t)\rangle - L(t,x_0(t),u_0(t))$ $(t \in T)$.

Estas relaciones corresponden precisamente a las del Lema 3.4 con la transformada de Legendre en vez de la transformada de Young-Fenchel. Sin embargo, para todo $(t,x,p) \in T_2((x_0,p_0);\rho)$, $(t,x,u) \in \mathcal{A}$,

$$\mathcal{E}(t,x,U(t,x,p),u) = \mathcal{L}(t,x,p) - \ \langle p, A(t)x + B(t)u\rangle + L(t,x,u)$$

y entonces, si (x_0,u_0) satisface la condición modificada de Weierstrass, ambas transformadas coinciden (localmente). Este hecho, junto con el Lema 3.4, implican el Teorema 3.10.

6 La desigualdad de Hamilton Jacobi.

La última prueba depende de la desigualdad de Hamilton Jacobi. El teorema de verificación para este caso no es equivalente al de la transformada de Young Fenchel sino que es más débil y, a diferencia de las otras pruebas, la suficiencia es derivada sin la ayuda de un campo de extremos.

Consideremos el siguiente Lema de verificación:

Lema 3.5 *Supongamos que $(x, u) \in Z_e(\mathcal{A})$ y que existe $\epsilon > 0$ y una función C^1 $W : T_0(x; \epsilon) \to \mathbf{R}$ tal que*

$$W_t(t, y) + \langle W_y(t, y), A(t)y + B(t)v\rangle - L(t, y, v)$$
$$\leq W_t(t, x(t)) + \langle W_y(t, x(t)), \dot{x}(t)\rangle - L(t, x(t), u(t))$$

para todo (t, y, v) en $(T_0(x; \epsilon) \times \mathbf{R}^m) \cap \mathcal{A}$. Entonces (x, u) es un mínimo fuerte para $P(\mathcal{A})$.

Esta afirmación se verifica fácilmente simplemente integrando la desigualdad anterior, la cual llamamos (DHJ). Observamos que, si W satisface (a) y (b) del Lema 3.4, entonces W satisface (DHJ) pero el recíproco no es necesariamente cierto.

Prueba del Teorema 3.10: Supongamos una vez más, que $L \in C^2(\mathcal{A}; x, u)$, $(\dagger)$ se cumple, para algún $p \in X$, $(x, u, p) \in R$ y, para algún $\eta > 0$,

$$(x, u) \in Z_e(\mathcal{A}) \cap E \cap J' \cap L' \cap W(\mathcal{A}; \eta).$$

Reemplazamos la condición $(x, u) \in E$ por la suposición de que

$$\dot{p}(t) + A(t)^T p(t) = L_x(t, x(t), u(t)) \qquad (t \in T).$$

No es difícil mostrar que (ver[**1**]), para alguna $\mu > 0$, no hay una solución no trivial (y, q) de (JE) con γ reemplazada por $\gamma - \mu I$, para la cual y se anula en ambos, en a y en algún punto s en $(a, b]$. En vista del Teorema 3.7, existe una solución matricial (Y, Q) de el sistema (JE) con γ reemplazada por $\gamma - \mu I$, tal que $|Y(t)| \neq 0$ y $Y(t)^T Q(t) = Q(t)^T Y(t)$ para todo $t \in T$.

Definiendo $V(t) = Q(t)Y^{-1}(t)$ (ver Corolario 4.2) se sigue que V es una solución matricial sobre T de la desigualdad matricial de Riccati:

$$\dot{V}(t) + V(t)\alpha(t) + \alpha(t)^T V(t) + V(t)\beta(t)V(t) - \gamma(t) < 0.$$

Definimos, para todo (t, y) en $T \times \mathbf{R}^n$,

$$W(t, y) = \langle p(t), y\rangle + \frac{1}{2}\langle y - x(t), V(t)(y - x(t))\rangle.$$

Mostraremos que el dominio de W puede ser disminuido de manera que W satisfaga (DHJ).

Observamos que, si definimos

$$G(t,y,v)=B(t)^Tp(t)+B(t)^TV(t)(y-x(t))-L_u(t,y,v)$$

entonces $G(t,x(t),u(t))=0$ y $|G_v(t,x(t),u(t))|\neq 0$ para todo $t\in T$. También G y G_v son continuas. Por el teorema de la función implícita, existe $\rho,\delta>0$ y una función continua $\xi:T_0(x;\delta)\to\mathbf{R}^m$ tal que

$$\xi(t,x(t))=u(t)\qquad (t\in T),$$

$$B(t)^Tp(t)+B(t)^TV(t)(y-x(t))=L_u(t,y,\xi(t,y))\qquad ((t,y)\in T_0(x;\delta)),$$

y las relaciones

$$B(t)^Tp(t)+B(t)^TV(t)(y-x(t))=L_u(t,y,v),\ |v-\xi(t,y)|<\rho,\ \ (t,y)\in T_0(x;\delta)$$

se cumplen sólo si $v=\xi(t,y)$. También $G\in C^r\Rightarrow\xi\in C^r$. Ahora definimos, para todo $(t,y)\in T_0(x;\delta)$,

$$F(t,y)=W_t(t,y)+\ \langle W_y(t,y),A(t)y+B(t)\xi(t,y)\rangle-L(t,y,\xi(t,y)).$$

Supongamos que ya probamos que F_y y F_{yy} son continuas y que para todo $t\in T$ satifacen que $F_y(t,x(t))=0$ y

$$F_{yy}(t,x(t))=\dot V(t)+V(t)\alpha(t)+\alpha(t)^TV(t)+V(t)\beta(t)V(t)\ \ -\gamma(t)<0.$$

En este evento, por la fórmula de Taylor, δ puede ser disminuido de manera que, para todo (t,y) en $T_0(x;\delta)$, $F(t,y)\le F(t,x(t))$. Si es necesario disminuimos nuevamente δ de manera que $(t,y,\xi(t,y))\in T_1((x,u);\eta)$. Observamos que, si $(t,y,v)\in(T_0(x;\delta)\times\mathbf{R}^m)\cap\mathcal{A}$ entonces, puesto que $(x,u)\in W(\mathcal{A};\eta)$ y

$$B(t)^TW_y(t,y)=L_u(t,y,\xi(t,y)),$$

tenemos que

$$\begin{aligned}
&W_t(t,y)+\langle W_y(t,y),A(t)y+B(t)v\rangle-L(t,y,v)\\
&\quad=\ W_t(t,y)+\langle W_y(t,y),A(t)y\rangle\ +\langle L_u(t,y,\xi(t,y)),v\rangle-L(t,y,v)\\
&\quad\le\ W_t(t,y)+\langle W_y(t,y),A(t)y\rangle\ +\langle L_u(t,y,\xi(t,y)),\xi(t,y)\rangle-L(t,y,\xi(t,y))\\
&\quad=\ F(t,y)\le F(t,x(t))\\
&\quad=\ W_t(t,x(t))+\langle W_y(t,x(t)),\dot x(t)\rangle-L(t,x(t),u(t)).
\end{aligned}$$

Una aplicación del Lema 3.5 produciría el resultado requerido.

Probemos la afirmación. Denotamos por $\tilde x(t)$ el punto $(t,x(t),u(t))$. Evaluando W_t y W_y, uno encuentra que F está dada por

$$\begin{aligned}
F(t,y)\ \ =\ \ &\langle\dot p(t),y\rangle-\langle\dot x(t),V(t)(y-x(t))\rangle\ +\frac{1}{2}\langle y-x(t),\dot V(t)(y-x(t))\rangle\\
&+\langle p(t)+V(t)(y-x(t)),A(t)y+B(t)\xi(t,y)\rangle\ -L(t,y,\xi(t,y)).
\end{aligned}$$

Por lo tanto,

$$F_y(t,x(t)) = \dot{p}(t) + A(t)^T p(t) - L_x(\tilde{x}(t)) - L_u(\tilde{x}(t))\ \xi_y(t,x(t))$$
$$+B(t)^T p(t)\xi_y(t,x(t)) + V(t)A(t)x(t) + V(t)B(t)u(t) - V(t)\dot{x}(t)$$

y entonces $F_y(t,x(t)) = 0$. Para la segunda relación, tenemos que

$$\begin{aligned} F_{yy}(t,x(t)) \quad = \quad & \dot{V}(t) + V(t)A(t) + A(t)^T V(t) + V(t)B(t)\xi_y(t,x(t)) \\ & + B(t)^T V(t)\xi_y(t,x(t)) - L_{xx}(\tilde{x}(t)) \ - L_{xu}(\tilde{x}(t))\xi_y(t,x(t)) \\ & - L_{ux}(\tilde{x}(t))\xi_y(t,x(t)) \ - L_{uu}(\tilde{x}(t))\xi_y(t,x(t))\xi_y(t,x(t)). \end{aligned}$$

Ahora, para todo (t,y) en $T_0(x;\delta)$,

$$B(t)^T p(t) + B(t)^T V(t)(y - x(t)) = L_u(t,y,\xi(t,y))$$

y entonces

$$\xi_y(t,x(t)) = L_{uu}^{-1}(\tilde{x}(t))\{B(t)^T V(t) - L_{ux}(\tilde{x}(t))\}.$$

Se sigue que

$$F_{yy}(t,x(t)) = \dot{V}(t) + V(t)A(t) + A(t)^T V(t) - V(t)B(t)L_{uu}^{-1}(\tilde{x}(t))\ L_{ux}(\tilde{x}(t))$$
$$-L_{xu}(\tilde{x}(t))L_{uu}^{-1}(\tilde{x}(t))B(t)^T V(t) \ + V(t)B(t)L_{uu}^{-1}(\tilde{x}(t))B(t)^T V(t)$$
$$-L_{xx}(\tilde{x}(t)) + L_{xu}(\tilde{x}(t))L_{uu}^{-1}(\tilde{x}(t))\ L_{ux}(\tilde{x}(t))$$
$$= \dot{V}(t) + V(t)\alpha(t) + \alpha(t)^T V(t) + V(t)\beta(t)V(t) - \gamma(t).$$

Esto completa la prueba.

4 La condición de punto conjugado.

Consideremos un intervalo compacto $I = [a, b]$, dos puntos $\xi_0, \xi_1 \in \mathbf{R}^n$, dos conjuntos abiertos $X \subset \mathbf{R}^n, U \subset \mathbf{R}^m$, una función $L : I \times X \times U \to \mathbf{R}$, y dos matrices de funciones continuas $A(t)$ y $B(t)$ de dimensiones $n \times n$ y $n \times m$ respectivamente.

El problema de control óptimo que estudiaremos consiste en

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

sobre todas las funciones suaves por fragmentos $x, x \in PWS(I, X)$ y continuas por fragmentos $u, u \in PWC(I, U)$, que satisfacen

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t) & t \in I, \\ x(a) = \xi_0, x(b) = \xi_1 \end{cases} \qquad (R)$$

donde $u \in PWC(I, U)$ significa que u tiene sólo un número finito de discontinuidades $t_1, \ldots, t_p$, en los cuales ambos límites derecho e izquierdo, $u(t_i^+)$ y $u(t_i^-)$ existen y pertencen a U.

Definición 4.1 *Una pareja (x, u) se dice admisible para (P) si $x \in PWS(I, X)$, $u \in PWC(I, U)$ y (x, u) satisface (R).*

Definición 4.2 *Una pareja admisible (x, u) es una solución (local) para (P) si (al contraer X si es necesario) cualquier otra pareja admisible (y, v) satisface que $I(y, v) \geq I(x, u)$.*

Definimos $\mathcal{A} = I \times X \times U$.

El principio de Pontryagin proporciona condiciones necesarias de primer orden para un óptimo.

En lo que sigue denotaremos $\alpha(t) = \alpha(t, x(t), u(t))$

Teorema 4.1 (Pontryagin) *Supongamos que $L \in C^1(\mathcal{A}; x, u)$. Si (x, u) es una solución para (P), entonces existe una función suave por fragmentos $p : I \to \mathbf{R}^n$ y un número λ_0, tal que*

(a) $\lambda_0 \geq 0, \ \lambda_0 + |p(b)| \neq 0;$
(b) $\dot{p}(t) + A(t)^T p(t) = \lambda_0 L_x(t), \quad t \in I$
(c) $max\{\langle A(t)x(t) + B(t)u, p(t)\rangle$
$\quad -\lambda_0 L(t, x(t), u) : (t, x(t), u) \in \mathcal{A}\}$
$\quad = \langle A(t)x(t) + B(t)u(t), p(t)\rangle - \lambda_0 L(t), \ t \in I,$

Hacemos notar que cuando en el Teorema 4.1, $\lambda_0 \neq 0$, la pareja (x,u) se dice ser normal.

Por la condición (c) del Teorema tenemos que

(c$'$) $B(t)^T p(t) \;\; - \lambda_0 L_u(t) = 0,\; t \in I$.

Definición 4.3 *Una pareja admisible (x,u) es un extremo para (P), si existen una función suave por fragmentos p y un número $\lambda_0 \in \mathbf{R}$ que satisfacen las condiciones (a), (b) y (c') del Teorema 4.1.*

Dado un extremo (x,u) con correspondiente p y $\lambda_0 \in \mathbf{R}$, estamos interesados en encontrar condiciones bajo las cuales $\lambda_0 \neq 0$.

Consideremos el sistema variacional

$$\begin{cases} \dot{y}(t) = A(t)y(t) + B(t)v(t),\; t \in I, \\ y(a) = 0,\; y(b) = 0,\; v \in PWC(I, \mathbf{R}^m). \end{cases} \qquad (S)$$

Definición 4.4 *Una pareja admisible (x,u) se dice fuertemente normal sobre I si el sistema*

$$\begin{array}{l} \text{(i)} \;\; -\dot{p}(t) = A(t)^T p(t),\; t \in I \\ \text{(ii)} \; p(t)^T B(t) = 0,\; t \in I \end{array}$$

tiene solamente la solución $p \equiv 0$.

Teorema 4.2 *Sea (x,u) un extremo con correspondiente p y λ_0. Si (x,u) es fuertemente normal sobre I, entonces $\lambda_0 \neq 0$.*

Prueba. Supongamos que $\lambda_0 = 0$. Entonces la condiciones (b) y (c') del Teorema 4.1 se reducen a las condiciones (i) y (ii) de la definición 4.4. Se sigue que $p \equiv 0$, de donde $|p(b)| = 0$ y por lo tanto $\lambda_0 + |p(b)| = 0$, contradiciendo la condición (a) del Teorema 4.1.

Observación 4.1 *Sea (x,u) un extremo fuertemente normal. Si $p(t)$ satisface las condiciones (a), (b) y (c') del Teorema 4.1 y $p(t)$ es reemplazada por $\hat{p}(t) = \gamma p(t)$(que reescribimos como $p(t)$) y λ_0 reemplazado por $\hat{\lambda}_0 = \gamma\lambda_0$ donde γ es un número real positivo, entonces (a), (b) y (c') también se cumplen. Entonces una constante γ puede ser seleccionada de manera que $\lambda_0 = 1$, y las condiciones del Teorema 4.1 pueden ser escritas como*

$$\begin{cases} \dot{p}(t) + A(t)^T p(t) = L_x(t) \qquad \forall t \in I \\ B(t)^T p(t) = L_u(t) \qquad \forall t \in I \end{cases} \qquad (3.1)$$

1 Direcciones admisibles y variaciones.

En esta sección mostraremos cómo construir, para una pareja admisible (x, u) la cual es fuertemente normal, una dirección admisible v, esto es, obtendremos una familia de controles admisibles $u(\cdot, \epsilon)$ que contiene a u y cuya variación en $\epsilon = 0$ es exactamente v.

Denotemos por $x(\cdot, u)$ y $y(\cdot, v)$ las soluciones de

$$\begin{cases} \dot{x}(t) = A(t)x(t) + B(t)u(t), \ t \in I \\ x(a) = \xi_0 \end{cases} \tag{4.1}$$

y

$$\begin{cases} \dot{y}(t) = A(t)y(t) + B(t)v(t), \ t \in I \\ y(a) = 0 \end{cases} \tag{4.2}$$

correspondientes a u y v respectivamente. Por la unicidad de la solución de ecuaciones diferenciales, se obtiene fácilmente que

$$(\partial x/\partial u)\mid_u (v)(t) = y(t, v) \quad \forall t \in I.$$

donde $(\partial x/\partial u)|_u$ es la derivada de Fréchet de $x(\cdot, u)$ evaluada en u.

Teorema 4.3 *Supongamos que (x, u) es una pareja admisible la cual es fuertemente normal sobre I. Si (y, v) resuelve el sistema variacional (S), entonces existe $\delta > 0$ y una función*

$$\begin{aligned} u &: I \times (-\delta, \delta) \to \mathbf{R}^m \\ u &: (t, \epsilon) \to u(t, \epsilon) \end{aligned}$$

continua por fragmentos en t, C^2 en ϵ tal que:

$$\begin{aligned} &\text{(i) } u(t, 0) = u(t), \ \forall t \in I \\ &\text{(ii) } (\partial u/\partial \epsilon)|_{\epsilon=0}(t) = v(t), \quad \forall t \in I \\ &\text{(iii) } \textit{para } x(t, \epsilon) = x(t, u(\cdot, \epsilon)), \end{aligned}$$

$$x(b, \epsilon) = \xi_1, \ \forall \epsilon \in (-\delta, \delta)$$

y

$$(\partial x/\partial \epsilon)|_{\epsilon=0}(t) = y(t, v), \quad \forall t \in I$$

Prueba. Sean $\tau_0, \tau_1, \ldots, \tau_n$ puntos de un $(n+1)$-simplex en $\mathbf{R}^n$ que contiene a cero en su interior, i.e. n cualesquiera de ellos son linealmente independientes y cero puede ser escrito de manera única como su combinación convexa. En otras palabras, existen únicos $\beta_0, \ldots, \beta_n$ tales que

$$\sum_{i=0}^{n} \beta_i = 1 \qquad y \qquad \sum_{i=0}^{n} \beta_i \tau_i = 0.$$

Dado que la normalidad fuerte de (x, u) sobre I es equivalente a la controlabilidad de la ecuación (4.2), i.e., $\mathcal{R} = \mathbf{R}^n$ donde $\mathcal{R}$ es el conjunto solución de la ecuación variacional al tiempo $t = b$, existen $v_0, v_1, \ldots, v_n$, $v_i \in PWC(I, \mathbf{R}^m)$ tales que $y(b, v_i) = \tau_i$. Sea (y, v) una solución del sistema variacional. Sin pérdida de generalidad podemos suponer que

$$v(t) = \sum_{i=0}^{n} \beta_i v_i(t) \qquad \forall t \in I,$$

(de otra manera definimos $v_n = (1/\beta_n)(v - \sum_{i=0}^{n-1} \beta_i v_i))$. Definimos

$$\begin{array}{ll} u(t,\lambda) = u(t) + \sum_{i=0}^{n} \lambda_i v_i(t) & t \in I \\ F : B_{n+1}(\delta_1) \to \mathbf{R}^{n+1} & \\ F : \lambda \to (\sum_{i=0}^{n} \lambda_i, x(b,\lambda) - \xi_1) & \end{array} \tag{4.3}$$

donde

$$x(t, \lambda) = x(t, u(\cdot, \lambda))$$

Notamos que $F(0) = (0, x(b, 0) - \xi_1) = (0, x(b, u(\cdot, 0)) - \xi_1) = (0, x(b, u) - \xi_1) = (0, \xi_1 - \xi_1) = (0, 0)$,

$$\frac{\partial F}{\partial \lambda_i}(0) = \left(1, \left.\frac{\partial x}{\partial u}\right|_u \left.\frac{\partial u}{\partial \lambda_i}\right|_{\lambda=0}(b)\right) = \left(1, \left.\frac{\partial x}{\partial u}\right|_u (v_i)(b)\right) = (1, y(b, v_i)) = (1, \tau_i)\ \forall i.$$

Por lo tanto, $(\partial F/\partial \lambda)(0)$ es no singular. Por el Teorema de la función inversa, existe $\delta > 0$ tal que F es invertible sobre $B_{n+1}(\delta)$. Sea la función

$$\begin{array}{l} \lambda : (-\delta, \delta) \to \mathbf{R}^{n+1} \\ \lambda : F^{-1}|_A; \end{array}$$

donde

$$A = \{(x_1, \ldots, x_{n+1}) \in \mathbf{R}^{n+1} : x_1 = \epsilon,\ \ x_i = 0,\ i = 2, \ldots, n+1, \ \text{ y } \ -\delta < \epsilon < \delta\}$$

Entonces, $\forall \epsilon \in (-\delta, \delta)$, $\lambda(\cdot)$ es C^2,

$$\sum_{i=0}^{n} \lambda_i(\epsilon) = \epsilon \qquad y \qquad x(b, \lambda(\epsilon)) = \xi_1 \tag{4.4}$$

con

$$\lambda(0) = 0$$

Derivando (4.4) tenemos que

$$\sum_{i=0}^{n} \dot{\lambda}_i(0) = 1$$

y

$$0 = \left.\frac{\partial x}{\partial u}\right|_u \left.\frac{\partial u}{\partial \lambda}\right|_{\lambda=0}(b)\dot{\lambda}(0) = \left.\frac{\partial x}{\partial u}\right|_u \sum_{i=0}^{n} \left.\frac{\partial u}{\partial \lambda_i}\right|_{\lambda=0} \dot{\lambda}_i(0)(b)$$

$$= \left.\frac{\partial x}{\partial u}\right|_u \sum_{i=0}^{n} v_i(b)\dot{\lambda}_i(0) = \sum_{i=0}^{n} y(b, v_i)\ \dot{\lambda}_i(0) = \sum_{i=0}^{n} \dot{\lambda}_i(0)\tau_i.$$

Por la definición de los τ_i's, tenemos que $\dot{\lambda}_i(0) = \beta_i\ \forall i = 0, \ldots, n$. Finalmente usando (4.3) definimos sobre $I \times (-\delta, \delta)$

$$u(t, \lambda(\epsilon)) = u(t, \epsilon) = u(t) + \sum_{i=0}^{n} \lambda_i(\epsilon)v_i(t).$$

Es claro que $u(\cdot, \epsilon)$ es continua por fragmentos y $u(t, \cdot)$ es C^2, $u(t, 0) = u(t)$; $(\partial u/\partial\epsilon)|_{\epsilon=0}(t) = \sum_{i=0}^{n} \dot{\lambda}_i(0)v_i(t) = \sum_{i=0}^{n} \beta_i v_i(t) = v(t)\ \ \forall t \in I$. Definiendo $x(t, \epsilon) = x(t, u(\cdot, \epsilon)) = x(t, u(\cdot, \lambda(\epsilon))) = x(t, \lambda(\epsilon))$, por (4.4) $x(b, \epsilon) = \xi_1$, $\forall \epsilon$ en $-\delta < \epsilon < \delta$. Además,

$$\left.\frac{\partial x}{\partial \epsilon}\right|_{\epsilon=0}(t) = \left.\frac{\partial x}{\partial u}\right|_u \left.\frac{\partial u}{\partial \epsilon}\right|_{\epsilon=0}(t) = \left.\frac{\partial x}{\partial u}\right|_u (v)(t) = y(t, v)\ \forall t \in I.$$

Definición 4.5 *Definimos el conjunto Y de direcciones admisibles como*

$$Y = \{\ (y, v) \text{ que satisfacen (S) }\ \}$$

2 El problema accesorio.

Para obtener el problema accesorio correspondiente a (P) en un extremo (x, u), uno necesitaría encontrar una familia de controles admisibles $u(\cdot, \epsilon)$ que contenga a u para calcular la segunda variación con respecto a ϵ de la función objetivo $I(x(\cdot, \epsilon), u(\cdot, \epsilon))$. El Teorema 4.3 nos proporciona una familia tal.

Teorema 4.4 *Supongamos que $L \in C^1(\mathcal{A}; x, u)$. Sea (x, u) una solución para (P) la cual es fuertemente normal sobre I. Entonces existe $p \in PWS(I, \mathbf{R}^n)$ tal que:*

(a) $\dot{p}(t) + A(t)^T p(t) = L_x(t),\ \forall t \in I$
(b) $B(t)^T p(t) = L_u(t)\ , \forall t \in I$

y si además, $L \in C^2(\mathcal{A}; x, u)$, entonces

(c) $I''(y, v) \geq 0,\ \forall (y, v) \in Y,$ *donde*
$I''(y, v) = \int_a^b 2\Omega(t, y(t), v(t))dt$
y
$2\Omega(t, y, v) = L_{xx}(t)(y, y)\ + 2L_{xu}(t)(v, y)\ + L_{uu}(t)(v, v)$

Prueba. Los incisos (a) y (b) se siguen de (3.1). Sea $(y, v) \in Y$ y sea $(x(\cdot, \epsilon), u(\cdot, \epsilon))$ la familia correspondiente a (y, v) del Teorema 4.3. Si

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

entonces

$$F'(\epsilon) = \int_a^b \{\langle L_x(t, x(t, \epsilon), u(t, \epsilon))\, , \partial x(t, \epsilon)/\partial\epsilon\rangle + \langle L_u(t, x(t, \epsilon), u(t, \epsilon)), \partial u(t, \epsilon)/\partial\epsilon\rangle\}dt$$

Como $x(t, 0) = x(t)$ y $u(t, 0) = u(t)$, cero es un mínimo local de F y por lo tanto $F'(0) = 0$ y

$$F''(0) = \int_a^b \{L_{xx}(t)(y(t), y(t)) + L_{xu}(t)(v(t), y(t)) + L_{uu}(v(t), v(t))\}dt \geq 0,\ (y, v) \in Y.$$

3 Puntos conjugados y Ecuación de Riccati.

En esta sección se introduce la definición de un punto conjugado en una pareja (x, u) y se muestra que ésta puede ser escrita en términos del sistema lineal de una ecuación diferencial. Finalmente se prueba que, la no existencia de un punto en (a, b) conjugado a a es necesaria para un óptimo el cual es fuertemente normal en cada subintervalo $[c, b]$ y $[a, c]$ de $[a, b]$. Esta condición necesaria es equivalente a la existencia de una solución Q sobre (a, b) de una cierta ecuación de Riccati.

Definición 4.6 *Un punto $c \in (a, b]$ es conjugado a $t = a$ sobre (x, u) si existe una tripleta no nula (y, q, v) tal que y y q son suaves por fragmentos, y v es continua por fragmentos satisfaciendo para todo $t \in I$*

$$\begin{cases} \dot{y}(t) = A(t)y(t) + B(t)v(t), & (6.1)\\ \dot{q}(t) + A(t)^T q(t) = L_{xx}(t)y(t) + L_{xu}(t)v(t), & (6.2)\\ y(a) = 0,\ y(c) = 0, & (6.3)\\ B(t)^T q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t) = 0. & (6.4) \end{cases}$$

Observación. Supongamos que $L \in C^2(\mathcal{A}; x, u)$ y sea (x, u) una pareja admisible con u continua. Sea $(\bar{y}, \bar{v})$ una pareja en Y que minimiza el problema accesorio, esto es,

$$(AP) \qquad minimizar \qquad \left\{\frac{1}{2}I''(y, v) :\ (y, v) \in Y\right\}$$

La normalidad fuerte de (x, u) produce la normalidad de cualquier extremo $(y, v) \in Y$. En consecuencia, para (AP) obtenemos una función $\bar{q} \in PWS(I; \mathbf{R}^n)$

tal que $(\bar{y}, \bar{q}, \bar{v})$ satisfacen $(6.1), (6.2), (6.3)$ para $c = b$, y (6.4). Vale la pena mencionar que usando el siguiente resultado se puede probar fácilmente lo siguiente:

$$(y, q, v) \neq (0, 0, 0) \text{ satisface } (6.1) - (6.4) \text{ si y sólo si}$$

$$(y, q, v) \text{ satisface } (6.1) - (6.4) \text{ con } y \not\equiv 0$$

Proposición 4.1 *Supongamos que $L \in C^2(\mathcal{A}; x, u)$, y para todo $t \in I$, $L_{uu}(t) > 0$. Entonces, un punto $c \in (a, b]$ es conjugado a a si y sólo si existen funciones suaves por fragmentos no nulas y y q que satisfacen para todo $t \in I$*

$$\begin{aligned} \dot{y}(t) &= \alpha(t)y(t) + \beta(t)q(t) \\ \dot{q}(t) &= \gamma(t)y(t) - \alpha(t)^T q(t) \end{aligned} \tag{6.5}$$

con

$$y(a) = 0, \qquad y(c) = 0.$$

donde

$$\begin{aligned} \alpha(t) &= A(t) - B(t)L_{uu}^{-1}(t)L_{ux}(t) \\ \beta(t) &= B(t)L_{uu}^{-1}(t)B(t)^T \\ \gamma(t) &= L_{xx}(t) - L_{xu}(t)L_{uu}^{-1}(t)L_{ux}(t). \end{aligned}$$

Definición 4.7 *Una pareja admisible (x, u) se dice fuertemente normal sobre cualquier intervalo de la forma $[a, c] \subset [a, b]$, si para cualquier $c \in (a, b]$ el sistema*

$$\begin{cases} -\dot{p}(t) = A(t)^T p(t), \ t \in [a, c], \\ p(t)^T B(t) = 0, \ t \in [a, c]. \end{cases}$$

tiene sólo la solución cero.

La pareja (x, u) se dice fuertemente normal sobre cualquier intervalo de la forma $[c, b] \subset [a, b]$, si para cualquier $c \in (a, b)$ el sistema

$$\begin{cases} -\dot{p}(t) = A(t)^T p(t), \ t \in [c, b], \\ p(t)^T B(t) = 0, \ t \in [c, b] \end{cases}$$

tiene sólo la solución cero. El siguiente resultado afirma que, la no existencia de un punto $c \in (a, b)$ conjugado a a es necesaria para un óptimo.

Teorema 4.5 *Supongamos que $L \in C^2(\mathcal{A}; x, u)$. Si (x, u) es una solución para (P) la cual es fuertemente normal sobre cada intervalo de la forma $[a, c]$ y $[c, b]$, u continua y además*

$$L_{uu}(t) > 0, \ \forall t \in I,$$

entonces no existe ningún punto en (a, b) conjugado a a.

Prueba. Supongamos que existe un punto c en (a, b) conjugado a a. Definimos sobre I,

$$(\bar{y}(t), \bar{q}(t), \bar{v}(t)) = \begin{cases} (y(t), q(t), v(t)) & \text{si } t \in [a, c) \\ 0 & \text{si } t \in [c, b] \end{cases}$$

donde $(y, q, v) \neq 0$ satisface la definición 4.6. Entonces, $(\bar{y}, \bar{v})$ es admisible para el problema accesorio (AP). Observamos que

$$\begin{aligned} I''(\bar{y}, \bar{v}) &= \int_a^b \{L_{xx}(t)(\bar{y}(t), \bar{y}(t)) \; + 2L_{xu}(\bar{v}(t), \bar{y}(t)) \; + L_{uu}(t)(\bar{v}(t), \bar{v}(t))\} dt \\ &= \int_a^c \{\langle y(t), L_{xx}(t) y(t) \\ &\qquad + L_{xu}(t) v(t)\rangle \; + \langle v(t), L_{ux}(t) y(t) + L_{uu}(t) v(t)\rangle\} dt \\ &= \int_a^c \{\langle y(t), \dot{q}(t) + A(t)^T q(t)\rangle \; + \langle v(t), B(t)^T q(t)\rangle\} dt \\ &= \int_a^c \{\langle y(t), \dot{q}(t)\rangle \; + \langle y(t), A(t)^T q(t)\rangle \; + \langle v(t), B(t)^T q(t)\rangle\} dt \\ &= \int_a^c \{\langle y(t), \dot{q}(t)\rangle \; + \langle A(t) y(t) + B(t) v(t), q(t)\rangle\} dt \\ &= \int_a^c \{\langle y(t), \dot{q}(t)\rangle \; + \langle \dot{y}(t), q(t)\rangle\} dt \\ &= \int_a^c \frac{d}{dt} \langle y(t), q(t)\rangle dt \\ &= [\langle y(c), q(c)\rangle - \langle y(a), \; q(a)\rangle] = 0 \end{aligned}$$

Así, $(\bar{y}, \bar{v})$ es una solución diferente de cero del problema accesorio (AP). Dado que u es continua, tenemos que $\Omega(t, y, v) \in C^1(I \times \mathbf{R}^n \times \mathbf{R}^m)$. Aplicando el principio de Pontryagin a (AP), tomando en cuenta que $L_{uu}(t) > 0, \;\; \forall t \in I$, obtenemos una función suave por fragmentos $\tilde{q}$ que satisface con $(\bar{y}, \bar{v})$

$$\dot{\tilde{q}}(t) + A(t)^T \tilde{q}(t) \; = L_{xx}(t)\bar{y}(t) + L_{xu}(t)\bar{v}(t), \; \forall t \in I.$$

$$B(t)^T \tilde{q}(t) - L_{ux}(t)\bar{y}(t) \; - L_{uu}(t)\bar{v}(t) = 0, \; \forall t \in I.$$

La normalidad fuerte sobre $[c, b]$ produce que $\tilde{q} \equiv 0$ sobre $[c, b]$. Pero como

$$\dot{q}(t) + A(t)^T q(t) \; = L_{xx}(t)\bar{y}(t) + L_{xu}(t)\bar{v}(t), \; t \in [a, c)$$

$$B(t)^T q(t) - L_{ux}(t)\bar{y}(t) - L_{uu}(t)\bar{v}(t) = 0, \;\; t \in [a, c)$$

entonces,

$$-[\dot{\tilde{q}}(t) - \dot{q}(t)] = A(t)^T[\tilde{q}(t) - q(t)], \;\; t \in [a, d]$$
$$B(t)^T(\tilde{q}(t) - q(t)) = 0, \; t \in [a, d]$$

para todo $a < d < c$. Por la normalidad fuerte en $[a, d]$, tenemos que $q = \tilde{q}$ sobre $[a, d]$ para $a < d < c$. Puesto que $\tilde{q}$ y q son continuas en $I = [a, b]$, entonces

$$0 = \tilde{q}(c) = \lim_{d \to c^-} \tilde{q}(d) = \lim_{d \to c^-} q(d) = q(c)$$

Por lo tanto, $q(c) = y(c) = 0$. Entonces por la Proposición 4.1, $q \equiv y \equiv v \equiv 0$, contradiciendo el hecho de que $(y, q, v) \neq 0$.

Consideremos el sistema variacional correspondiente a (6.5):

$$\begin{aligned} \dot{Y}(t) &= \alpha(t)Y(t) + \beta(t)Q(t) \qquad (6.6)\\ \dot{Q}(t) &= \gamma(t)Y(t) - \alpha(t)^T Q(t) \end{aligned}$$

$$Y(a) = 0.$$

La Proposición 4.1 y el Teorema 4.5 implican el siguiente resultado.

Corolario 4.1 *Sea $L \in C^2(\mathcal{A}; x, u)$ y supongamos que $L_{uu}(t) > 0$ $\forall t \in I$, entonces (a) y (b) son equivalentes, donde*
(a). No existen puntos conjugados a $t = a$ en (a, b).
(b). Si (Y, Q) es una solución de (6.6) con $Y(a) = 0$, entonces $det\ Y(t) \neq 0$ $\forall t \in (a, b)$.

Corolario 4.2 *Sea $L \in C^2(\mathcal{A}; x, u)$ y supongamos que $L_{uu}(t) > 0$ $\forall t \in I$. La existencia de una solución (Y, Q) que satisface las ecuaciones (6.6) y det $Y(t) \neq 0$ sobre (a, b) es equivalente a la existencia de una función matricial simétrica suave por fragmentos Λ, tal que, para todo $t \in (a, b)$,*

$$L(\Lambda) = \dot{\Lambda}(t) + \Lambda(t)\beta(t)\Lambda(t) + \Lambda(t)\alpha(t) + \alpha(t)^T\Lambda(t) - \gamma(t) = 0.$$

Prueba. Sea (Y, Q) que satisfaga las ecuaciones (6.6) y *det* $Y(t) \neq 0$ sobre (a, b). Sea $\Lambda = QY^{-1}$. Tenemos que

$Y(t)^T Q(t) - Q(t)^T Y(t)$ es igual a cero en $t = a$ y su derivada es cero $\forall t \in I$.

En efecto,

$$\begin{aligned} & \frac{d}{dt}\{Y(t)^T Q(t) - Q(t)^T Y(t)\} \\ = \;& \dot{Y}(t)^T Q(t) + Y(t)^T \dot{Q}(t) - \dot{Q}(t)^T Y(t) - Q(t)^T \dot{Y}(t) \\ = \;& \{Y(t)^T \alpha(t)^T + Q(t)^T \beta(t)^T\} Q(t) + Y(t)^T \{\gamma(t) Y(t) - \alpha(t)^T Q(t)\} \\ & -\{Y(t)^T \gamma(t)^T - Q(t)^T \alpha(t)\} Y(t) - Q(t)^T \{\alpha(t) Y(t) + \beta(t) Q(t)\} \\ = \;& 0 \end{aligned}$$

Por lo tanto,

$$Y(t)^T Q(t) = Q(t)^T Y(t), \ \forall t \in I.$$

Además,

$$\Lambda Y = Q \Rightarrow (Y^{-1})^T Y^T \Lambda^T = (Y^{-1})^T Q^T \;\Rightarrow \Lambda^T Y = (Y^{-1})^T Q^T Y$$

$$\Rightarrow \Lambda^T Y = (Y^{-1})^T Y^T Q \Rightarrow \Lambda^T = QY^{-1} = \Lambda$$

Notamos que,

$$\frac{d}{dt}\{\Lambda Y\} = \dot{Q} = \gamma Y - \alpha^T Q = \dot{\Lambda} Y \; + \Lambda \dot{Y},$$

entonces

$$\dot{\Lambda} Y + \Lambda\{\alpha Y + \beta Q\} - \gamma Y + \alpha^T Q = 0 \;\Rightarrow \dot{\Lambda} + \Lambda\alpha + \Lambda\beta\Lambda \; - \gamma + \alpha^T \Lambda$$

$$= \dot{\Lambda} + \Lambda\beta\Lambda + \Lambda\alpha + \alpha^T\Lambda \; - \gamma = L(\Lambda) = 0,\; t \in (a,b).$$

Recíprocamente, sea Λ una función matricial simétrica suave por fragmentos sobre (a,b). Sea $Y(\cdot)$ la solución sobre (a,b) de

$$\dot{Y}(t) = [\alpha(t) + \beta(t)\Lambda(t)]Y(t)$$

$$Y(s) = I, \text{ para algún } s \in (a,b).$$

Entonces Y es invertible sobre (a,b), y $Q = \Lambda Y$ es tal que

$$\dot{Y}(t) = \alpha(t)Y(t) + \beta(t)Q(t)$$

Por lo tanto,

$$L(\Lambda)Y = \dot{\Lambda}Y + \Lambda\beta\Lambda Y + \Lambda\alpha Y + \alpha^T\Lambda Y - \gamma Y = 0 \;\text{ y } \dot{Q} = \Lambda\dot{Y} + \dot{\Lambda}Y$$

$$\Rightarrow 0 = \dot{Q} - \Lambda\dot{Y} + \Lambda\beta Q + \Lambda\alpha Y + \alpha^T Q - \gamma Y$$

Así,

$$\begin{aligned}
\dot{Q} &= \gamma Y - \alpha^T Q + \Lambda\dot{Y} - \Lambda\beta Q - \Lambda\alpha Y \\
&= \gamma Y - \alpha^T Q + \Lambda\{\alpha Y + \beta Q\} - \Lambda\beta Q - \Lambda\alpha Y \\
&= \gamma Y - \alpha^T Q + \Lambda\alpha Y + \Lambda\beta Q - \Lambda\beta Q - \Lambda\alpha Y \\
&= \gamma Y - \alpha^T Q.
\end{aligned}$$

5 Puntos conjugados generalizados.

Consideremos el problema de control

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

sobre todas las trayectorias suaves por fragmentos x con controles continuos por fragmentos u, que satisfacen

$$\begin{cases} \dot{x}(t)=A(t)x(t)+B(t)u(t) \\ x(a)=\xi_0, x(b)=\xi_1 \\ u(t)\in U \end{cases} \qquad (D)$$

donde U es un subconjunto abierto de $\mathbf{R}^m$ y $A(\cdot), B(\cdot)$ son matrices de funciones continuas sobre $[a,b]$.

Terminología. Un arco es una función suave por fragmentos $x:[a,b]\to\mathbf{R}^n$. Implícito en la declaración del problema (P) es un subconjunto abierto relativo Ω de $[a,b]\times\mathbf{R}^n$ para el cual el dominio de $f(t,x,u)=A(t)x+B(t)u$ y $L(t,x,u)$ es $\Omega\times U=\mathcal{A}$. Una pareja estado-control (x,u) es admisible para (P) si ésta satisface (D) y obedece que $(t,x(t))\in\Omega$ para todo $t\in[a,b]$. Una pareja admisible (x,u) es una solución (local) para (P) si (al contraer Ω si es necesario) cualquier otra pareja admisible (y,v) satisface que $I(y,v)\geq I(x,u)$. Una pareja admisible (x,u) es normal si la única solución al sistema

$$-\dot{p}(t)=A(t)^T p(t) \qquad t\in[a,b]$$

$$p(t)^T B(t)=0 \qquad t\in[a,b]$$

es $p\equiv 0$.

1 Condiciones necesarias de primer orden.

Supongamos que $L\in C^1(\mathcal{A};x,u)$ y que (x,u) es una solución normal para (P). Por el principio del máximo existe una función suave por fragmentos p que satisface;

(a.) $\dot{p}(t)=-H_x(t,x(t),p(t),u(t)); t\in[a,b]$
(b.) $H(t,x(t),p(t),u(t))\geq H(t,x(t),p(t),v), v\in\mathbf{R}^m$ con $(t,x(t),v)\in\Omega\times U$.

La condición (b) implica que

$$(\mathrm{b}'.)B(t)^T p(t)-L_u(t,x(t),u(t))=0 \qquad t\in[a,b].$$

donde $H(t,x,p,u)=\langle A(t)x+B(t)u,p\rangle-L(t,x,u)$. En lo que sigue al referirnos a una pareja admisible (x,u) denotaremos $L(t)=L(t,x(t),u(t))$.

Definición 5.1 *Una pareja admisible (x, u) es un extremo si existe un arco p, tal que (a) y (b') se cumplen.*

2 Condiciones necesarias de segundo orden.

Sea (x, u) un extremo con correspondiente p. Consideremos la siguiente familia de funciones

$$\mathcal{V} = \{\alpha v : \alpha \geq 0, v \in PWC[a, b], v(t) \in \mathbf{R}^m\}$$

Entonces $\mathcal{V}$ es un cono convexo de funciones continuas por fragmentos.

Definición 5.2 *El conjunto de control $\mathcal{V}$ se dice un conjunto de dirección admisible si para cualquier $v \in \mathcal{V}$, existen constantes $\tau > 0, M > 0$, y una función $u : [a, b] \times [-\tau, \tau] \to \mathbf{R}^m$ que satisface lo siguiente:*
$u(\cdot, \theta)$ es continua por fragmentos y $u(t, \cdot)$ es C^2;
$u(t, 0) = u(t), u_\theta(t, 0) = v(t)$;
$\|u(t, \cdot)\|_\infty$, $\|u_\theta(t, \cdot)\|_\infty \leq M$.

Ahora consideramos el sistema variacional asociado con (D),

$$\begin{cases} \dot{y}(t) = A(t)y(t) + B(t)v(t), \\ y(a) = 0, \\ v \in \mathcal{V} \end{cases} \qquad (L)$$

donde $\mathcal{V}$ es un conjunto de dirección admisible.

De manera similar al Teorema 4.4 de la sección anterior obtenemos el siguiente teorema.

Teorema 5.1 *Supongamos que $L \in C^2(\mathcal{A}; x, u)$ y sea (x, u) normal. Si (x, u) resuelve el problema (P), entonces para cualquier solución (y, v) de (L) que satisface $y(b) = 0$, tenemos que*

$$I''(y, v) = \int_a^b 2\Omega(t, y(t), v(t))dt \geq 0$$

donde

$$2\Omega(t, y, v) = L_{xx}(t)(y, y) + 2L_{xu}(t)(v, y) + L_{uu}(t)(v, v).$$

3 Puntos conjugados generalizados.

Definición 5.3 *Un punto $c \in (a, b]$ es un punto conjugado generalizado a $t = a$ sobre (x, u) si existe (y, q, v) sobre $[a, c]$ tal que*

$$\begin{cases} \dot{y}(t) = A(t)y(t) + B(t)v(t) \\ y(a) = 0, y(c) = 0, v \in PWC, v(t) \in \mathbf{R}^m, \\ \dot{q}(t) + A(t)^T q(t) = L_{xx}(t)y(t) + L_{xu}(t)v(t), \\ q(c) \neq 0 \\ \langle B(t)^T q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t), v(t)\rangle \geq 0 \end{cases}$$

y si además (a) o (b) se cumplen, donde

(a) $\langle B(t)^T q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t), v(t)\rangle > 0$ *sobre un conjunto de medida positiva; o*

(b) *existen un control v_1 y un arco y_1 tales que para todo $t \in [a,b]$*

$$\begin{cases} \dot{y}_1(t) = A(t)y_1(t) + B(t)v_1(t) \\ y_1(a) = 0,\ \langle y_1(c), q(c)\rangle < 0, y_1(b) = 0, \\ \langle B(t)^T q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t), v_1(t)\rangle \geq 0. \end{cases}$$

Observación. El problema clásico en cálculo de variaciones es minimizar el funcional $\int_a^b L(t, x(t), \dot{x}(t))dt$ sobre todos los arcos x que satisfacen las restricciones de punto final fijo $x(a) = \xi_0$ y $x(b) = \xi_1$. Para pasarlo a un problema de control óptimo, podemos tomar $f(t,x,u) = u, U = \mathbf{R}^n, A(t) = 0, B(t) = I$, y $H = \langle p, u\rangle - L(t,x,u)$. Claramente (x,u) es normal. Las condiciones de la definición 5.3 se refieren a una tripleta (y,q,v) y a un punto c para los cuales

$$\begin{cases} \dot{y}(t) = v(t), y(a) = 0, y(c) = 0, \\ \dot{q}(t) = L_{xx}(t)y(t) + L_{xu}(t)v(t), \\ q(c) \neq 0, v \in PWC, v(t) \in \mathbf{R}^n, \\ \langle q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t), v(t)\rangle \geq 0 \end{cases}$$

Estas condiciones ciertamente se cumplen para cualquier pareja suave por fragmentos (y,q) tal que

$$\begin{cases} \dot{q}(t) = L_{xx}(t)y(t) + L_{xu}(t)\dot{y}(t), y(a) = 0, y(c) = 0, \\ q(t) = L_{ux}(t)y(t) + L_{uu}(t)\dot{y}(t), q(c) \neq 0 \end{cases}$$

Si $L_{uu}(t) > 0$ entonces cualquier solución no trivial de la ecuación de Jacobi cumplirá las condiciones anteriores, puesto que

$$\dot{y}(t) = L_{uu}^{-1}(t)[q(t) - L_{ux}(t)y(t)]$$

por lo tanto $0 \neq \dot{y}(c) = L_{uu}^{-1}(c)q(c) \Rightarrow q(c) \neq 0$. Para ver que cualquier función y de este tipo también cumple (b), escogemos $y_1(t) = (a-t)(b-t)q(c), v_1(t) = \dot{y}_1$, entonces,

$$\begin{cases} \dot{y}_1(t) = v_1(t), \\ y_1(a) = 0, \langle y_1(c), q(c)\rangle = (a-c)(b-c)|q(c)|^2 < 0, y_1(b) = 0 \end{cases}$$

De la discusión anterior vemos que si c es un punto conjugado en cálculo de variaciones y la condición modificada de Legendre $L_{uu}(t) > 0$ se cumple, entonces c debe ser un punto conjugado generalizado. Escogiendo controles v de esta manera podemos encontrar todos los puntos conjugados en cálculo de variaciones. Por su puesto, podemos encontrar otros puntos conjugados generalizados escogiendo otros controles v. Por ejemplo, supongamos que x es un extremo suave para el cual la condición de Legendre no se cumple. Esto es, supongamos que existe un punto $c \in (a,b)$ y un vector unitario $\hat{u}$ tal que $\langle L_{uu}(c)\hat{u},\hat{u}\rangle < 0$. Mostraremos que c es un punto conjugado generalizado a a. Puesto que $\langle L_{uu}(c)\hat{u},\hat{u}\rangle < 0$, existe una $\delta > 0$, tal que $\langle L_{uu}(t)\hat{u},\hat{u}\rangle < 0$ sobre $[c-\delta,c]$. Por la continuidad de las funciones $L_{xx}(t), L_{xu}(t), L_{uu}(t)$, existe $\epsilon > 0$ y $M > 0$ tal que $\langle L_{uu}(t)\hat{u},\hat{u}\rangle \leq -\epsilon$ y $|L_{xx}(t)|$, $|L_{xu}(t)|$ están acotadas por M para todo $c-\delta \leq t \leq c$. Ahora sea $0 < d < 1$ lo suficientemente pequeño de tal manera que $[c-2d,c] \subset [c-\delta,c]$. Sean $c_1 = c-d$ y $c_2 = c-2d$, y consideremos las funciones

$$v(t) = \begin{cases} 0 & t \in [a,c_2) \\ -\hat{u} & t \in [c_2,c_1) \\ \hat{u} & t \in [c_1,c] \end{cases}$$

$$y(t) = \begin{cases} 0 & t \in [a,c_2) \\ -\hat{u}(t-c_2) & t \in [c_2,c_1) \\ \hat{u}(t-c) & t \in [c_1,c] \end{cases}$$

y $q(t) = d\hat{u} - \int_t^c [L_{xx}(s)y(s) + L_{xu}(s)v(s)]ds$. Esta selección de (y,q,v) satisface todas las condiciones básicas en la definición 5.3. Las tres primeras líneas son

$$\dot{y}(t) = v(t), y(a) = 0, y(c) = 0;$$

$$\dot{q}(t) = L_{xx}(t)y(t) + L_{xu}(t)v(t);$$

$$q(c) = d\hat{u} \neq 0, v \in PWC, v(t) \in \mathbf{R}^n.$$

Para confirmar la cuarta, necesitamos solamente considerar el intervalo $[c_2,c]$. Es fácil ver que

$$|q(t)| \leq d + 4Md \qquad \Rightarrow \qquad |\langle q(t) - L_{ux}(t)y(t), v(t)\rangle| \leq d + 5Md.$$

Y por lo tanto,

$$\langle q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t), v(t)\rangle \geq -(1+5M)d + \epsilon$$

Si, además a las restricciones impuestas anteriormente, insistimos en que $d < \epsilon/(1+5M)$, entonces (y,q,v) satisface todas las condiciones básicas en la definición 5.3 y la condición suplementaria (a). Por lo tanto, c es un punto conjugado generalizado a $t=a$.

El resultado principal.

El siguiente resultado proporciona una construcción concreta de cómo mostrar que si una pareja (x,u) tiene un punto conjugado generalizado en (a,b), entonces la segunda variación es negativa para algún control admisible.

Teorema 5.2 *Supongamos que $L \in C^2(\mathcal{A};x,u)$. Si existe un punto conjugado generalizado a $t=a$ en (a,b), entonces existe una pareja (y,v) que satisface el sistema (L) con $y(b)=0$ tal que $I''(y,v) < 0$.*

Prueba. Supongamos que $c \in (a,b)$ es un punto conjugado generalizado a a. Usamos la tripleta (y,q,v) de la definición 5.3 para definir

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

y

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

Entonces $(\bar{y},\bar{v})$ es una solución de (L). Tenemos que,

$$\begin{aligned}
I''(\bar{y},\bar{v}) &= \int_a^b 2\Omega(t,\bar{y}(t),\bar{v}(t))dt \\
&= \int_a^c \{L_{xx}(t)(y(t),y(t)) + 2L_{xu}(t)(v(t),y(t)) + L_{uu}(t)(v(t),v(t))\}dt \\
&\leq \int_a^c \{\langle \dot{q}(t) + A(t)^T q(t), y(t)\rangle + \langle B(t)^T q(t), v(t)\rangle\}dt \\
&= \int_a^c \{\langle \dot{q}(t), y(t)\rangle + \langle q(t), A(t)y(t) + B(t)v(t)\rangle\}dt \\
&= \int_a^c \{\langle \dot{q}(t), y(t)\rangle + \langle q(t), \dot{y}(t)\rangle\}dt \\
&= \int_a^c \frac{d}{dt}\langle q(t), y(t)\rangle dt = 0
\end{aligned}$$

Si la condición (a) de la definición 5.3 se cumple, la anterior desigualdad es estricta, y por lo tanto $I''(\bar{y},\bar{v}) < 0$. De otra manera, la condición (b) se cumple, y se tiene $I''(\bar{y},\bar{v}) = 0$. Para $\alpha > 0$, definimos el control $v_\alpha = v_1 + \alpha\bar{v}$. Claramente $v_\alpha \in \mathcal{V}$, y el arco $y_\alpha(t) = y_1(t) + \alpha\bar{y}(t)$ es la solución al sistema (L) correspondiente a v_α. Por lo tanto,

$$\begin{aligned}
I''(y_\alpha,v_\alpha) &= \int_a^b 2\Omega(t,y_\alpha(t),v_\alpha(t))dt \\
&= \int_a^b 2\Omega(t,y_1(t) + \alpha\bar{y}(t), v_1(t) + \alpha\bar{v}(t))dt \\
&= \Bigg\{ \int_a^b \{L_{xx}(t)(y_1,y_1) + 2L_{xu}(t)(v_1,y_1) + L_{uu}(t)(v_1,v_1)\}dt \\
&\quad + \int_a^b \{\alpha^2 L_{xx}(t)\,(\bar{y},\bar{y}) + 2\alpha^2 L_{xu}(t)\,(\bar{v},\bar{y}) + \alpha^2 L_{uu}(t)(\bar{v},\bar{v})\}dt \\
&\quad + \int_a^c 2\alpha\{L_{xx}(t)(y(t),y_1(t)) + L_{xu}(t)(v(t),y_1(t)) \\
&\qquad + L_{xu}(t)(v_1(t),y(t)) + L_{uu}(t)(v_1(t),v(t))\}dt \Bigg\}
\end{aligned}$$

$$
\begin{aligned}
\leq\ & \Big\{ I''(y_1, v_1) + \alpha^2 I''(\bar{y}, \bar{v}) \\
& + 2\alpha \int_a^c \{\langle \dot{q}(t) + A(t)^T q(t), y_1(t)\rangle + \langle B(t)^T q(t), v_1(t)\rangle\} dt \Big\} \\
=\ & I''(y_1, v_1) + 2\alpha \int_a^c \{\langle \dot{q}(t), y_1(t)\rangle + \langle A(t) y_1(t) + B(t) v_1(t), q(t)\rangle\} dt \\
=\ & I''(y_1, v_1) + 2\alpha \int_a^c \{\langle \dot{q}(t), y_1(t)\rangle + \langle q(t), \dot{y}_1(t)\rangle\} dt \\
=\ & I''(y_1, v_1) + 2\alpha \int_a^c \frac{d}{dt}\{\langle q(t), y_1(t)\rangle\} dt \\
=\ & I''(y_1, v_1) + 2\alpha \langle y_1(c), q(c)\rangle.
\end{aligned}
$$

El lado derecho se aproxima a $-\infty$ cuando $\alpha \to \infty$, entonces para α suficientemente grande, tenemos que $I''(y_\alpha, v_\alpha) < 0$. Combinando los Teoremas 5.1 y 5.2, se obtiene el resultado principal.

Teorema 5.3 *Supongamos que $L \in C^2(\mathcal{A}; x, u)$ y sea (x, u) normal. Si (x, u) resuelve el problema (P), entonces el intervalo (a, b) no contiene puntos conjugados generalizados a $t = a$.*

Ejemplo 5.1. Consideremos el siguiente problema de cálculo de variaciones con espacio de estados dos-dimensional x.

$$
\textit{minimizar} \qquad I(x, u) = \int_0^T (u_2(t)^2 - x_1(t)^2) dt
$$

sujeto a

$$
\dot{x} = u, x(0) = 0, x(T) = 0, u \in \mathbf{R}^2, T > 1.
$$

Tenemos que $H(x, p, u) = p_1 u_1 + p_2 u_2 - u_2^2 + x_1^2$. El problema es normal. La pareja $(x, u) = (0, 0)$ es un extremo con $p = 0$,

$$
A(t) = 0, B(t) = I_2, L_{xu} = 0,
$$

$$
L_{xx} = \begin{pmatrix} -2 & 0 \\ 0 & 0 \end{pmatrix}
$$

$$
L_{uu} = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}
$$

Un punto $c \in (0, T)$ es un punto conjugado generalizado a $t = 0$ si existe una tripleta (y, q, v), tal que para todo $t \in [0, c]$

$$
\begin{cases} \dot{y} = v, y(0) = 0, y(c) = 0, \\ -\dot{q}^1 = 2y^1, \dot{q}^2 = 0, q(c) \neq 0, \\ q^1 v^1 + (q^2 - 2v^2) v^2 \geq 0, v \in \mathbf{R}^2. \end{cases} \qquad (*)
$$

y si además (a) o (b) se cumplen, donde

(a) $q^1v^1 + (q^2 - 2v^2)v^2 > 0$ sobre un conjunto de medida positiva; o

(b) existen un control v_1 y un arco y_1 tales que para todo $t \in [0, T]$

$$\begin{cases} \dot{y}_1 = v_1, \\ y_1(0) = 0, \langle y_1(c), q(c)\rangle < 0, y_1(T) = 0, \\ q^1v_1^1 + (q^2 - 2v^2)v_2^1 \geq 0 \end{cases}$$

donde $v_1 = (v_1^1, v_2^1)$. Sea $c = 1, v^1(t) = 1$ sobre $[0, 1/3]$ y $v^1(t) = -1/2$ sobre $(1/3, 1]$, y $v^2(t) = 0$ en $(*)$. Así, se obtienen las soluciones

$$y^1(t) = \begin{cases} t & 0 \leq t \leq 1/3 \\ -1/2(t-1) & 1/3 \leq t \leq 1 \end{cases}$$

$$q^1(t) = \begin{cases} (-t + 1/3)(t + 1/3) & 0 \leq t \leq 1/3 \\ 1/2(t - 1/3)(t - 5/3) & 1/3 \leq t \leq 1 \end{cases}$$

y $y^2 = 0, q^2 = 0$. Claramente, $q^1v^1 + (q^2 - 2v^2)v^2 = q^1v^1 > 0$ en un conjunto de medida positiva. Por lo tanto, $c = 1$ es un punto conjugado generalizado a $t = 0$ para todo $T > 1$. Entonces $(x, u) = (0, 0)$ no es óptimo para ningún valor de $T > 1$.

Nótese que la condición modificada de Legendre $L_{uu}(t) > 0$ no se cumple en este problema, puesto que $\langle L_{uu}(t)(a, 0), (a, 0)\rangle = 0$ $(a \neq 0)$. En consecuencia, la teoría clásica de puntos conjugados no se puede aplicar. Esta condición es requerida por los resultados de ambos [Zeidan & Zezza] y [J.F.Rosenblueth], por lo tanto el ejemplo cae más allá del alcance de ambos trabajos. Por su puesto, estas hipótesis no pueden ser simplemente ignoradas: aplicando la definición 4.6 [Zeidan & Zezza] a este problema tenemos que $c \in (0, T]$ es conjugado a $t = 0$ si existe una tripleta no nula (y, q, v) tal que y y q son suaves por fragmentos, y v es continua por fragmentos satisfaciendo para todo $t \in [0, T]$

$$\begin{cases} \dot{y} = v; \; y(0) = 0; \; y(c) = 0, \\ \dot{q}^1 = -2y^1; \; \dot{q}^2 = 0, \\ q^1 = 0; \; q^2 = 2v^2. \end{cases}$$

Por lo tanto $y \equiv 0$, lo cual implica que no existen puntos conjugados sobre $(0, T]$. La Definición 3.2 [J.F.Rosenblueth] de punto conjugado difiere de la anterior sólo por la condición de que $(y, v) \in E(x, u)$, pero en cálculo de variaciones, esta condición es equivalente a la condición de Jacobi. Por lo tanto también nos lleva a la no existencia de puntos conjugados sobre $(0, T]$.

6 Comparación entre diferentes trabajos.

Recordemos algunos resultados importantes.

Definición 4.6 Un punto $c \in (a, b]$ es conjugado a $t = a$ sobre (x, u) si existe una tripleta no nula (y, q, v) tal que y y q son suaves por fragmentos, y v es continua por fragmentos satisfaciendo para todo $t \in I$

$$\begin{cases} \dot{y}(t) = A(t)y(t) + B(t)v(t) \\ \dot{q}(t) + A(t)^T q(t) = L_{xx}(t)y(t) + L_{xu}(t)v(t), \\ y(a) = 0, \; y(c) = 0, \\ B(t)^T q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t) = 0, \end{cases}$$

Teorema 4.5 Supongamos que $L \in C^2(\mathcal{A}; x, u)$. Si (x, u) es una solución para (P) la cual es fuertemente normal sobre cada intervalo de la forma $[a, c]$ y $[c, b]$, u continua y además

$$L_{uu}(t) > 0, \; \forall t \in I,$$

entonces no existe ningún punto en (a, b) conjugado a a.

Definición 3.2 Un punto $c \in (a, b]$ es *conjugado a $t = a$* sobre (x, u) si existen $(y, v) \in E(x, u) \cap D$ y $q \in X$ que satisfacen

$$B(t)^T q(t) = L_{ux}(\tilde{x}(t))y(t) + L_{uu}(\tilde{x}(t))v(t) \qquad (t \in T)$$

tal que $y(a) = y(c) = 0$ y $y(t) \not\equiv 0$ sobre (a, c).

Teorema 3.4 Supongamos que $L \in C^2(\mathcal{A}; x, u)$, $(x, u) \in S(\mathcal{A})$ no singular, $(\dagger)$ se cumple, y u continua, entonces $(x, u) \in J$.

Definición 5.3 Un punto $c \in (a, b]$ es un punto conjugado generalizado a $t = a$ sobre (x, u) si existe (y, q, v) sobre $[a, c]$ tal que

$$\begin{cases} \dot{y}(t) = A(t)y(t) + B(t)v(t) \\ y(a) = 0, y(c) = 0, v \in PWC, v(t) \in \mathbf{R}^m, \\ \dot{q}(t) + A(t)^T q(t) = L_{xx}(t)y(t) + L_{xu}(t)v(t), \\ q(c) \neq 0 \\ \langle B(t)^T q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t), v(t) \rangle \geq 0 \end{cases}$$

y si además (a) o (b) se cumplen, donde

(a) $\langle B(t)^T q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t), v(t) \rangle > 0$ sobre un conjunto de medida positiva; o

(b) existen un control v_1 y un arco y_1 tales que para todo $t \in [a, b]$

$$\begin{cases} \dot{y}_1(t) = A(t)y_1(t) + B(t)v_1(t) \\ y_1(a) = 0, \; \langle y_1(c), q(c) \rangle < 0, y_1(b) = 0, \\ \langle B(t)^T q(t) - L_{ux}(t)y(t) - L_{uu}(t)v(t), v_1(t) \rangle \geq 0. \end{cases}$$

Teorema 5.3 Supongamos que $L \in C^2(\mathcal{A}; x, u)$ y sea (x, u) normal. Si (x, u) resuelve el problema (P), entonces el intervalo (a, b) no contiene puntos conjugados generalizados a $t = a$.

Comparación. [J.F.Rosenblueth] vs [Zeidan & Zezza]

Teorema 6.1 *Sea (x, u) una pareja admisible. Entonces, (x, u) es fuertemente normal sobre cada intervalo de la forma $[a, c] \subset [a, b]$ $(c \in (a, b])$ si y sólo si $(\dagger)$ se cumple.*

Prueba. Supongamos que (x, u) no es fuertemente normal en $[a, b]$, entonces existe $p \not\equiv 0$ en X tal que p satisface el sistema

$$\begin{cases} -\dot{p}(t) = A(t)^T p(t) & t \in [a, b] \\ p(t)^T B(t) = 0 & t \in [a, b] \end{cases}$$

Sea $(y_1, v_1) \in Y$ tal que $y_1(a) = y_1(c) = y_1(b) = 0$, y $p(c)^T \dot{y}_1(c) \neq 0$, para algún $c \in (a, b)$. Así,

$$\begin{aligned} p(t)^T \dot{y}_1(t) &= p(t)^T A(t) y_1(t) + p(t)^T B(t) v_1(t) \qquad (t \in [a, b]) \\ &= -\dot{p}(t)^T y_1(t) + p(t)^T B(t) v_1(t) \end{aligned}$$

Entonces,

$$p(t)^T \dot{y}_1(t) + \dot{p}(t)^T y_1(t) = p(t)^T B(t) v_1(t) = 0 \qquad (t \in [a, b]),$$

Así,

$$0 = p(c)^T \dot{y}_1(c) + \dot{p}(c)^T y_1(c) = p(c)^T \dot{y}_1(c) \neq 0$$

lo cual es una contradicción.

Supongamos que (x, u) es fuertemente normal sobre cada intervalo de la forma $[a, c]$, y sea y un arco que cumple con $y(a) = y(b) = 0$. Tenemos que si un arco w satisface la ecuación diferencial

$$\dot{w}(t) = A(t) w(t) + B(t) v(t) \qquad (L)$$

sobre $[a, b]$ para algún control v, entonces existe una vecindad $\mathcal{F}_w$ de w en el *espacio* $- tw$ tal que si z es un arco definido en un intervalo $[a, c] \subset [a, b]$ para el cual sus elementos $(t, z(t))$ están en $\mathcal{F}_w$, entonces existe un control v_z que satisface con z la ecuación diferencial (L).

Por la contralabilidad sobre todo intervalo de la forma $[a, c] \subset [a, b]$, existe $\epsilon_1 > 0$ lo suficientemente pequeño, y un arco z_1 que satisface (L) sobre el intervalo $[a, a + \epsilon_1]$ y las condiciones $z_1(a) = 0$ y $z_1(a + \epsilon_1) = y(a + \epsilon_1)$, y tal que $y(t)$ $(t \in [a, a + \epsilon_1])$ está en la vecindad $\mathcal{F}_{z_1}$ de z_1, y en consecuencia existe un control v_0 tal que

$$\dot{y}(t) = A(t) y(t) + B(t) v_0(t) \quad t \in [a, a + \epsilon_1].$$

Luego entonces, existe $\epsilon_2 > 0$ lo suficientemente pequeño y un arco z_2 que satisface (L) sobre el intervalo $[a, a+\epsilon_1+\epsilon_2]$ y las condiciones $z_2(a) = 0$ y $z_2(a+\epsilon_1+\epsilon_2) = y(a+\epsilon_1+\epsilon_2)$, y tal que $y(t)$ $(t \in [a+\epsilon_1, a+\epsilon_1+\epsilon_2])$ está en la vecindad $\mathcal{F}_{z_2}$ de z_2, y en consecuencia existe un control v_1 sobre $[a+\epsilon_1, a+\ \epsilon_1+\epsilon_2]$ tal que

$$\dot{y}(t) = A(t)y(t) + B(t)v_1(t) \quad t \in [a+\epsilon_1, a+\epsilon_1+\epsilon_2]$$

Procediendo de esta manera, y dejando que $t_0 = a, t_1 = a+\epsilon_1,\ t_2 = a+\epsilon_1+\epsilon_2, \ldots, t_n = a + \sum_{i=1}^{n} \epsilon_i = b$, para algún $n \in \mathbf{N}$, obtenemos una colección de arcos $\{z_j\}$ $(j = 1, \ldots, n)$, y una colección de controles $\{v_i\}$ $(i = 0, \ldots, n-1)$ tal que

$$\dot{y}(t) = A(t)y(t) + B(t)v_i(t) \quad t \in [t_i, t_{i+1}]$$

y $\bigcup_{i=0}^{n-1}[t_i, t_{i+1}] = [a, b]$. Por lo tanto (†) se cumple.

Notamos que las definiciones 3.2 y 4.6, sólo difieren en los conjuntos

$$E(x,u) = \{(y,v) \in Z \mid J_{(x,u)}((y,v);(z,w)) = 0 \text{ para todo } (z,w) \in Y\}$$

$$\text{y } \left\{(y,v) \in Z \mid \dot{q}(t) + A(t)^T q(t) \right.$$

$$\left. = L_{xx}(\tilde{x}(t))y(t) + L_{xu}(\tilde{x}(t))v(t) \ \ (t \in T) \text{ para algún } q \in X \right\} \qquad (\bullet)$$

respectivamente. Por la Proposición 3.9, si

$$B(t)^T q(t) = L_{ux}(\tilde{x}(t))y(t) + L_{uu}(\tilde{x}(t))v(t) \ \ (t \in T)$$

se cumple, y (y,v) satisface $(\bullet) \Rightarrow (y,v) \in E(x,u)$, y si (†) se cumple, entonces ambas afirmaciones son equivalentes.

Por lo tanto, el conjunto de punto conjugado de la definición 4.6 [Zeidan & Zezza] está contenido en el conjunto de punto conjugado de la definición 3.2 [J.F. Rosenblueth]. Sin embargo, al aplicar alguno de los Teoremas 4.5 o 3.4, estos conjuntos son iguales, por lo tanto, ambos trabajos coinciden.

Comparación.
[J.F.Rosenblueth, Zeidan & Zezza] vs [P.D.Lowen & H.Zheng]

Notamos que la definición 4.6 y (si (†) se cumple) la definición 3.2 cumplen todas las condiciones de la definición 5.3 [P.D.Lowen & H.Zheng] excepto $q(c) \neq 0$ y la condición (b). Vemos que, la definicón 4.6 y la definición 3.2 son una extensión natural del cálculo de variaciones, y en consecuencia son más simples que la definición 5.3. Pero hay un precio que pagar por esto. Para aplicar el Teorema 4.5, (x,u) debe de ser fuertemente normal en ambas esquinas del intervalo $[a,b]$, así como para aplicar el Teorema 3.4, (†) se debe verificar y en ambos, una versión de la condición modificada de Legendre se debe cumplir. Sin embargo, bajo estas condiciones los conjuntos de punto conjugado de las definiciones 4.6 y 3.2 son un subconjunto del conjunto derivado de la definición 5.3. Si (y, q, v) satisfacen las definiciones 4.6 y 3.2, al ser no nulos, y si se cumple la versión modificada de Legendre $L_{uu}(t) > 0, \forall t \in I$, entonces por la Proposición 4.1 [Zeidan & Zezza] (o por la Proposición 3.9 [J.F.Rosenblueth]) implica que

$$\begin{aligned}&0 \neq \dot{y}(c) = -B(c)L_{uu}^{-1}(c)B(c)^T q(c)\\&\Rightarrow q(c) \neq 0\end{aligned}$$

Además, escogiendo un arco y_1 tal que

$$y_1(a) = y_1(b) = 0,\ y_1(c) = -q(c) \qquad y \qquad \dot{y}_1(t) = A(t)y_1(t) + B(t)v_1(t)$$

para algún control v_1, la condición (b) de la definición 5.3 es satisfecha. En el Teorema 5.3 [P.D.Lowen & H.Zheng] no se requiere la continuidad de u, la normalidad fuerte sobre la esquina $t = a$, ni la condición modificada de Legendre. También, en la definición de punto conjugado generalizado la condición (b) es más débil que la condición de controlabilidad fuerte. El ejemplo 5.1 muestra cómo el extremo $(x, u) = (0, 0)$ no es óptimo para ningún valor de $T > 1$. Los resultados de [Zeidan & Zezza] y [J.F.Rosenblueth] no pueden detectar esto. Por lo tanto la definición de punto conjugado generalizado es más precisa y más general.

Bibliografía

1 Frank H. Clarke & Vera Zeidan (1985) Sufficiency and the Jacobi Condition in the Calculus of Variations.

2 Hestenes, M. R, Calculus of variations and optimal control theory, John Wiley, New York, (1966)

3 H. L. Royden, Real Analysis, Macmillan Publishing Company (New York), Collier Macmillan Publishing (London), (1988)

4 P. D. Loewen & H. Zheng (1994) Generalized Conjugate Points for Optimal Control Problems: Nonlinear Analysis, Theory, Methods and Applications. Vol 22, No. 6: 771-791

5 Rosenblueth JF (1996) Jacobi's Conditions for the Lagrange Problem with Linear Dynamics, Memorias del XXVIII Congreso Nacional de la Sociedad Matemática Mexicana, Colima, Aportaciones Matemáticas, Serie Comunicaciones **18**: 125-138

6 Rosenblueth JF (1996) Sufficiency and Conjugate Points for the Lagrange Problem with Linear Dynamics, Memorias del XXVIII Congreso Nacional de la Sociedad Matemática Mexicana, Colima, Aportaciones Matemáticas, Serie Comunicaciones **18**: 139-153

7 Rosenblueth JF (1996) Variational Conditions for the Lagrange Problem with Linear Dynamics, Memorias del XXVIII Congreso Nacional de la Sociedad Matemática Mexicana, Colima, Aportaciones Matemáticas, Serie Comunicaciones **18**: 155-170

8 Rosenblueth JF (1999) Variational Conditions and Conjugate Points for the Fixed-Endpoint Control Problem, IMA Journal of Mathematical Control and Information, **16**: 1-17

9 V. Zeidan & P. Zezza (1988) The Conjugate Point Condition for Smooth Control Sets: Journal of Mathematical Analysis and Applications **32**: 572-589

10 V. Zeidan & P. Zezza (1988) Necessary Conditions for Optimal Control Problems: Conjugate Points: SIAM J. Control and Optimization Vol. 26: 592-608

11 Wendell H. Fleming & Raymond W. Rishel Deterministic and Stochastic Optimal Control, Springer-Verlag New York Heidelberg Berlin (1975)