UNIVERSIDAD NACIONAL AUTÓNOMA DE MEXICO
PROGRAMA DE MAESTRÍA Y DOCTORADO EN CIENCIAS MATEMÁTICAS Y
DE LA ESPECIALIZACIÓN EN ESTADÍSTICA APLICADA

NORMALIDAD DÉBIL PARA CONDICIONES DE SEGUNDO ORDEN EN EL PROBLEMA ISOPERIMÉTRICO DE LAGRANGE

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

PRESENTA:
JORGE ANTONIO BECERRIL GÓMEZ

TUTOR PRINCIPAL
JAVIER FERNANDO ROSENBLUETH LAGUETTE
UNAM-IIMAS

COMITÉ TUTOR
RICARDO BERLANGA ZUBIAGA
UNAM-IIMAS
GERARDO SÁNCHEZ LICEA
UNAM-DEPARTAMENTO DE MATEMÁTICAS

CIUDAD DE MÉXICO, SEPTIEMBRE DEL 2017

Para mis padres, Hilda y Antonio.

Agradecimientos

Antes que a nadie debo agradecer a mis padres por su apoyo y amor incondicional. Sin ellos este trabajo hubiera resultado imposible. También quiero agradecer a Javier por todo el tiempo, la paciencia y las palabras de aliento que me dio mientras trabajaba en este proyecto. Finalmente, pero no menos importante, quiero agradecer a Karla, quien estuvo apoyándome todo el tiempo y con quien compartí largas horas de estudio, frustración y alegría.

Contenido

Introducción

Sin duda una de las áreas más ricas de las matemáticas en cuanto a historia, teoría y aplicaciones se refiere, es el cálculo de variaciones. Con sus herramientas se puede resolver una gran diversidad de problemas, por ejemplo el de encerrar la mayor área posible con una cuerda de perímetro fijo (el problema isoperimétrico), hallar el camino más corto sobre una superficie dada (geodésicas) o el de hallar una curva en el plano que una dos puntos fijos, uno arriba del otro, de forma que una partícula que viaje sobre ella hará el recorrido en el menor tiempo posible (la braquistocrona). Como se puede apreciar, la finalidad de estos problemas es el de hallar un «óptimo» en una familia de elementos que satisfacen restricciones prescritas.

En este trabajo de investigación nos centraremos en los problemas del tipo isoperimétrico. Estos reciben este nombre por ser similares al famoso problema mencionado anteriormente. La familia de problemas así denominados tiene la particularidad de que, tanto la función que se busca minimizar, como las restricciones que se deben de satisfacer, pueden ser escritas como una integral que dependerá suavemente de una curva y de su derivada.

En la literatura podemos encontrar condiciones similares a las usadas en el cálculo diferencial para hallar condiciones de primer y segundo orden que debe de cumplir una curva óptima. Estas condiciones son la anulación de la primera variación (condición equivalente a las famosas ecuaciones de Euler-Lagrange), y la no negatividad de la segunda variación en cierto conjunto. Existen distintas hipótesis bajo las cuales estas condiciones se cumplen y de acuerdo a estas también varía el conjunto donde se satisface la no negatividad de la segunda variación. Ejemplo de esto puede encontrarse en los trabajos mencionados en la bibliografía.

Son dos las contribuciones principales de este trabajo a las condiciones necesarias de segundo orden. La primera es ver que en gran parte de la literatura la hipótesis usada es equivalente a lo que llamaremos *normalidad fuerte* o *normalidad relativa a* S_0 (más adelante quedará claro el motivo de esta nomenclatura), y la segunda es debilitar esta hipótesis sin afectar el conjunto donde la segunda variación es no negativa.

Para entender claramente de dónde viene la nueva noción que sustituirá a la normalidad fuerte, es necesario recordar el caso básico en dimensión finita y notar que las condiciones clásicas para problemas isoperimétricos no son análogas a estas. Esto se hará al inicio del Capítulo 1. Después se fijará la notación usada a lo largo de este trabajo y se planteará el problema que nos concierne. Para finalizar el capítulo se dan resultados básicos necesarios en el trabajo posterior.

En el Capítulo 2 se discutirán las condiciones clásicas de primer y segundo orden con hipótesis de regularidad y normalidad, y la relación que hay entre ellas, partiendo del caso más simple con restricciones de igualdad para llegar al caso general con restricciones de igualdad y desigualdad. A continuación se ven algunas equivalencias de normalidad fuerte usadas ampliamente en la literatura, y para concluir el capítulo, se da un ejemplo que ilustra algunas de las ideas más importantes.

En el Capítulo 3 se generalizan las herramientas presentadas en el caso de dimensión finita para que también abarquen nuestro problema. Primero se usan para deducir las condiciones de primer y segundo orden dadas en el Capítulo 2, y después se plantea la noción de normalidad débil (normalidad relativa a S). Se hace notar que nuestra nueva condición de segundo orden será cierta bajo la suposición de que esta nueva noción de normalidad implica regularidad. El Capítulo 4 se dedica a la prueba de este hecho, y se concluye con un ejemplo donde es imposible aplicar la teoría clásica para discernir si un arco es solución o no, pero la condición de segundo orden derivada en el Capítulo 3 nos permitirá llegar a una conclusión.

Por último, en el Capítulo 5 se da un breve resumen comparando nuestros resultados principales con los resultados clásicos y se extiende el trabajo realizado en los capítulos anteriores al contexto de control óptimo.

Capítulo 1

Antecedentes

1.1. El Caso en Dimensión Finita

Para entender claramente la generalización al caso de dimensión infinita, se dará una breve exposición del caso en dimensión finita. Las definiciones y resultados aquí presentados son en su mayoría de [19], mientras que en [14] se puede hallar uno de los tratados más completos del tema.

Supongamos que tenemos funciones $f, g_i : \mathbf{R}^n \to \mathbf{R}$ $(i \in A \cup B)$. El *conjunto de restricciones* asociado está dado por

$$S = \{x \in \mathbf{R}^n \mid g_\alpha(x) \leq 0 \ (\alpha \in A),\ g_\beta(x) = 0 \ (\beta \in B)\},$$

donde $A = \{1, \dots, q\}$ y $B = \{q+1, \dots, m\}$. Consideremos el problema $P(S)$ de minimizar f en S. Supondremos adicionalmente que las funciones f, g_i son continuas en una vecindad de $x_0 \in S$ y poseen primera y segunda diferencial en x_0.

Definición 1.1.1. *1. El conjunto de índices activos en x_0 está dado por*

$$I_a(x_0) = \{\alpha \in A \mid g_\alpha(x_0) = 0\}.$$

2. El conjunto de restricciones tangenciales de S en x_0 está definido por

$$R_S(x_0) = \{h \in \mathbf{R}^n \mid g'_\alpha(x_0; h) \leq 0 \ (\alpha \in I_a(x_0)),\ g'_\beta(x_0; h) = 0 \ (\beta \in B)\},$$

donde $g'_\alpha(x_0; h) = g'_\alpha(x_0)h$ *es la derivada de* g_α *en* x_0 *con dirección* h*.*

Es conveniente aproximar al conjunto de restricciones, al menos localmente, por uno más sencillo de estudiar. Esta idea motivó una gran variedad de «conos tangentes» (ver [1] y [14]). El cono tangente con el que trabajaremos aquí será llamado cono tangente secuencial o simplemente cono tangente, mientras que el cono tangente más conocido en la literatura será llamado cono tangente curvilíneo.

Definición 1.1.2. *1. El cono tangente secuencial (o simplemente cono tangente) de* S *en* x_0 *está definido por*

$$T_S(x_0) = \left\{ h \in \mathbf{R}^n \mid \exists [\{x_i\}_{i=1}^{\infty} \subset S, \{t_i\}_{i=1}^{\infty} \subset \mathbf{R}^+] \ni t_i \to 0,\ \frac{x_i - x_0}{t_i} \to h \right\}.$$

2. El cono tangente curvilíneo de S *en* x_0 *está dado por*

$$C_S(x_0) = \{h \in \mathbf{R}^n \mid \exists [\delta > 0, y : [0, \delta] \to S] \ni y \text{ es continua, } y(0) = x_0,\ y'(0) = h\}.$$

En general $C_S(x_0) \subset T_S(x_0) \subset R_S(x_0)$, pero no se tienen las contenciones contrarias. Los siguientes importantes resultados son bien conocidos y se pueden consultar en [19, C.4, Sección 4].

Proposición 1.1.3. *1.* $T_S(x_0)$ *es un cono cerrado.*
2. $C_S(x_0) \subset T_S(x_0) \subset R_S(x_0)$*.*
3. Si S *es un conjunto convexo entonces* $C_S(x_0) = T_S(x_0)$*.*

El segundo inciso de esta proposición es la motivación para la siguiente definición.

Definición 1.1.4. *El punto* x_0 *será llamado regular de* S *cuando* $T_S(x_0) = R_S(x_0)$*.*

El siguiente teorema nos da condiciones necesarias de primer orden para una solución de $P(S)$. Para su demostración consultar [19, p. 224].

Teorema 1.1.5. *Supongamos que* x_0 *es regular de* S *y es solución local de* $P(S)$*. Entonces existen multiplicadores* $\lambda_i \in \mathbf{R}$ $(i \in A \cup B)$*, con*

$$\lambda_\alpha \geq 0, \quad \lambda_\alpha g_\alpha(x_0) = 0 \quad (\alpha \in A), \tag{1.1.1}$$

$$F'(x_0) = 0, \tag{1.1.2}$$

donde $F := f + \langle \lambda, g \rangle$, $\lambda = (\lambda_1, \ldots, \lambda_m)$ *y* $g = (g_1, \ldots, g_m)$*.*

A la función F se le llama el Lagrangiano. Para estudiar las condiciones necesarias de segundo orden requerimos de la noción de *extremos* que es motivada por el teorema anterior.

Definición 1.1.6. *El conjunto de extremos $\mathcal{E}$ del problema $P(S)$ son los elementos $(x, \lambda) \in S \times \mathbf{R}^m$ tales que, junto con F definida como en el teorema anterior satisfacen las relaciones (1.1.1) y (1.1.2).*

Otra forma de expresar el Teorema 1.1.5 es la siguiente: si x_0 es una solución local regular de $P(S)$ entonces existe $\lambda \in \mathbf{R}^m$ con $(x_0, \lambda) \in \mathcal{E}$.

Las condiciones de segundo orden son más delicadas. Consideremos el conjunto

$$S_0 = \{x \in \mathbf{R}^n \mid g_i(x) = 0 \ (i \in I_a(x_0) \cup B)\},$$

junto con sus restricciones tangenciales asociadas

$$R_{S_0}(x_0) = \{h \in \mathbf{R}^n \mid g_i'(x_0; h) = 0 \ (i \in I_a(x_0) \cup B)\}.$$

Es claro que si x_0 es solución local de $P(S)$ entonces también es solución local de $P(S_0)$. De la condición clásica de Lagrange de segundo orden para un mínimo local de un problema con solo restricciones de igualdad $P(S_0)$ se tiene el siguiente resultado (ver [19, p. 186] o [14, p. 213]).

Teorema 1.1.7. *Si $(x_0, \lambda) \in \mathcal{E}$, x_0 es una solución local de $P(S_0)$ y la matriz jacobiana $\bar{g}'(x_0)$ de la función $\bar{g} = (g_i : (i \in I_a(x_0) \cup B))$ evaluada en x_0 tiene rango completo, entonces*

$$F''(x_0; h) := \langle F''(x_0)h; h\rangle \geq 0$$

para todo $h \in R_{S_0}(x_0)$, donde $F''(x_0)$ es la matriz Hessiana de F en x_0.

Se puede probar el siguiente resultado similar como sigue.

Teorema 1.1.8. *Si $(x_0, \lambda) \in \mathcal{E}$ y x_0 es una solución local de $P(S_0)$ regular de S_0, entonces*

$$F''(x_0; h) \geq 0$$

para todo $h \in R_{S_0}(x_0)$.

Demostración. Supongamos que $h \in T_{S_0}(x_0)$ con sucesiones asociadas $\{x_i\} \subset S_0$ y $\{t_i\} \subset \mathbf{R}^+$. De la fórmula de Taylor de segundo orden obtenemos

$$\frac{F(x_i) - F(x_0) - \langle F'(x_0); x_i - x_0 \rangle}{t_i^2} = \frac{1}{2} F'' \left(x_0; \frac{x_i - x_0}{t_i} \right) + \frac{R_2(x_0; x_i - x_0)}{t_i^2},$$

y, tomando el límite de ambos lados de la igualdad, obtenemos

$$\lim_i \frac{F(x_i) - F(x_0)}{t_i^2} = \frac{1}{2} F''(x_0; h),$$

ya que $F'(x_0) = 0$ y R_2 es $o(|x_i - x_0|^2)$. Como $x_i \in S_0$ tenemos que

$$\frac{F(x_i) - F(x_0)}{t_i^2} = \frac{f(x_i) - f(x_0)}{t_i^2} \geq 0$$

ya que x_0 resuelve $P(S_0)$. □

De acuerdo al siguiente resultado (ver [19, p. 166]) y al inciso (2) de la Proposición 1.1.3, el Teorema 1.1.8 es más general que el Teorema 1.1.7.

Teorema 1.1.9. *Si $x_0 \in S$ y la matriz jacobiana $\bar{g}'(x_0)$ del Teorema 1.1.7 es de rango completo, entonces $C_S(x_0) = R_S(x_0)$. En particular, x_0 es regular de S.*

Una diferencia importante es que en el Teorema 1.1.7 los multiplicadores son únicos, mientras que en el Teorema 1.1.8 no necesariamente.

Podemos mejorar el conjunto donde las condiciones de segundo orden del Teorema 1.1.8 son válidas tomando en cuenta el signo de los multiplicadores. Para $\lambda \in \mathbf{R}^m$ que satisface (1.1.1) definimos los conjuntos

$$\Gamma_+(\lambda) = \{\alpha \in A \mid \lambda_\alpha > 0\}, \quad \Gamma_0(\lambda) = \{\alpha \in A \mid \lambda_\alpha = 0\}.$$

Asociados a estos conjuntos definimos las restricciones

$$S_1(\lambda) = \{x \in \mathbf{R}^n \mid g_\alpha(x) \leq 0 \ (\alpha \in \Gamma_0(\lambda)), \ g_\beta(x) = 0 \ (\beta \in \Gamma_+(\lambda) \cup B)\}.$$

junto con sus restricciones tangenciales en $x_0 \in S_1(\lambda)$ por

$$R_{S_1}(x_0) = \{h \in \mathbf{R}^n \mid g'_\alpha(x_0; h) \leq 0 \ (\alpha \in I_a(x_0) \cap \Gamma_0(\lambda)), \ g'_\beta(x_0; h) = 0 \ (\beta \in \Gamma_+(\lambda) \cup B)\}.$$

Teorema 1.1.10. *Si* $(x_0, \lambda) \in \mathcal{E}$, x_0 *es solución local de* $P(S)$ *y* x_0 *es regular con respecto de* $S_1(\lambda)$, *entonces*

$$F''(x_0; h) \geq 0$$

para toda $h \in R_{S_1}(x_0)$.

Para la demostración ver [19, p. 227].

En general verificar regularidad es difícil, por lo que requerimos criterios más sencillos que impliquen esta condición. A este tipo de criterios se les llamará de *normalidad*.

Definición 1.1.11. *Un punto* $x_0 \in S$ *será llamado normal relativo a* S *si la única solución* $\lambda \in \mathbf{R}^m$ *al sistema*

1. $\lambda_\alpha \geq 0$, $\lambda_\alpha g_\alpha(x_0) = 0$ $(\alpha \in A)$,
2. $\lambda^* g'(x_0) = 0$, *donde* $*$ *significa transpuesta,*

es $\lambda = 0$.

Esta es la definición de normalidad que aparece en [19, p. 243], mientras que en [19, p. 241] se demuestra lo siguiente.

Proposición 1.1.12. *Los siguientes enunciados son equivalentes.*

1. x_0 *es normal de* S.
2. *Los gradientes* $\{g'_\beta(x_0) : \beta \in B\}$ *son linealmente independientes y existe* $h \in R_S(x_0)$ *de forma que* $g'_\alpha(x_0)h < 0$ $(\alpha \in I_a(x_0))$.

La Definición 1.1.11 depende del conjunto de restricciones S. Podemos aplicar esta misma definición a los conjuntos S_0 y S_1, de donde obtenemos las siguientes definiciones.

Definición 1.1.13. *1.* x_0 *es normal con respecto de* S_0 *si*

$$[\lambda_\alpha g_\alpha(x_0) = 0 \ (\alpha \in A),\ \lambda^* g'(x_0) = 0] \Rightarrow \lambda = 0.$$

2. x_0 *es normal con respecto de* S_1 *si*

$$[\mu_\alpha \geq 0,\ \mu_\alpha g_\alpha(x_0) = 0 \ (\alpha \in \Gamma_0(\lambda)),\ \mu^* g'(x_0) = 0] \Rightarrow \mu = 0.$$

Notemos que la definición de normalidad con respecto a S_0 coincide con la hipótesis usada en el Teorema 1.1.7 sobre el rango de la matriz jacobiana.

En algunos textos (por ejemplo [18, 20]) el resultado básico es el siguiente.

Teorema 1.1.14. *Si $x_0 \in S$ es normal de S_0, entonces es regular de S_0, S_1 y S.*

Por medio de este resultado podemos modificar el Teorema 1.1.7 de la siguiente manera.

Teorema 1.1.15. *Si $x_0 \in S$ es solución local de $P(S)$ y normal de S_0, entonces existe $\lambda \in \mathbf{R}^m$ con $(x_0, \lambda) \in \mathcal{E}$ y*

$$F''(x_0; h) \geq 0$$

para todo $h \in R_{S_1}(x_0)$.

Es decir, suponiendo normalidad con respecto a S_0 tenemos condiciones de segundo orden no sólo en S_0, sino también en el conjunto S_1 que lo contiene. Este resultado puede verificarse en [19, 20]. Más aún, es posible obtener el mismo resultado con hipótesis menos restrictivas. El siguiente teorema es crucial para continuar con este esfuerzo y se puede consultar en [19, p. 241, Teorema 10.4].

Teorema 1.1.16. *Si $x_0 \in S$ es normal con respecto de S, entonces también es regular con respecto de S.*

En vista de este resultado podemos enunciar condiciones de segundo orden más fuertes que las anteriores, en las que podemos reemplazar normalidad de S_0 por normalidad de S_1. Esto será consecuencia de [19, p. 227 Teorema 7.5 y p. 241 Teorema 10.4] y de la definición de normalidad relativa a S. A continuación lo escribimos explícitamente.

Teorema 1.1.17. *Sea $x_0 \in S$ solución local de $P(S)$ y $\lambda \in \mathbf{R}^m$ con $(x_0, \lambda) \in \mathcal{E}$. Si x_0 es normal relativo a S_1, entonces*

$$F''(x_0; h) \geq 0$$

para todo $h \in R_{S_1}(x_0)$.

En gran parte de la literatura del cálculo de variaciones y control óptimo, las condiciones de segundo orden se demuestran bajo hipótesis de normalidad fuerte (normalidad con respecto al conjunto que corresponde a S_0), dejando abierta la pregunta de la existencia de un resultado análogo al Teorema 1.1.17. En la Sección 3 se mostrará que la respuesta es afirmativa en el caso del cálculo de variaciones, mientras que en [9] también se da una respuesta afirmativa para casos particulares en problemas de control.

Por último, estaríamos tentados a preguntarnos si con nuestras mismas hipótesis las condiciones de segundo orden se satisfacen en el conjunto más general $R_S(x_0)$. La respuesta es no. En [7, p. 159, Teorema 3.3] aparentemente se dan hipótesis más restrictivas que las nuestras bajo las cuales condiciones necesarias de segundo orden se satisfacen en este conjunto. Sin embargo, aún con estas hipótesis más restrictivas el teorema es falso, y un contraejemplo se puede hallar en [14, p. 307].

1.2. El Caso en Dimensión Infinita

En esta sección introducimos la notación y el contexto en el que trabajaremos. Sea $\mathbf{X}$ el espacio de funciones vector-valuadas definidas en un compacto, $x : [a,b] \to \mathbf{R}^n$, que son continuas y poseen derivada continua a trozos. Es decir existe una partición finita del intervalo $[a,b]$

$$a = t_0 < t_1 < \cdots < t_k = b,$$

de forma que x restringida a cada subintervalo $[t_i, t_{i+1}]$ $(i = 0, \ldots, k-1)$ es de clase C^1.

Usaremos la notación $\dot{x}$ para denotar a la derivada de x con respecto a t. Si t_i es un punto esquina de x (es decir, la derivada no es continua en ese punto) denotaremos por $\dot{x}(t_i)$ al límite $\dot{x}(t_i-)$ si estamos considerando el intervalo $[t_{i-1}, t_i]$ o el límite $\dot{x}(t_i+)$ si estamos hablando del intervalo $[t_i, t_{i+1}]$. Por $\tilde{x}(t)$ denotaremos al elemento $(t, x(t), \dot{x}(t)) \in [a,b] \times \mathbf{R}^n \times \mathbf{R}^n$.

Trataremos a $\mathbf{X}$ como espacio vectorial real normado.

Definición 1.2.1. *Definimos la norma débil en* $\mathbf{X}$ *por*

$$\|x\| = \sup_{a \le t \le b} \{|x(t)|^2 + |\dot{x}(t)|^2\}^{1/2}.$$

Conservaremos gran parte de la notación de la Sección 1.1. Definimos los conjuntos de índices $A = \{1, \ldots, q\}, B = \{q+1, \ldots, m\}$, el conjunto $\mathbf{X}_e$ por

$$\mathbf{X}_e = \{x \in \mathbf{X} \mid x(a) = X_a,\ x(b) = X_b\},$$

donde X_a y X_b son elementos fijos de $\mathbf{R}^n$, el *conjunto de restricciones*

$$S = \{x \in \mathbf{X}_e \mid I_\alpha(x) \leq 0 \ (\alpha \in A),\ I_\beta(x) = 0 \ (\beta \in B)\},$$

donde los funcionales I_γ $(\gamma \in A \cup B)$ están definidos por $I_\gamma(x) = \int_a^b f_\gamma(\tilde{x}(t))dt$ y las funciones f, f_γ : $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ $(\gamma \in A \cup B)$ se suponen de clase C^2. A los elementos de $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$ los denotaremos por las ternas $(t, x, \dot{x})$ sin temor a confundir la variable independiente $\dot{x}$ con la derivada de una función $\dot{x}(t)$. El contexto aclarará cualquier confusión. Las parciales de las funciones serán denotadas por $f_{x^i}, f_{\dot{x}^j}, f_{x^i\dot{x}^j}$, etc. mientras que f_x y $f_{\dot{x}}$ son los gradientes

$$f_x = (f_{x^1}, \ldots, f_{x^n}), \quad f_{\dot{x}} = (f_{\dot{x}^1}, \ldots, f_{\dot{x}^n}),$$

y $f_{xx}, f_{x\dot{x}}, f_{\dot{x}\dot{x}}$ son las matrices de segundas derivadas parciales correspondientes.

El problema que nos concierne, nuevamente llamado $P(S)$, es el *problema isoperimétrico de Lagrange con puntos extremos fijos*, que consiste en minimizar el funcional $I(x) = \int_a^b f(\tilde{x}(t))dt$ sobre el conjunto S. A los elementos de $\mathbf{X}$ les llamaremos *arcos*, mientras que a los elementos de S les llamaremos *arcos admisibles.*

El siguiente resultado es fundamental en cálculo de variaciones. Nos dice explícitamente quién es la primera y segunda derivada (en el sentido de Fréchet) de los funcionales que tienen la forma de I con respecto de la normal débil. Podemos hallar la demostración en [18, C.2, Sección 6]

Proposición 1.2.2. *1. La primera derivada de I en el punto $x_0 \in \mathbf{X}$ es la función*

$$I'(x_0; \cdot) : \mathbf{X} \to \mathbf{R},$$

donde

$$I'(x_0; y) = \int_a^b \{f_x(\tilde{x}_0(t))y(t) + f_{\dot{x}}(\tilde{x}_0(t))\dot{y}(t)\}dt.$$

2. La segunda derivada de I en el punto $x_0 \in \mathbf{X}$ es la función

$$I''(x_0; \cdot) : \mathbf{X} \to \mathbf{R},$$

donde

$$I''(x_0; y) = \int_a^b 2\Omega(\tilde{y}(t))dt$$

y

$$2\Omega(t, y, \dot{y}) := \langle y; f_{xx}(\tilde{x}_0(t))y\rangle + 2\langle y; f_{x\dot{x}}(\tilde{x}_0(t))\dot{y}\rangle + \langle \dot{y}; f_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{y}\rangle.$$

La primera y segunda derivada de los funcionales I_γ $(\gamma \in A \cup B)$ están dadas de manera análoga, cambiando en el integrando la función f por f_γ respectivamente. En la literatura es muy común usar la palabra *variación* como sinónimo de derivada de un funcional. Nosotros usaremos estas dos palabras indistintamente, aunque en general nos referiremos a la primera y segunda derivada de un funcional como su primera y segunda variación.

Definición 1.2.3. *A* $\mathbf{Y}$*, el subespacio normado de* $\mathbf{X}$ *cuyos elementos satisfacen* $y(a) = 0 = y(b)$ *se le llamará el conjunto de variaciones admisibles.*

Asociado al conjunto de restricciones, tenemos el conjunto de restricciones tangenciales

$$R_S(x_0) = \{y \in \mathbf{Y} \mid I'_\alpha(x_0; y) \leq 0 \ (\alpha \in I_a(x_0)),\ I'_\beta(x_0; y) = 0 \ (\beta \in B)\},$$

y los conjuntos

$$S_0(x_0) = \{x \in \mathbf{X}_e \mid I_\gamma(x) = 0 \ (\gamma \in I_a(x_0) \cup B)\},$$

$$S_1(\lambda) = \{x \in \mathbf{X}_e \mid I_\alpha(x) \leq 0 \ (\alpha \in \Gamma_0(\lambda)),\ I_\beta(x) = 0 \ (\beta \in \Gamma_+(\lambda) \cup B)\},$$

con sus respectivas restricciones tangenciales

$$R_{S_0}(x_0) = \{y \in \mathbf{Y} \mid I'_i(x_0; y) = 0 \ (i \in I_a(x_0) \cup B)\},$$

$$R_{S_1}(x_0) = \{y \in \mathbf{Y} \mid I'_\alpha(x_0; y) \leq 0 \ (\alpha \in I_a(x_0) \cap \Gamma_0(\lambda)),\ I'_\beta(x_0; y) = 0 \ (\beta \in \Gamma_+(\lambda) \cup B)\},$$

donde los conjuntos $I_a(x_0)$, $\mathcal{E}$, $\Gamma_0(\lambda)$ y $\Gamma_+(\lambda)$ tienen el mismo significado que en la sección anterior al sustituir

$g'_\gamma(x_0;\cdot)$ por $I'_\gamma(x_0;\cdot)$. Explícitamente el conjunto $\mathcal{E}$ son los pares $(x,\lambda) \in S \times \mathbf{R}^m$ con

1. $\lambda_\alpha \geq 0, \quad \lambda_\alpha I_\alpha(x_0) = 0 \quad (\alpha \in A)$,

2. Si $J(x) := I(x) + \sum_{\gamma=1}^m \lambda_\gamma I_\gamma(x)$, entonces $J'(x_0;y) = 0$ para todo $y \in \mathbf{Y}$.

En la literatura existen diversas condiciones de primer y segundo orden para el problema que acabamos de definir y su equivalente en teoría de control. Algunas de estas condiciones son análogas a las vistas en la sección anterior, mientras que otras tienen diferencias cruciales. Se remite al lector a los artículos [3, 4, 10, 15–17, 21, 22, 26–28, 31] y a las referencias en estos para profundizar en ellas.

Como se mencionó antes, en el Capítulo 2 se muestra que las hipótesis usadas en la mayoría de las referencias corresponden al concepto de normalidad fuerte, o normalidad relativa a S_0 con la notación de la sección anterior. Esta normalidad se define como sigue.

Definición 1.2.4. *Se dirá que $x_0 \in S$ es fuertemente normal o normal relativo a S_0 si las primeras variaciones $\{I'_\gamma(x_0;\cdot) : \gamma \in I_a(x_0) \cup B\}$ son linealmente independientes en $\mathbf{Y}$.*

Por lo tanto nuestra pregunta persiste, ¿se tiene un análogo al Teorema 1.1.17 en dimensión infinita?. Esta pregunta se responde en los Capítulos 3 y 4.

1.3. Resultados Importantes

En esta sección se enuncian algunos resultados necesarios para abordar las demostraciones de los capítulos siguientes.

Los siguientes resultados son válidos para funcionales lineales en un espacio vectorial real $\mathbf{W}$.

Lema 1.3.1. [18, p. 12] *Sean $L_1, \ldots, L_m$ funcionales lineales en $\mathbf{W}$. Los siguientes enunciados son equivalentes.*

1. *$\{L_i\}_{i=1}^m$ son linealmente independientes.*

2. *Existen elementos $w_1, \ldots, w_m \in \mathbf{W}$ de forma que el determinante $|L_i(w_j)| \neq 0$ $(i, j = 1, \ldots, m)$.*

Lema 1.3.2. [18, p. 12] *Sean $L, L_1, \ldots, L_m$ funcionales lineales en $\mathbf{W}$. Supongamos que $L(w) = 0$ en el subconjunto de $\mathbf{W}$ definido por las restricciones*

$$L_i(w) = 0 \quad (1 \le i \le m).$$

Entonces existen multiplicadores $\lambda_1, \ldots, \lambda_m$ *de forma que*

$$L(w) + \sum_{j=1}^{m} \lambda_j L_j(w) = 0.$$

Si los funcionales $L_1, \ldots, L_m$ *son linealmente independientes, entonces los multiplicadores son únicos.*

Lema 1.3.3. [18, p. 13] *Sean* $L, L_1, \ldots, L_m$ *funcionales lineales en* $\mathbf{W}$*. Supongamos que* $L(w) \ge 0$ *en el subconjunto de* $\mathbf{W}$ *definido por las restricciones*

$$L_i(w) \le 0 \quad (1 \le i \le q), \qquad L_j(w) = 0 \quad (q+1 \le j \le m).$$

Entonces existen multiplicadores $\lambda_1, \ldots, \lambda_m$ *con* $\lambda_i \ge 0$ $(1 \le i \le q)$ *y de forma que*

$$L(w) + \sum_{j=1}^{m} \lambda_j L_j(w) = 0$$

para todo $w \in \mathbf{W}$*. Si los funcionales* $L_1, \ldots, L_m$ *son linealmente independientes, entonces los multiplicadores son únicos.*

Lema 1.3.4. [19, p. 201] *Sean* $L_1, \ldots, L_m$ *funcionales lineales en* $\mathbf{W}$*. Si éstos se anulan en el conjunto definido por los elementos* $w \in \mathbf{W}$ *con*

$$L_i(w) \le 0 \quad (1 \le i \le q), \qquad L_j(w) = 0 \quad (q+1 \le j \le m),$$

entonces existen multiplicadores $\lambda_1, \ldots, \lambda_m$*, con* λ_i *positivos para* $i = 1, \ldots, q$ *y de forma que*

$$\lambda_1 L_1(w) + \cdots + \lambda_m L_m(w) = 0$$

para todo $w \in \mathbf{W}$*.*

Lema 1.3.5. [19, p. 59] *Sea* $\langle \cdot ; \cdot \rangle$ *un producto interno en* $\mathbf{W}$*. Entonces un conjunto de vectores* $y_1, \ldots, y_n$ *es linealmente independiente si y solo si el determinante*

$$|(\langle y_i ; y_j \rangle)_{ij}| \quad (i, j = 1, \ldots, n)$$

es distinto de cero.

Requeriremos de la siguiente versión del teorema de la función implícita. Consideremos el sistema de ecuaciones

$$f_i(t^1,\ldots,t^m,x^1,\ldots,x^n)=0 \quad (i=1,\ldots,n),$$

donde f está definida en un conjunto abierto $F \subset \mathbf{R}^m \times \mathbf{R}^n$ y es de clase C^1 con respecto de las variables $x_1,\ldots,x_n$. Definamos $D(t,x)$ por

$$D(t,x)=\left|\left(\frac{\partial f_i}{\partial x^j}(t,x)\right)_{ij}\right| \quad (i,j=1,\ldots,n).$$

Teorema 1.3.6. [18, p. 23] (**Función Implícita**) Supongamos que existen funciones continuas

$$x_0^i(t) \quad (i=1,\ldots,n)$$

definidas en un conjunto compacto T_0, de forma que si $t\in T_0$ se cumple que $(t,x_0(t))\in F$ y

$$f_i(t,x_0(t))=0, \quad D(t,x_0(t))\neq 0.$$

Entonces existen funciones continuas

$$x^i(t) \;\; (i=1,\ldots,n)$$

definidas en una vecindad T de T_0 y una constante $\epsilon>0$ de forma que

$$x^i(t)=x_0(t) \text{ si } t\in T_0, \quad f_i(t,x(t))=0 \text{ si } t\in T,$$

y además las relaciones

$$f_i(t,x)=0, \quad |x-x(t)|<\epsilon \;\; (i=1,\ldots,n)$$

se cumplen simultáneamente si y solo si $x=x(t)$. Más aún, si las funciones f_i son de clase C^k, las funciones x^i también serán de clase C^k en T.

El siguiente resultado engloba algunos teoremas fundamentales de la teoría clásica de ecuaciones diferenciales lineales. Sea $A(t)$ una matriz de dimensión $n\times n$ con componentes continuas definidas en $[a,b]$, y consideremos la ecuación diferencial lineal homogénea

$$\dot{x}(t) = A(t)x(t) \quad (t \in [a,b]). \tag{1.3.1}$$

Podemos resumir [2, p. 123, Teorema 11.2] y las notas debajo de éste como sigue.

Teorema 1.3.7. *1. Cualquier combinación lineal de soluciones del sistema (1.3.1) sigue siendo solución.*

2. Si dos soluciones x, y del sistema (1.3.1) coinciden en un punto, entonces $x \equiv y$. En particular, si $x(t) = 0$ para algún $t \in [a,b]$, entonces $x \equiv 0$.

3. Existe una matriz $X(t)$ de dimensión $n \times n$ cuyas columnas $x_1(t), \ldots, x_n(t)$ son soluciones linealmente independientes del sistema (1.3.1). Más aún, podemos elegir $t_0 \in [a,b]$ de forma que $X(t_0) = I_n$ donde I_n es la matriz identidad.

A la matriz $X(t)$ del inciso (3) se le llama la matriz fundamental del sistema (1.3.1) basada en t_0.

Capítulo 2

Normalidad y Regularidad

En este capítulo enunciaremos condiciones de primer y segundo orden para el problema $P(S)$ bajo distintas hipótesis de regularidad y normalidad. Además veremos que en gran parte de bibliografía, las hipótesis bajo las que se cumplen estas condiciones son equivalentes a normalidad fuerte.

2.1. Independencia Lineal con Restricciones de Igualdad

En esta sección consideramos el caso donde $A = \emptyset$ y $B = \{1, \ldots, m\}$.

Definición 2.1.1. *Dado $x_0 \in S$ definimos los siguientes conjuntos.*

1. El conjunto de restricciones tangenciales en x_0 por

$$R_S(x_0) = \{y \in \mathbf{Y} \mid I'_\beta(x_0; y) = 0 \ (\beta \in B)\}.$$

2. El conjunto de tangentes curvilíneos en x_0 por

$$\begin{aligned} C_S(x_0) = \{y \in \mathbf{Y} \mid \exists[\delta > 0,\ x(\cdot, \epsilon) \in S \ (|\epsilon| < \delta)] \ni x_t(\cdot,\cdot) \ \ x_\epsilon(\cdot,\cdot) \ \textit{y} \ x_{t\epsilon}(\cdot,\cdot) \ \textit{son continuas a trozos} \\ \textit{en } t,\ x(t, 0) = x_0(t) \ \textit{y} \ x_\epsilon(t, 0) = y(t) \ (t \in [a, b])\}. \end{aligned}$$

En el inciso (2) de la definición anterior y en adelante, ser continua a trozos en t significará que existe una partición finita $T_i = [t_{i-1}, t_i]$ $(i = 1, \ldots, N)$ de $[a, b]$ de forma que $t_0 = a,\ t_N = b$ y la función es continua en cada conjunto $T_i \times [-\delta, \delta]$.

Lema 2.1.2. $C_S(x_0) \subset R_S(x_0)$.

Demostración. Si $y \in C_S(x_0),\ \delta > 0$ y $x(\cdot, \epsilon)$ son como en la definición anterior, entonces $I_\beta(x(\cdot, \epsilon)) = 0$ para todo $|\epsilon| < \delta$ y $\beta \in B$. Derivando esta igualdad obtenemos

$$\begin{aligned}
0 &= \frac{d}{d\epsilon} I_\beta(x(\cdot, \epsilon)) = \frac{d}{d\epsilon} \int_a^b f_\beta(t, x(t, \epsilon), \dot{x}(t, \epsilon))dt \\
&= \frac{d}{d\epsilon} \left[\sum_{i=1}^N \int_{t_{i-1}}^{t_i} f_\beta(t, x(t, \epsilon), \dot{x}(t, \epsilon))dt \right] \\
&= \sum_{i=1}^N \frac{d}{d\epsilon} \int_{t_{i-1}}^{t_i} f_\beta(t, x(t, \epsilon), \dot{x}(t, \epsilon))dt \\
&= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \frac{\partial}{\partial\epsilon} f_\beta(t, x(t, \epsilon), \dot{x}(t, \epsilon))dt \\
&= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \{f_{\beta x}(t, x(t, \epsilon), \dot{x}(t, \epsilon))x_\epsilon(t, \epsilon) + f_{\beta \dot{x}}(t, x(t, \epsilon), \dot{x}(t, \epsilon))\dot{x}_\epsilon(t, \epsilon)\}dt \\
&= \int_a^b \{f_{\beta x}(t, x(t, \epsilon), \dot{x}(t, \epsilon))x_\epsilon(t, \epsilon) + f_{\beta \dot{x}}(t, x(t, \epsilon), \dot{x}(t, \epsilon))\dot{x}_\epsilon(t, \epsilon)\}dt \\
&= I'_\beta(x(t, \epsilon); x_\epsilon(t, \epsilon)),
\end{aligned}$$

por lo que evaluando en $\epsilon = 0$ obtenemos $I'_\beta(x_0; y) = 0$, que es lo que queríamos demostrar. □

Nota 2.1.3. En la demostración anterior para concluir que $I'_\beta(x_0; y) = 0$ se usó implícitamente el hecho que las derivadas cruzadas $x_{t\epsilon}$ y $x_{\epsilon t}$ son iguales. Esto es posible gracias a las hipótesis sobre las funciones $x_t,\ x_\epsilon$ y $x_{t\epsilon}$ (ver [25, p. 235, Teorema 9.41]). Del mismo modo se justifica el intercambio de la derivada con la integral en la cuarta igualdad (ver [11] o [25]). Será importante tener en cuenta esto en las demostraciones parecidas que aparezcan más adelante.

Definición 2.1.4. *Diremos que x_0 es c-regular si $C_S(x_0) = R_S(x_0)$.*

Como se ha mencionado anteriormente, probar regularidad es complicado en la práctica por lo que requerimos de una noción más sencilla de verificar que implique esta propiedad. El siguiente es uno de estos criterios.

Definición 2.1.5. *Un arco $x_0 \in S$ es llamado normal fuerte o f-normal si las primeras variaciones $I'_\beta(x_0; \cdot)$ $(\beta \in B)$ son linealmente independientes en $\mathbf{Y}$.*

Por el Lema 1.3.1, f-normalidad es equivalente a la existencia de variaciones $y_1, \ldots, y_m \in \mathbf{Y}$ de forma que el determinante $|I'_\beta(x_0; y_i)|$ $(\beta, i \in B)$ no se anula.

Teorema 2.1.6. *Si $x_0 \in S$ es un arco f-normal, entonces x_0 es c-regular.*

Demostración. Supongamos que $x_0 \in S$ es normal. Por el Lema 2.1.2 solo falta probar que $R_S(x_0) \subset C_S(x_0)$. Sea $y \in R_S(x_0)$ y tomemos $y_1, \ldots, y_m \in \mathbf{Y}$ de forma que el determinante $|I'_\beta(x_0; y_i)|$ $(\beta, i \in B)$ no se anula. Ahora definimos $G = (G_1, \ldots, G_m) : \mathbf{R} \times \mathbf{R}^m \to \mathbf{R}^m$ por

$$G_\beta(\epsilon, c) = I_\beta\left(x_0 + \epsilon y + \sum_{\gamma=1}^{m} c_\gamma y_\gamma\right) \quad (\beta \in B).$$

Notemos que $G(0,0) = 0$ y $|G_c(0,0)| = |I'_\beta(x_0; y_i)| \neq 0$. Por el Teorema 1.3.6, existe $\delta > 0$ y $c : [-\delta, \delta] \to \mathbf{R}^m$ función de clase C^2 con $c(0) = 0$ y $G(\epsilon, c(\epsilon)) = 0$ $(|\epsilon| < \delta)$. Diferenciando esta última igualdad con respecto a ϵ y evaluando en $\epsilon = 0$ obtenemos

$$0 = G_\epsilon(0,0) + G_c(0,0)c'(0) = G_c(0,0)c'(0),$$

y por lo tanto $c'(0) = 0$. Definiendo $x(t, \epsilon) := x_0(t) + \epsilon y(t) + \sum_{i=1}^{m} c_i(\epsilon) y_i(t)$ obtenemos que $y \in C_S(x_0)$. □

Nota 2.1.7. En la demostración anterior $x(t, \epsilon)$ tiene primera y segunda derivada continua $x_\epsilon(t, \epsilon), x_{\epsilon\epsilon}(t, \epsilon)$ con respecto de ϵ que definen arcos en $\mathbf{Y}$ para $|\epsilon| < \delta$. Además las parciales cumplen las propiedades de continuidad del inciso (2) de la Definición 2.1.1.

La derivación de condiciones de primer y segundo orden con hipótesis de normalidad fuerte de la solución del problema se puede hacer vía el Lema 1.3.2.

Teorema 2.1.8. *Supongamos que $x_0 \in S$ es f-normal y resuelve $P(S)$ localmente. Entonces existen multiplicadores únicos $\lambda \in \mathbf{R}^m$ de forma que $(x_0, \lambda) \in \mathcal{E}$. Más aún $J''(x_0; y) \geq 0$ para todo $y \in R_S(x_0)$.*

Demostración. Sea $y_0 \in R_S(x_0)$, $\delta > 0$ y $x(\cdot, \epsilon) \in S$ $(|\epsilon| < \delta)$ con $x(t, 0) = x_0(t)$ y $x_\epsilon(t, 0) = y_0(t)$ $(t \in [a, b])$. Si definimos $g(\epsilon) := I(x(\cdot, \epsilon))$, entonces g tiene un mínimo local en $\epsilon = 0$, por lo que $g'(0) = I'(x_0; y_0) = 0$. Por el Lema 1.3.2, existe $\lambda \in \mathbf{R}^m$ de forma que

$$I'(x_0; y) + \sum_{\gamma=1}^{m} \lambda_\gamma I'_\gamma(x_0; y) = 0 \text{ para todo } y \in \mathbf{Y}.$$

Por lo tanto si definimos J como en la definición del conjunto $\mathcal{E}$ de la Sección 1.2.2, $J'(x_0; y) = 0$ para todo $y \in \mathbf{Y}$. Para demostrar la segunda conclusión, observemos que, por las igualdades $I_\gamma(x(\cdot, \epsilon)) = 0$ podemos escribir

$$g(\epsilon) = J(x(\cdot, \epsilon)) = I(x(\cdot, \epsilon)) + \sum_{\gamma=1}^{m} \lambda_\gamma I_\gamma(x(\cdot, \epsilon)),$$

de donde $g'(\epsilon) = J'(x(\cdot,\epsilon); x_\epsilon(\cdot,\epsilon))$. Como $\epsilon = 0$ minimiza g localmente y $x_{\epsilon\epsilon}(\cdot,\epsilon) \in \mathbf{Y}$ tenemos

$$0 \leq g''(0) = J''(x_0; y_0) + J'(x_0; x_{\epsilon\epsilon}(\cdot, 0)) = J''(x_0; y_0). \qquad \square$$

Nota 2.1.9. Sean x_0, λ y J como en el teorema anterior. Entonces $J'(x_0; y) = 0$ para todo $y \in \mathbf{Y}$ si y solo si existe $c \in \mathbf{R}^n$ de forma que

$$F_{\dot{x}}(t, x(t), \dot{x}(t)) = \int_a^t F_x(s, x(s), \dot{x}(s))ds + c \quad (t \in [a,b]),$$

donde $F = f + \sum \lambda_\gamma f_\gamma$ (ver [18, p. 69]). Esta es la forma integral de la ecuación de Euler

$$\frac{d}{dt} F_{\dot{x}}(\tilde{x}(t)) = F_x(\tilde{x}(t)) \quad (t \in [a,b]).$$

Si $\dot{x}$ tiene una discontinuidad, d/dt se interpreta como la derivada por la derecha o por la izquierda, y esta fórmula se cumple aunque x no tenga segunda derivada. Es importante recordar que, si $J'(x_0; y) = 0$ para todo $y \in \mathbf{Y}$ entonces x_0 satisface la ecuación de Euler con respecto a F, pero a menos que la función $t \mapsto F_{\dot{x}}(\tilde{x}_0(t))$ $(t \in [a,b])$ esté en $\mathbf{X}$, el contrario puede no satisfacerse. Por ejemplo si $F(t,x,\dot{x}) = (\dot{x}^2 - x^2)/2, T = [0, 2\pi],\ x_0(t) := \operatorname{sen} t$ si $t \in [0,\pi]$ y $x_0(t) := 0$ si $t \in [\pi, 2\pi]$, entonces x_0 satisface la ecuación de Euler, pero no existe $c \in \mathbf{R}$ de forma que $\dot{x}_0(t) = \int_0^t -x_0(s)ds + c$ para todo $t \in T$.

2.2. Restricciones de Igualdad con Índices Activos

Consideremos ahora el caso general donde $A = \{1, \ldots, q\}$ y $B = \{q+1, \ldots, m\}$. En esta sección daremos condiciones necesarias de segundo orden para el problema $P(S)$, basándonos en los resultados solo con igualdades de la sección anterior.

Recordemos que el conjunto S_0 está definido por

$$S_0 = \{x \in \mathbf{X}_e \mid I_\alpha(x) = 0 \ (\alpha \in I_a(x_0) \cup B)\},$$

mientras que las restricciones tangenciales asociadas a este conjunto están dadas por

$$R_{S_0}(x_0) = \{y \in \mathbf{Y} \mid I'_\alpha(x_0; y) = 0 \ (\alpha \in I_a(x_0) \cup B)\}.$$

Lema 2.2.1. *Si $x_0 \in S$ resuelve $P(S)$ localmente, también resuelve $P(S_0)$ localmente.*

Demostración. Si $I_\alpha(x_0) < 0$, tomemos $\epsilon_\alpha > 0$ de modo que $I_\alpha(x) < 0$ para $\|x - x_0\| < \epsilon_\alpha$. Definamos $N := \{x \in \mathbf{X} \mid \|x - x_0\| < \epsilon\}$ donde $\epsilon = \text{mín}\{\epsilon_\alpha \mid I_\alpha(x_0) < 0\}$ si $I_a(x_0) \neq A$ o $N := \mathbf{X}$ en caso contrario. Como $S_0 \cap N \subset S$, x_0 también resuelve $P(S_0)$ localmente. $\square$

Como en el caso de $P(S_0)$ estamos considerando únicamente restricciones de igualdad, podemos derivar condiciones necesarias de segundo orden a partir del Teorema 2.1.8 y del lema anterior.

Teorema 2.2.2. *Supongamos que x_0 resuelve $P(S)$ localmente y existe $\lambda \in \mathbf{R}^m$ con $(x_0, \lambda) \in \mathcal{E}$. Si x_0 es f-normal con respecto de S_0 entonces*

$$J''(x_0; y) \geq 0 \textit{ para todo } y \in R_{S_0}(x_0).$$

Explícitamente, f-normalidad con respecto de S_0 quiere decir la independencia lineal de los operadores $I'_\alpha(x_0; \cdot)$ $(\alpha \in I_a(x_0) \cup B)$ en $\mathbf{Y}$, o equivalentemente por el Lema 1.3.1, la existencia de elementos $y_\gamma \in \mathbf{Y}$ que le permiten al determinante $|I'_\alpha(x_0; y_\gamma)|$ $(\alpha, \gamma \in I_a(x_0) \cup B)$ no anularse.

En la siguiente sección veremos cómo derivar condiciones de segundo orden que se satisfacen en un conjunto más grande que $R_{S_0}(x_0)$.

2.3. Independencia Lineal con Restricciones de Igualdad y Desigualdad

Para trabajar este caso requerimos la siguiente modificación de la definición del conjunto de tangentes curvilíneos.

Definición 2.3.1. *El conjunto de tangentes curvilíneos positivos de S en x_0 es el conjunto*

$$\begin{aligned} C_S^+(x_0) = \{y \in \mathbf{Y} \mid \exists[\delta > 0,\ x(\cdot, \epsilon) \in S\ (0 \leq \epsilon < \delta)] \ni x_t(\cdot, \cdot),\ x_\epsilon(\cdot, \cdot) \textit{ y } x_{t\epsilon}(\cdot, \cdot) \textit{ son continuas a trozos} \\ \textit{en } t,\ x(t, 0) = x_0(t) \textit{ y } x_\epsilon(t, 0) = y(t)\ (t \in [a, b])\}. \end{aligned}$$

La derivada con respecto a ϵ en la definición anterior es la derivada por la derecha.

Es importante notar que ahora necesitamos del conjunto de tangentes curvilíneos positivos en vez del conjunto de tangentes curvilíneos definidos en la sección anterior. Esto se debe a que, al demostrar que s-normalidad implica p-regularidad (ver Definición 2.3.3 y Teorema 2.3.5), tendremos que lidiar con la igualdad

$I_\gamma(x(\cdot,\epsilon)) = I_\gamma(x_0) + \epsilon I'_\gamma(x_0; y)$ con $y \in R_S(x_0)$, de donde podemos asegurar que $x(\cdot,\epsilon) \in S$ únicamente si $\epsilon \in [0,\delta)$.

Lema 2.3.2. *Si $x_0 \in S$, entonces $C_S^+(x_0) \subset R_S(x_0)$.*

Demostración. Si $y \in C_S^+(x_0), \delta > 0$ y $x(\cdot,\epsilon)$ son como en la definición anterior, entonces para cualquier $\alpha \in I_a(x_0)$ tenemos que $I_\alpha(x(\cdot,\epsilon)) \leq 0$ para $0 \leq \epsilon < \delta$, por lo que

$$0 \geq \frac{d}{d\epsilon} I_\alpha(x(\cdot,\epsilon))\Big|_{\epsilon=0} = \int_a^b \{f_{\alpha x}(\tilde{x}_0(t))y(t) + f_{\alpha\dot{x}}(\tilde{x}_0(t))\dot{y}(t)\}dt = I'_\alpha(x_0; y).$$

El Lema 2.1.2 nos da la igualdad que necesitamos para $I_\beta(x(\cdot,\epsilon))$ cuando $\beta \in B$. □

También modificamos la definición de regularidad de la sección anterior de forma natural.

Definición 2.3.3. *1. Decimos que $x_0 \in S$ es p-regular si $C_S^+(x_0) = R_S(x_0)$.*

2. El arco $x_0 \in S$ será llamado normal fuerte, o s-normal, si las primeras variaciones $I'_\gamma(x_0; y)$ $(\gamma \in I_a(x_0) \cup B)$ de I_γ a lo largo de x_0 son linealmente independientes en $\mathbf{Y}$.

A continuación demostramos que s-normalidad implica p-regularidad. La prueba está basada en el siguiente lema.

Lema 2.3.4. *Sea $x_0 \in S$ un arco s-normal. Si $y \in R_S(x_0)$, existen $\delta > 0$ y una familia $x(\cdot,\epsilon)$ $(0 \leq \epsilon \leq \delta)$ con las propiedades de la Definición 2.3.1. Además para $\gamma \in I_a(x_0) \cup B$ se satisface*

$$I_\gamma(x(\cdot,\epsilon)) = \epsilon I'_\gamma(x_0; y).$$

Demostración. Por el Lema 1.3.1, existen $y_\sigma \in \mathbf{Y}$ $(\sigma \in I_a(x_0) \cup B)$ de forma que $|I'_\gamma(x_0; y_\sigma)| \neq 0$ $(\gamma, \sigma \in I_a(x_0) \cup B)$. Definamos la función

$$\bar{x}(t,\epsilon,\alpha) = x_0(t) + \epsilon y(t) + \sum_\sigma \alpha_\sigma y_\sigma(t) \quad (t \in [a,b],\ \epsilon, \alpha_\sigma \in \mathbf{R}),$$

donde $\alpha = (\alpha_\sigma :\ \sigma \in I_a(x_0) \cup B)$, y consideremos el sistema de ecuaciones

$$g^\gamma(\epsilon,\alpha) := I_\gamma(\bar{x}(\cdot,\epsilon,\alpha)) - \epsilon I'_\gamma(x_0; y) = 0 \quad (\gamma \in I_a(x_0) \cup B).$$

Notemos que en $(\epsilon,\alpha) = (0,0)$ las ecuaciones se satisfacen y, además

$$\left| \frac{\partial g^\gamma}{\partial \alpha_\sigma}(0,0) \right| = |I'_\gamma(x_0; y_\sigma)| \neq 0 \quad (\gamma, \sigma \in I_a(x_0) \cup B).$$

Por el Teorema 1.3.6, existen $\delta > 0$ y funciones continuas $\alpha_\sigma : [-\delta, \delta] \to \mathbf{R}$ ($\sigma \in I_a(x_0) \cup B$) de clase C^2 con $\alpha_\sigma(0) = 0$ y $g(\epsilon, \alpha(\epsilon)) = 0$ ($|\epsilon| < \delta$), donde $g = (g^\gamma : \ \gamma \in I_a(x_0) \cup B)$. Derivando esta última igualdad con respecto de ϵ y evaluando en $\epsilon = 0$ obtenemos

$$0 = g_\epsilon(0,0) + g_\alpha(0,0)\alpha'(0) = g_\alpha(0,0)\alpha_\epsilon(0),$$

de donde $\alpha_\epsilon(0) = 0$. Con esto podemos concluir que $x(t,\epsilon) := \bar{x}(t, \epsilon, \alpha(\epsilon))$ satisface las propiedades deseadas para $0 \leq \epsilon \leq \delta$. $\square$

Teorema 2.3.5. *s-normalidad implica p-regularidad.*

Demostración. Supongamos que x_0 es s-normal. Queremos probar que $R_S(x_0) \subset C_S^+(x_0)$. Tomemos $y \in R_S(x_0)$ y $\delta > 0$, $x(\cdot, \epsilon) \in \mathbf{X}$ ($0 < \epsilon \leq \delta$) con las propiedades del lema anterior. Notemos que por la definición de $x(t,\epsilon)$, $x(\cdot,\epsilon)$ converge uniformemente a x_0 en $[a,b]$ cuando ϵ tiende a cero, por lo que de ser necesario, por la continuidad de los operadores podemos disminuir $\delta > 0$ para que la desigualdad $I_\gamma(x(\cdot,\epsilon)) < 0$ se mantenga en $0 \leq \epsilon \leq \delta$ para los γ que cumplan $I_\gamma(x_0) < 0$. Esto claramente implica la contención deseada. $\square$

Para demostrar las condiciones necesarias de primer y segundo orden usaremos el Lema 1.3.3 (que claramente es una generalización del Lema 1.3.2) y el siguiente resultado.

Lema 2.3.6. *Sea $x_0 \in S_1(\lambda)$. Entonces x_0 es s-normal relativo a S si y solo si es s-normal relativo a $S_1(\lambda)$.*

Demostración. Definamos los conjuntos

$$\bar{I}_a(\lambda) = \{\alpha \in \Gamma_0(\lambda) \mid I_\alpha(x_0) = 0 \} = \{\alpha \in I_a(x_0) \mid \lambda_\alpha = 0\}, \quad \bar{B} = \Gamma_+(\lambda) \cup B.$$

Entonces x_0 es s-normal de $S_1(\lambda)$ si y solo si las primeras variaciones $\{I'_\gamma(x_0; \cdot) \mid \gamma \in \bar{I}_a(x_0) \cup \bar{B}\}$ son linealmente independientes en $\mathbf{Y}$. Por lo tanto basta mostrar que $I_a(x_0) \cup B = \bar{I}_a(x_0, \lambda) \cup \bar{B}$, pero como $I_a(x_0) = \bar{I}_a(x_0, \lambda) \cup \Gamma_+(\lambda)$ tenemos que

$$\bar{I}_a(x_0, \lambda) \cup \bar{B} = \bar{I}_a(x_0, \lambda) \cup \Gamma_+(\lambda) \cup B = I_a(x_0) \cup B. \qquad \square$$

Teorema 2.3.7. *Supongamos que x_0 es s-normal y resuelve $P(S)$ localmente. Entonces existe $\lambda \in \mathbf{R}^m$ de forma que $(x_0, \lambda) \in \mathcal{E}$. Más aún, $J''(x_0, y) \geq 0$ para todo $y \in R_{S_1}(x_0)$.*

Demostración. Sea $y_0 \in R_S(x_0) = C_S^+(x_0)$ y $\delta > 0, x(\cdot,\epsilon)$ como en la Definición 2.3.1. Como la función $g(\epsilon) := I(x(\cdot,\epsilon))$ tiene un mínimo local en $\epsilon = 0$, tenemos que $0 \le g'(0) = I'(x_0; y_0)$ (se tiene desigualdad y no igualdad ya que g sólo está definida en $[0,\delta)$). Por el Lema 1.3.3, existen λ_γ ($\gamma \in I_a(x_0) \cup B$) de forma que $\lambda_\alpha \ge 0$ ($\alpha \in I_a(x_0)$) e

$$I'(x_0; y) + \sum_{\gamma=1}^{m} \lambda_\gamma I'_\gamma(x_0; y) = 0 \quad \text{para todo } y \in \mathbf{Y},$$

donde definimos $\lambda_\gamma = 0$ para $\gamma \in A - I_a(x_0)$. Es decir definiendo J como antes tendremos que $J'(x_0; y) = 0$ para todo $y \in \mathbf{Y}$. Para demostrar la segunda conclusión tomemos $y_1 \in R_{S_1}(x_0)$. Como x_0 es s-normal relativo a S, también es s-normal relativo a $S_1(\lambda)$ y por lo tanto p-regular con respecto a $S_1(\lambda)$. Entonces $y_1 \in C_{S_1}^+(x_0)$ por lo que existe $\delta > 0, x(\cdot,\epsilon) \in S_1(\lambda)$ para $0 \le \epsilon < \delta$ como en la definición. Como antes, si $g_1(\epsilon) := I(x(\cdot,\epsilon))$, entonces g_1 tiene un mínimo local en $\epsilon = 0$ y $g_1'(0) = I'(x_0; y_1) \ge 0$.

Nuevamente por el Lema 1.3.3, existe $\mu \in \mathbf{R}^m$ con $\mu_\alpha \ge 0$ ($\alpha \in \Gamma_0(\lambda)$) tal que

$$I'(x_0; y) + \sum_{\gamma=1}^{m} \mu_\gamma I'_\gamma(x_0; y) = 0 \quad \text{para todo } y \in \mathbf{Y},$$

donde $\mu_\gamma = 0$ para $\gamma \in A - \Gamma_+(\lambda)$. Por unicidad $\mu_\gamma = \lambda_\gamma$. Como $I(x(\cdot,\epsilon)) = J(x(\cdot,\epsilon))$ podemos escribir $g_1(\epsilon) = J(x(\cdot,\epsilon))$. Entonces $g_1'(\epsilon) = J'(x(\cdot,\epsilon); x_\epsilon(\cdot,\epsilon))$ y como $\epsilon = 0$ es un mínimo local de g_1 y $x_{\epsilon\epsilon}(\cdot,\epsilon) \in \mathbf{Y}$ tenemos

$$0 \le g_1'(0) = J''(x_0; y_1) + J'(x_0; x_{\epsilon\epsilon}(\cdot,0)) = J''(x_0; y_1). \qquad \square$$

Para cerrar esta sección, notemos que s-normalidad con respecto de S es lo mismo que f-normalidad con respecto de S_0. Por lo tanto el teorema anterior representa una mejoría del Teorema 2.2.2 ya que $R_{S_0}(x_0) \subset R_{S_1}(x_0)$.

2.4. El Método del Determinante

El objetivo de esta sección es derivar condiciones de segundo orden usando las técnicas de [18, p. 283-286] para problemas de control con restricciones integrales y ver que la noción de normalidad definida con este método es equivalente a la de normalidad fuerte. A grandes rasgos, el método consiste en derivar condiciones de segundo orden para el conjunto $C_{S_1}^+(x_0)$ y asumir una hipótesis adecuada que implique p-regularidad para obtener de nuevo el Teorema 2.3.6.

Teorema 2.4.1. *Supongamos que* $(x_0, \lambda) \in \mathcal{E}$. *Si* x_0 *resuelve* $P(S)$ *localmente tendremos que* $J''(x_0; y) \geq 0$ *para todo* $y \in C^+_{S_1}(x_0)$.

Demostración. Definamos $p(t) := F^*_{\dot{x}}(\tilde{x}_0(t))$ $(t \in [a, b])$, donde $F = f + \sum_{\gamma=1}^m \lambda_\gamma f_\gamma$. Por la Nota 2.1.9 tenemos que $\dot{p}(t) = F^*_x(\tilde{x}_0(t))$ $(t \in [a, b])$. Sea $G(t, x, \dot{x}) := F(t, x, \dot{x}) - \langle p(t), \dot{x}\rangle - \langle \dot{p}(t), x\rangle$ y definamos

$$K(x) := \langle p(b), X_b\rangle - \langle p(a), X_a\rangle + \int_a^b G(t, x(t), \dot{x}(t))dt.$$

Notemos que $K(x) = J(x)$ para todo x con $x(a) = X_a$ y $x(b) = X_b$, por lo que

$$S_1(\lambda) = \{x \in S \mid K(x) = I(x)\}.$$

Como $x_0 \in S_1(\lambda)$, x_0 minimiza $K(x)$ en $S_1(\lambda)$. Ahora sea $y \in C^+_{S_1}(x_0)$ y $\delta > 0$ de forma que existe $x(\cdot, \epsilon) \in S_1(\lambda)$ para cada $0 \leq \epsilon < \delta$ tal que $x(t, 0) = x_0(t)$ y $x_\epsilon(t, 0) = y(t)$ $(t \in [a, b])$. Entonces si $g(\epsilon) := K(x(\cdot, \epsilon))$ $(0 \leq \epsilon < \delta)$ se satisface que

$$g(\epsilon) = I(x(\cdot, \epsilon)) \geq I(x_0) = K(x_0) = g(0) \quad (0 \leq \epsilon < \delta).$$

Como para $t \in [a, b]$ se cumplen las igualdades

$$G_x(\tilde{x}_0(t)) = F_x(\tilde{x}_0(t)) - \dot{p}^*(t) = 0 \;\text{ y }\; G_{\dot{x}}(\tilde{x}_0(t)) = F_{\dot{x}}(\tilde{x}_0(t)) - p^*(t) = 0$$

tendremos que

$$g'(0) = K'(x_0; y) = \int_a^b \{G_x(\tilde{x}_0(t))y(t) + G_{\dot{x}}(\tilde{x}_0(t))\dot{y}(t)\}dt = 0$$

lo que implica que $g''(0) \geq 0$. El resultado se sigue de que $g''(0) = K''(x_0; y) = J''(x_0; y)$. $\square$

Nota 2.4.2. En lo que sigue supondremos sin pérdida de generalidad que todas las restricciones son activas con respecto a la solución x_0 de $P(S)$. Esto se puede hacer debido a que solo estamos interesados en arcos admisibles en una vecindad de x_0, por lo que si una restricción es inactiva, i.e. $I_\alpha(x_0) < 0$, existirá una vecindad suficientemente pequeña de x_0 en donde cualquier arco también satisface esta desigualdad.

La siguiente definición aparece en [18, p. 273].

Definición 2.4.3. *Diremos que* $x_0 \in S$ *es v-normal si existen* $m + n$ *arcos* $y_\sigma \in \mathbf{X}$ *con* $y_\sigma(a) = 0$ $(\sigma = 1, \ldots, m + n)$ *y de forma que el determinante*

$$\Delta := \left| \begin{array}{c} I'_\gamma(x_0; y_\sigma) \\ y^i_\sigma(b) \end{array} \right| \begin{array}{c} (\gamma = 1, \ldots, m; i = 1, \ldots, n) \\ (\sigma = 1. \ldots, m+n) \end{array}$$

no se anula (γ e i son índices para las filas mientras que σ es índice para las columnas).

Teorema 2.4.4. *v-normalidad implica p-regularidad.*

Demostración. Supongamos que $x_0 \in S$ es v-normal. Únicamente necesitamos probar que $R_S(x_0) \subset C^+_S(x_0)$. Sea $y \in R_S(x_0)$, $N = m + n$ y seleccionamos N arcos $y_\sigma \in \mathbf{X}$ como en la definición anterior.

Definamos

$$\bar{x}(t, \epsilon, \alpha) := x_0(t) + \epsilon y(t) + \sum_{\sigma=1}^{N} \alpha_\sigma y_\sigma(t) \quad (t \in [a, b], \ \epsilon, \alpha_\sigma \in \mathbf{R}),$$

donde $\alpha = (\alpha_1, \ldots, \alpha_N)$. Notemos que $\bar{x}(t, 0, 0) = x_0(t)$ y

$$\bar{x}_\epsilon(t, 0, 0) = y(t), \quad \bar{x}_{\alpha_\sigma}(t, 0, 0) = y_\sigma(t) \quad (t \in [a, b]).$$

Definamos a $g : \mathbf{R} \times \mathbf{R}^N \to \mathbf{R}^N$ por sus componentes mediante las fórmulas

$$g^\gamma(\epsilon, \alpha) := I_\gamma(\bar{x}(\cdot, \epsilon, \alpha)) - \epsilon I'_\gamma(x_0; y) \quad (\gamma = 1, \ldots, m)$$
$$g^{m+i}(\epsilon, \alpha) := \bar{x}^i(b, \epsilon, \alpha) - X^i_b \quad (i = 1, \ldots, n).$$

Notemos que, como todos los índices los suponemos activos, $g(0, 0) = 0$ y $|g_\alpha(0, 0)| = \Delta \neq 0$. Por el Teorema 1.3.6, $\exists \delta > 0$ y $\beta : [-\delta, \delta] \to \mathbf{R}^N$ de clase C^2 de forma que $\beta(0) = 0$ y $g(\epsilon, \beta(\epsilon)) = 0$ ($|\epsilon| < \delta$). Derivando esta identidad con respecto de ϵ y evaluando en $\epsilon = 0$ obtenemos

$$0 = g_\epsilon(0, 0) + g_\alpha(0, 0)\beta'(0) = g_\alpha(0, 0)\beta'(0),$$

por lo que $\beta'(0) = 0$. Claramente el arco $x(t, \epsilon) := \bar{x}(t, \epsilon, \beta(\epsilon))$ satisface que $x(\cdot, \epsilon) \in S$ $(0 \leq \epsilon < \delta)$, $x(t, 0) = x_0(t)$ y $x_\epsilon(t, 0) = y(t)$ $(t \in [a, b])$ y se cumplen los requisitos de continuidad sobre las parciales, por lo que $y \in C^+_S(x_0)$. $\square$

Nota 2.4.5. Observemos que v-normalidad está definida en términos de las primeras variaciones y no diferencia los índices de A de los de B. Por lo tanto, al igual que con s-normalidad, esta noción no variará si se aplica a los conjuntos S, S_0 o S_1. Combinando este hecho con el teorema anterior se demuestra el siguiente resultado.

Teorema 2.4.6. *Supongamos que* $(x_0, \lambda) \in \mathcal{E}$. *Si* x_0 *es solución de* $P(S)$ *y es* v*-normal, entonces* $J''(x_0; y) \geq 0$ *para todo* $y \in R_{S_1}(x_0)$.

A continuación demostraremos que las nociones de v-normalidad y s-normalidad son equivalentes. Primero veremos que s-normalidad implica v-normalidad y después usaremos una noción auxiliar para dar otra caracterización de v-normalidad que nos permitirá demostrar la implicación contraria.

Teorema 2.4.7. *Sea* $x_0 \in S$. *Si* x_0 *es* s*-normal entonces es* v*-normal.*

Demostración. Supongamos que x_0 es s-normal. Entonces existen $y_\sigma \in \mathbf{Y}$ $(\sigma = 1, \ldots, m)$ de manera que $|M| \neq 0$, donde $M = (I'_\gamma(x_0; y_\sigma))$ $(\gamma, \sigma = 1, \ldots, m)$. Para $i, j = 1, \ldots, n$ definimos

$$y^i_{m+j}(t) = \begin{cases} (t-a)/(b-a) & \text{si } i = j \\ 0 & \text{de otra forma.} \end{cases}$$

Entonces el determinante

$$\Delta = \begin{vmatrix} I'_\gamma(x_0; y_\sigma) & I'_\gamma(x_0; y_{m+i}) \\ y^i_\sigma(b) & y^i_{m+i}(b) \end{vmatrix} = \begin{vmatrix} M & I'_\gamma(x_0; y_{m+i}) \\ 0 & I_{n\times n} \end{vmatrix}$$

$(\gamma, \sigma = 1, \ldots, m; i = 1, \ldots, n)$ es distinto de cero y x_0 es v-normal. $\square$

Ahora daremos la noción auxiliar de la que hablábamos en el párrafo anterior al Teorema 2.4.7, esta aparece en [18, p. 279] y es una definición más de normalidad, pero ahora en lugar de aplicarse a un problema de optimización, se aplica a una ecuación diferencial ordinaria. Supongamos que f es un mapeo de clase C^1 de $[a, b] \times \mathbf{R}^n \times \mathbf{R}^p$ a $\mathbf{R}^n$.

Definición 2.4.8. *Sea* (x_0, u_0) *es una solución de la ecuación diferencial*

$$\dot{x}(t) = f(t, x(t), u(t)) \quad (t \in [a, b]),$$

y definamos $A(t) := f_x(t, x_0(t), u_0(t))$ *y* $B(t) := f_u(t, x_0(t), u_0(t))$. *Diremos que* (x_0, u_0) *es* d*-normal si no existe solución no nula de la ecuación adjunta*

$$\dot{z}(t) = -A^*(t)z(t), \quad B^*(t)z(t) = 0 \quad (t \in [a, b]).$$

Lema 2.4.9. *Los siguientes enunciados son equivalentes.*

1. (x_0, u_0) *es d-normal.*

2. *Existen n soluciones* (y_j, v_j) $(j = 1, \ldots, n)$ *de*

$$\dot{y}(t) = A(t)y(t) + B(t)v(t) \quad (t \in [a, b]), \quad y(a) = 0$$

de forma que la matriz $(y_j^i(b))_{ij}$*, cuyas columnas son los vectores* $y_j(b)$*, tiene rango* n*.*

Demostración. $(a) \Rightarrow (b)$: Por el inciso (3) del Teorema 1.3.7, existe la matriz $Z(t) \in \mathbf{R}^{n\times n}$ con

$$\dot{Z}(t) = -Z(t)A(t) \quad (t \in [a, b]), \quad Z(b) = I_n,$$

donde I_n es la matriz identidad correspondiente. Denotemos por $z_1(t), \ldots, z_n(t)$ los vectores fila de $Z(t)$. Entonces se cumple que $\dot{z}_i(t) = -A^*(t)z_i(t)$ $(t \in [a, b],\ i = 1, \ldots, n)$. Si definimos $v_i(t) := B^*(t)z_i(t)$ (i=1,. . . ,n), las funciones $v_1, \ldots, v_n$ serán linealmente independientes, ya que de lo contrario existirán constantes $c_1, \ldots, c_n$, no todas cero de modo que

$$0 = \sum_{i=1}^{n} c_i v_i(t) = \sum_{i=1}^{n} c_i B^*(t) z_i(t) \quad \text{para todo } t \in [a, b].$$

Por el inciso (1) del Teorema 1.3.7, la función $z(t) := \sum_{i=1}^{n} c_i z_i(t)$ sería una solución no nula de la ecuación adjunta, lo que contradice la d-normalidad de (x_0, u_0). Ahora sean y_j $(j = 1, \ldots, n)$ soluciones de

$$\dot{y}(t) = A(t)y(t) + B(t)v_j(t) \quad (t \in [a, b]), \quad y(a) = 0.$$

Notemos que

$$\frac{d}{dt}\langle z_i(t); y_j(t)\rangle = z_i^*(t)[A(t)y_j(t) + B(t)v_j(t)] - z_i^*(t)A(t)y_j(t) = z_i^*(t)B(t)v_j(t),$$

por lo que

$$\langle z_i(b); y_j(b)\rangle = \int_a^b \langle v_i(t); v_j(t)\rangle dt \quad (i, j = 1, \ldots, n).$$

Podemos reescribir esta ecuación en notación matricial como

$$Z(b)(y_j^i(b))_{ij} = \left(\int_a^b \langle v_i(t); v_j(t)\rangle dt\right)_{ij}. \tag{2.4.1}$$

Como los vectores v_i son linealmente independientes, podemos aplicar el Lema 1.3.5 al espacio $\mathbf{X}$ con el producto interno usual

$$\langle x; y\rangle = \int_a^b \langle x(t); y(t)\rangle dt \quad (x, y \in \mathbf{X})$$

para observar que el lado derecho de la ecuación (2.4.1) es no singular. Además por el inciso (2) del Teorema 1.3.7, la matriz $Z(b)$ también es no singular. Por lo tanto, de la ecuación (2.4.1) podemos concluir que la matriz $(y_j^i(b))_{ij}$ es no singular.

$(b) \Rightarrow (a)$: Supongamos que z es solución de la ecuación adjunta

$$\dot{z}(t) = -A^*(t)z(t), \quad B^*(t)z(t) = 0 \quad (t \in [a, b]),$$

y tomemos y_j como en (b). Queremos demostrar que z es idénticamente nula. Para mostrar esto primero veamos que $\langle y_j(t); z(t)\rangle$ es constante para todo $t \in [a, b]$ y $j = 1, \ldots, n$. Esto se debe a que

$$\frac{d}{dt}\langle y_j(t); z(t)\rangle = z^*(t)B(t)v_j(t) = 0 \quad \text{para todo } t \in [a, b],\ j = 1, \ldots, n. \tag{2.4.2}$$

Por otro lado, como por hipótesis $y_j(a) = 0$, tendremos la igualdad

$$(\langle z(a); y_1(a)\rangle, \ldots, \langle z(a); y_n(a)\rangle) = z^*(a)(y_j^i(a))_{ij} = 0. \tag{2.4.3}$$

De las ecuaciones (2.4.2) y (2.4.3) podemos notar que

$$\langle z(b); y_j(b)\rangle = 0 \quad \text{para } j = 1, \ldots, n$$

pero esto es equivalente a la igualdad matricial

$$z^*(b)(y_j^i(b))_{ij} = 0$$

y como la matriz $(y_j^i(b))_{ij}$ es no singular tendremos que $z(b) = 0$. El resultado se sigue del inciso (2) del Teorema 1.3.7. □

Regresemos ahora al problema original $P(S)$ en cálculo de variaciones. Sea (w_0, u_0) solución de la ecuación diferencial

$$\dot{w}(t) = L(t, w(t), u(t)) \quad (t \in [a, b]),$$

donde $w = (x^1, \ldots, x^n, x^{n+1}, \ldots, x^{n+m})$, y $L : [a, b] \times \mathbf{R}^{n+m} \times \mathbf{R}^n \to \mathbf{R}^{n+m}$ está dada por

$$\begin{aligned} L^i(t, w, u) &= u^i \quad (i = 1, \ldots, n) \\ L^{n+\gamma}(t, w, u) &= f_\gamma(t, x^1, \ldots, x^n, u^1, \ldots, u^n) = f_\gamma(t, x, u) \quad (\gamma \in A \cup B). \end{aligned}$$

Entonces

$$\begin{aligned} \dot{x}_0^i(t) &= u_0^i(t) \quad (i = 1, \ldots, n, \ t \in [a, b]), \\ \dot{x}_0^{n+\gamma}(t) &= f_\gamma(t, x_0(t), \dot{x}_0(t)) \quad (\gamma \in A \cup B, \ t \in [a, b]). \end{aligned}$$

Sean $\bar{A}(t) = L_w(t, w_0(t), u_0(t))$, $\bar{B}(t) = L_u(t, w_0(t), u_0(t))$. Notemos que $\bar{A}(t)$ es la matriz de $(n+m)\times(n+m)$ dada por

$$\bar{A}(t) = \left(\frac{\partial L^i}{\partial w^j} \right) = \begin{pmatrix} 0 & 0 \\ C(t) & 0 \end{pmatrix} \quad (i, j = 1, \ldots, n + m),$$

donde

$$C(t) = \begin{pmatrix} f_{1x} \\ \vdots \\ f_{mx} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial x^1} & \cdots & \dfrac{\partial f_1}{\partial x^n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial x^1} & \cdots & \dfrac{\partial f_m}{\partial x^n} \end{pmatrix},$$

y $\bar{B}(t)$ es la matriz de $(n + m) \times n$ dada por

$$\bar{B}(t) = \left(\frac{\partial L^i}{\partial u^j} \right) = \begin{pmatrix} I_n \\ D(t) \end{pmatrix} \quad (i = 1, \ldots, n + m, \quad j = 1, \ldots, n),$$

donde

$$D(t) = \begin{pmatrix} f_{1u} \\ \vdots \\ f_{mu} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial f_1}{\partial u^1} & \cdots & \dfrac{\partial f_1}{\partial u^n} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial f_m}{\partial u^1} & \cdots & \dfrac{\partial f_m}{\partial u^n} \end{pmatrix},$$

donde todas las parciales de f están evaluadas en $(t, x_0(t), \dot{x}_0(t))$.

Por definición, (w_0, u_0) es d-normal si no existe solución no trivial de

$$\dot{z}(t) = -\bar{A}^*(t)z(t), \quad \bar{B}^*(t)z(t) = 0 \quad (t \in [a, b]).$$

Este sistema corresponde a

$$\dot{z}(t) = -\begin{pmatrix} 0 & C^*(t) \\ 0 & 0 \end{pmatrix} z(t), \quad (I_n, D^*(t))z(t) = 0,$$

es decir, para $i = 1, \ldots, n$ y $\gamma = 1, \ldots, m$

$$\dot{z}^i(t) = -\frac{\partial f_1}{\partial x^i} z^{n+1}(t) - \cdots - \frac{\partial f_m}{\partial x^i} z^{n+m}(t), \quad \dot{z}^{n+\gamma}(t) = 0,$$

$$0 = z^i(t) + \frac{\partial f_1}{\partial u^i} z^{n+1}(t) + \cdots + \frac{\partial f_m}{\partial u^i} z^{n+m}(t).$$

Haciendo $\lambda_\gamma = -z^{n+\gamma}(t)$ $(\gamma = 1, \ldots, m)$, se sigue que (w_0, u_0) es d-normal si $\lambda = 0$ es la única solución de

$$z(t) = \sum_{\gamma=1}^{m} \lambda_\gamma f^*_{\gamma\dot{x}}(\tilde{x}_0(t)), \quad \dot{z}(t) = \sum_{\gamma=1}^{m} \lambda_\gamma f^*_{\gamma x}(\tilde{x}_0(t)) \quad (t \in [a, b]).$$

Ahora, por definición tenemos que

$$I_\gamma(x) = \int_a^b f_\gamma(t, x(t), \dot{x}(t))dt,$$

así que, si definimos $F(t, x, \dot{x}) = \sum_{\gamma=1}^{m} \lambda_\gamma f_\gamma(t, x, \dot{x})$ entonces $J := \sum_{\gamma=1}^{m} \lambda_\gamma I_\gamma(x)$ satisface

$$J(x) = \int_a^b F(t, x(t), \dot{x}(t))dt.$$

Entonces

$$J'(x_0; y) = \sum_{\gamma=1}^{m} \lambda_\gamma I'_\gamma(x_0; y) = 0 \text{ para todo } y \in \mathbf{Y},$$

que es equivalente a la existencia de $c \in \mathbf{R}^n$ de forma que

$$F_{\dot{x}}(\tilde{x}_0(t)) = \int_a^t F_x(\tilde{x}_0(s))ds + c,$$

y que a su vez es equivalente a la existencia de $z \in \mathbf{X}$ que cumple

$$z(t) = F^*_{\dot{x}}(\tilde{x}_0(t)), \quad \dot{z}(t) = F^*_x(\tilde{x}_0(t)) \quad (t \in [a, b]).$$

Con esto hemos demostrado el siguiente Lema.

Lema 2.4.10. (w_0, u_0) *es d-normal si y solo si* x_0 *es normal relativo a* S_0*.*

El último resultado auxiliar que necesitaremos es el siguiente.

Lema 2.4.11. (w_0, u_0) *es d-normal si y solo si* x_0 *es v-normal.*

Demostración. Por el Lema 2.4.9, (w_0, u_0) es d-normal si y solo si $\exists n + m$ soluciones (w_j, v_j) de

$$\dot{w}(t) = \bar{A}(t)w(t) + \bar{B}(t)v(t), \quad w(a) = 0 \quad t \in [a, b],$$

cuya matriz $(w^i_j(b))_{ij}$ tiene rango $n+m$. Pero esto a su vez es equivalente a la existencia de $n+m$ soluciones y_j de

$$\dot{y}(t) = v(t), \quad y(a) = 0,$$

$$\dot{y}^{n+\gamma}(t) = f_{\gamma x}(\tilde{x}_0(t))y(t) + f_{\gamma \dot{x}}(\tilde{x}_0(t))\dot{y}(t), \quad y^{n+\gamma}(a) = 0,$$

cuya matriz $(y^i_j(b))_{ij}$ tiene rango $n + m$. Esta condición es precisamente v-normalidad de x_0. $\square$

Combinando los dos lemas anteriores tenemos el siguiente resultado.

Teorema 2.4.12. *Sea* $x_0 \in S$*. Entonces* x_0 *es v-normal si y solo si* x_0 *es s-normal.*

Hemos demostrado que las nociones de s-normalidad, v-normalidad, f-normalidad relativa a S_0 y d-normalidad para una ecuación diferencial apropiada son todas equivalentes. Mencionaremos sin demostración dos nociones adicionales también equivalentes a las anteriores. El método de Valentine [30] transforma el problema original a uno con solo restricciones de igualdad (ver [10] para una situación similar), y con la hipótesis de s-normalidad, las condiciones de segundo orden que ahí se derivan, se satisfacen en el conjunto $R_{S_0}(x_0)$. En [13], la hipótesis que genera condiciones de segundo orden con multiplicador de costo positivo es equivalente a f-normalidad con respecto de S_0, y la no negatividad se satisface en $R_{S_1}(x_0)$.

2.5. Ejemplo

Para concluir el capítulo, daremos un ejemplo que ilustra algunas de las características más importantes acerca de las condiciones de segundo orden que hemos discutido. Se pueden consultar más ejemplos en [24] y en las referencias contenidas en ese artículo.

Sea $a = 2\pi$ y consideremos el problema de minimizar

$$I(x) = \int_0^a \{t(\dot{x}^2(t) - x^2(t)) + x^3(t)\}dt$$

sujeto a $x(0) = x(a) = 0$ y $\int_0^a -x(t)dt \leq 0$.

Para este problema $n = 1,\ q = m = 1,\ X_a = 0 = X_b$,

$$f(t, x, \dot{x}) = t(\dot{x}^2 - x^2) + x^3 \quad \text{y} \quad f_1(t, x, \dot{x}) = -x.$$

El conjunto en el que estamos interesados en minimizar el funcional I es

$$S = \left\{ x \in \mathbf{X} \mid x(0) = x(a) = 0,\ I_1(x) = -\int_0^a x(t)dt \leq 0 \right\}.$$

Queremos saber si $x_0 \equiv 0$ es solución. Comencemos notando que x_0 es s-normal ya que existe $y_1 \in \mathbf{Y}$ tal que $|I_1'(x_0; y_1)| = |\int_0^a -y_1(t)dt| \neq 0$. Ahora, si $J := I + \lambda_1 I_1$ de forma que

$$J(x) = \int_0^a \{t(\dot{x}^2(t) - x^2(t)) + x^3(t) - \lambda_1 x(t)\}dt,$$

tendremos que

$$J'(x; y) = \int_0^a \{(3x^2(t) - 2tx(t) - \lambda_1)y(t) + 2t\dot{x}(t)\dot{y}(t)\}dt,$$

por lo tanto

$$J'(x_0;y) = \int_0^a -\lambda_1 y(t)dt.$$

Entonces, si $\lambda_1 = 0$ tendremos que $J'(x_0;y) = 0$ para todo $y \in \mathbf{Y}$. Observemos que en este caso

$$S_1(\lambda_1) = \{x \in S \mid J(x) = I(x)\} = S$$

y por lo tanto $R_{S_1}(x_0, \lambda_1) = R_S(x_0) = \{y \in \mathbf{Y} \mid \int_0^a y(t)dt \geq 0\}$. Ahora consideremos la segunda variación de J con respecto a x_0 que está dada por

$$J''(x_0;y) = \int_0^a 2t(\dot{y}^2(t) - y^2(t))dt.$$

Tomemos $c \in (\pi, 2\pi)$ y definamos

$$y(t) = \begin{cases} \operatorname{sen}(t) & \text{si } t \in [0, c/2] \\ \operatorname{sen}(c-t) & \text{si } t \in [c/2, c] \\ 0 & \text{de otra forma.} \end{cases}$$

Entonces $y \in R_{S_1}(x_0, \lambda_1)$ y tendremos que

$$\begin{aligned} J''(x_0;y) &= \int_0^{c/2} 2t(\cos^2 t - \operatorname{sen}^2 t)dt + \int_{c/2}^{c} 2t(\cos^2(c-t) - \operatorname{sen}^2(c-t))dt \\ &= 2c\int_0^{c/2} (\cos^2 t - \operatorname{sen}^2 t)dt = 2c\int_0^{c/2} \cos(2t)dt = c \operatorname{sen} c < 0 \end{aligned}$$

ya que $c \in (\pi, 2\pi)$. En vista del Teorema 2.3.6, podemos concluir que $x_0 \equiv 0$ no minimiza I en S.

Capítulo 3

Conos Tangentes y Normalidad Débil

En este capítulo se demostrará que los resultados del Capítulo 2 también pueden ser derivados por los métodos basados en conos tangentes para problemas de dimensión finita. Más aún, se verá que estos métodos son ideales para obtener una versión más general del Teorema 2.3.7.

3.1. Definiciones y Resultados Básicos

En la Sección 1.2 se hizo notar que $I'(x_0;\cdot)$ e $I''(x_0;\cdot)$ son la primera y segunda variación del funcional I en x_0 respecto a la norma débil. Esto quiere decir que se satisfacen las fórmulas de Taylor de primer y segundo orden

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

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

donde

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

Notemos que la definición en la Sección 1.1 de cono tangente en $\mathbf{R}^n$ se puede trasladar a cualquier espacio normado. A continuación recordamos esta definición.

Definición 3.1.1. *El cono tangente de $\mathcal{P} \subset \mathbf{X}$ en el punto $x_0 \in \mathcal{P}$ es el conjunto $T_{\mathcal{P}}(x_0)$ definido por*

$$T_{\mathcal{P}}(x_0) = \left\{ y \in \mathbf{X} \mid \exists [\{t_i\}_{i=1}^{\infty} \subset \mathbf{R}^+, \{x_i\}_{i=1}^{\infty} \subset \mathcal{P}] \ni t_i \to 0,\ \frac{x_i - x_0}{t_i} \to y \right\}.$$

Tomando a $\mathcal{P}$ como el conjunto de restricciones S con únicamente igualdades (i.e. $A = \emptyset$), podemos probar el siguiente lema.

Lema 3.1.2. *Si $x_0 \in S$, entonces $T_S(x_0) \subset R_S(x_0)$.*

Demostración. Sean $y \in T_S(x_0)$ y $\{t_i\}_{i=1}^{\infty}, \{x_i\}_{i=1}^{\infty}$ como en la Definición 3.1.1. Claramente $y \in \mathbf{Y}$. Por otro lado el resultado se sigue de la identidad

$$\lim_{i\to\infty} \frac{I_\gamma(x_i) - I_\gamma(x_0)}{t_i} = I'_\gamma(x_0; y), \tag{3.1.1}$$

que a su vez es consecuencia de la fórmula de Taylor antes mencionada. Como $x_i \in S$ tenemos que

$$0 = \lim_{i\to\infty} \frac{I_\gamma(x_i) - I_\gamma(x_0)}{t_i} = I'_\gamma(x_0; y). \qquad \square$$

Para la siguiente definición, S es el conjunto de restricciones general definido por igualdades y desigualdades.

Definición 3.1.3. *Diremos que $x_0 \in S$ es un punto regular de S si $T_S(x_0) = R_S(x_0)$.*

Por último, enunciamos dos propiedades básicas en el siguiente lema. Las demostraciones que aparecen en [19, C. 4, Sección 1] se pueden trasladar sin mayor dificultad a nuestro contexto en dimensión infinita.

Lema 3.1.4. *1. $T_{\mathcal{P}}(x_0)$ es un cono cerrado en $\mathbf{Y}$.*

2. Si $x_0 \in \mathcal{K} \subset \mathcal{P}$, entonces $T_{\mathcal{K}}(x_0) \subset T_{\mathcal{P}}(x_0)$.

3.2. Restricciones de Igualdad

Comencemos derivando las condiciones de primer y segundo orden para el caso de restricciones de igualdad (i.e. $A = \emptyset$ y $B = \{1, \ldots, m\}$). Recordemos que en este caso las restricciones tangenciales están dadas por el conjunto

$$R_S(x_0) = \{y \in \mathbf{Y} \mid I'_\beta(x_0; y) = 0\ (\beta \in B)\}.$$

Teorema 3.2.1. *Supongamos que x_0 es un arco regular de S y resuelve $P(S)$ localmente. Entonces existe $\lambda \in \mathbf{R}^m$ tal que $(x_0, \lambda) \in \mathcal{E}$.*

Demostración. Sea $y \in R_S(x_0) = T_S(x_0)$ y $\{t_i\}_{i=1}^{\infty}, \{x_i\}_{i=1}^{\infty}$ sucesiones como en la definición. Para valores grandes de i tenemos que $I(x_i) \geq I(x_0)$ por lo que

$$0 \leq \lim_{i\to\infty} \frac{I(x_i) - I(x_0)}{t_i} = I'(x_0; y).$$

Notemos además que si $y \in R_S(x_0)$ entonces $-y \in R_S(x_0)$, por lo que la desigualdad de arriba se convierte en igualdad. El resultado se sigue del Lema 1.3.2. $\square$

Teorema 3.2.2. *Supongamos que $x_0 \in S$ y existe $\lambda \in \mathbf{R}^m$ con $(x_0, \lambda) \in \mathcal{E}$. Si x_0 resuelve $P(S)$ localmente entonces $J''(x_0; y) \geq 0$ para todo $y \in T_S(x_0)$.*

Demostración. Sea $y \in T_S(x_0)$ y $\{t_i\}_{i=1}^{\infty}, \{x_i\}_{i=1}^{\infty}$ sucesiones como en la definición. De la fórmula de Taylor de segundo orden y del hecho que $J'(x_0; y) = 0$ para todo $y \in \mathbf{Y}$ tenemos la identidad

$$\lim_{i\to\infty} \frac{J(x_i) - J(x_0)}{t_i^2} = \frac{1}{2} J''(x_0; y). \tag{3.2.1}$$

Como $J(x) = I(x)$ para todo $x \in S$ y para valores grandes de i se cumple la desigualdad $J(x_i) \geq J(x_0)$, tendremos que

$$0 \leq \lim_{i\to\infty} \frac{J(x_i) - J(x_0)}{t_i^2} = \frac{1}{2} J''(x_0; y). \qquad \square$$

Nota 3.2.3. Si en el Teorema 3.2.2 además tenemos que x_0 es regular de S, tendremos que $J''(x_0; y) \geq 0$ para todo $y \in R_S(x_0)$.

Veamos que para un arco f-normal lo anterior se cumple. Para esto necesitamos el siguiente resultado.

Proposición 3.2.4. *Sea $x_0 \in S$ un arco f-normal de S. Entonces $C_S(x_0) \subset T_S(x_0)$.*

Demostración. Por la demostración del Teorema 2.1.6, si x_0 es un arco f-normal de S e $y \in C_S(x_0)$, entonces podemos tomar la función $x(\cdot, \epsilon) \in S$ $(|\epsilon| < \delta)$ que satisface $x(t, 0) = x_0(t)$ y $x_\epsilon(t, 0) = y(t)$ de la forma

$$x(t, \epsilon) := x_0(t) + \epsilon y(t) + \sum_{i=1}^{m} c_i(\epsilon) y_i(t)$$

donde los y_i son elementos en $\mathbf{Y}$ y $c = (c_1, \ldots, c_m) : [-\delta, \delta] \to \mathbf{R}^m$ satisface $c(0) = 0$ y $c'(0) = 0$. Definamos $x_k := x(\cdot, 1/k)$ y $t_k := 1/k$. Tenemos que demostrar que

$$\lim_{k\to\infty} \frac{x_k - x_0}{t_k} = y$$

en la norma débil. Pero esto es equivalente a demostrar que

$$\lim_{k\to\infty} \frac{x_k(t) - x_0(t)}{t_k} = y(t), \qquad \lim_{k\to\infty} \frac{\dot{x}_k(t) - \dot{x}_0(t)}{t_k} = \dot{y}(t).$$

y que la convergencia es uniforme en $[a, b]$. Para esto notemos que, como $y_i \in \mathbf{Y}$, existe $M > 0$ de forma que $|y_i(t)| < M$ para $i = 1, \dots, m$ y $t \in [a, b]$. Entonces

$$\left| \frac{x_k(t) - x_0(t)}{t_k} - y(t) \right| = \left| \frac{\sum_{i=1}^m c_i(1/k) y_i(t)}{1/k} \right| \leq \sum_{i=1}^m \left| \frac{c_i(1/k)}{1/k} \right| |y_i(t)| < \sum_{i=1}^m \left| \frac{c_i(1/k)}{1/k} \right| M$$

pero como

$$\lim_{k\to\infty} \left| \frac{c_i(1/k)}{1/k} \right| = \lim_{k\to\infty} \left| \frac{c_i(1/k) - c_i(0)}{1/k} \right| = |c_i'(0)| = 0,$$

dado cualquier $\eta > 0$ existe $N > 0$ de forma que si $k > N$

$$\left| \frac{c_i(1/k)}{1/k} \right| < \frac{\eta}{mM}$$

para $i = 1, \dots, m$. Entonces si $k > N$

$$\left| \frac{x_k(t) - x_0(t)}{t_k} - y(t) \right| < \sum_{i=1}^m \frac{\eta}{mM} M = \eta$$

para todo $t \in [a, b]$. De esto que la convergencia del primer límite es uniforme. Análogamente demostramos que la convergencia del segundo límite es uniforme. □

De este resultado podemos concluir que f-normalidad implica regularidad.

Teorema 3.2.5. *Sea $x_0 \in S$ un arco f-normal de S. Entonces x_0 es regular de S.*

Demostración. Por el Teorema 2.1.6, x_0 es c-regular de S. Por lo tanto

$$R_S(x_0) = C_S(x_0) \subset T_S(x_0),$$

por lo que x_0 es regular de S. □

Notemos que combinando los Teoremas 3.2.5, 3.2.1 y 3.2.2 regresamos al Teorema 2.1.8.

3.3. Restricciones de Igualdad y Desigualdad

Regresemos al caso general con igualdades y desigualdades (i.e $A = \{1, \ldots, q\}$, $B = \{q+1, \ldots, m\}$). Comencemos observando que la contención $T_S(x_0) \subset R_S(x_0)$ sigue siendo cierta.

Teorema 3.3.1. *Sea $x_0 \in S$. Entonces $T_S(x_0) \subset R_S(x_0)$.*

Demostración. Sea $y \in T_S(x_0)$ y $\{t_i\}_{i=1}^{\infty}, \{x_i\}_{i=1}^{\infty}$ sucesiones como en la Definición 3.1.1. De la ecuación (3.3.1) tenemos que si $\gamma \in I_a(x_0)$ entonces

$$0 \leq \lim_{i\to\infty} \frac{I_\gamma(x_i) - I_\gamma(x_0)}{t_i} = I'_\gamma(x_0; y),$$

mientras que si $\gamma \in B$

$$0 = \lim_{i\to\infty} \frac{I_\gamma(x_i) - I_\gamma(x_0)}{t_i} = I'_\gamma(x_0; y). \qquad \square$$

Como antes, llamaremos al punto $x_0 \in S$ regular de S si $T_S(x_0) = R_S(x_0)$.

Teorema 3.3.2. *Supongamos que $x_0 \in S$ es regular de S y resuelve $P(S)$ localmente. Entonces existe $\lambda \in \mathbf{R}^m$ con $(x_0, \lambda) \in \mathcal{E}$.*

Demostración. Sea $y \in T_S(x_0)$ y $\{t_i\}_{i=1}^{\infty}, \{x_i\}_{i=1}^{\infty}$ sucesiones como en la Definición 3.1.1. Para valores grandes de i tenemos que $I(x_i) \geq I(x_0)$. Por la ecuación (3.3.1)

$$0 \leq \lim_{i\to\infty} \frac{I(x_i) - I(x_0)}{t_i} = I'(x_0; y).$$

Por regularidad, $I'(x_0; y) \geq 0$ para todo $y \in R_S(x_0)$ por lo que el Lema 1.3.3 nos da el resultado deseado. $\square$

Teorema 3.3.3. *Supongamos que $x_0 \in S$ resuelve $P(S)$ localmente y existe $\lambda \in \mathbf{R}^m$ de forma que $(x_0, \lambda) \in \mathcal{E}$. Entonces $J''(x_0; y) \geq 0$ para todo $y \in T_{S_1}(x_0)$.*

Demostración. Sea $y \in T_{S_1}(x_0)$ y $\{t_i\}_{i=1}^{\infty}, \{x_i\}_{i=1}^{\infty}$ sucesiones como en la Definición 3.1.1. Como $I(x) = J(x)$ para todo $x \in S_1$ tendremos que $J(x_i) \geq J(x_0)$ para valores suficientemente grandes de i. De la fórmula de Taylor de segundo orden y del hecho que $J'(x_0; y) = 0$ para todo $y \in \mathbf{Y}$ obtenemos

$$0 \leq \lim_{i\to\infty} \frac{J(x_i) - J(x_0)}{t_i^2} = \frac{1}{2} J''(x_0; y). \qquad \square$$

Nota 3.3.4. Si en el Teorema 3.3.3 además tenemos que x_0 es regular de S_1, entonces $J''(x_0; y) \geq 0$ para todo $y \in R_{S_1}(x_0)$.

Ahora damos una definición de normalidad análoga a la del caso en dimensión finita.

Definición 3.3.5. *Diremos que un arco x_0 es normal relativo a S, si $\lambda = 0 \in \mathbf{R}^m$ es la única solución a*

1. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha I_\alpha(x_0) = 0$ $(\alpha \in A)$
2. $\sum_{\gamma=1}^m \lambda_\gamma I'_\lambda(x_0; y) = 0$ *para todo* $y \in \mathbf{Y}$.

El origen de esta definición proviene de la siguiente condición clásica de primer orden que podemos consultar en [18, p 263, Teorema 5.1].

Teorema 3.3.6. *Supongamos que x_0 resuelve $P(S)$ localmente. Entonces existe $\lambda_0 \geq 0$ y $\lambda_1, \ldots, \lambda_m$ no todos cero de forma que*

1. $\lambda_\alpha \geq 0$ *y* $\lambda_\alpha I_\alpha(x_0) = 0$ $(\alpha \in R)$.
2. $\exists c \in \mathbf{R}^n$ *tal que*

$$F_{\dot{x}}(\tilde{x}_0(t)) = \int_a^t F_x(\tilde{x}_0(s))ds + c \quad (t \in [a,b])$$

donde $F = \lambda_0 f + \sum_{\gamma=1}^m \lambda_\gamma f_\gamma$.

Como lo mencionamos anteriormente, la condición (2) es equivalente a que $J'_0(x_0; y) = 0$ para todo $y \in \mathbf{Y}$ donde $J_0 = \lambda_0 I + \sum_{\gamma=1}^m \lambda_\gamma I_\gamma$. Además si x_0 es una solución normal del problema necesariamente $\lambda_0 > 0$ y los multiplicadores en el teorema anterior se pueden escoger de modo que $\lambda_0 = 1$, en cuyo caso serán únicos.

Esta noción de normalidad puede aplicarse al conjunto $S_1(\lambda)$. Recordemos que si $\lambda = (\lambda_1, \ldots, \lambda_m)$ cumple $\lambda_\alpha \geq 0$ y $\lambda_\alpha I_\alpha(x_0) = 0$ para $\alpha \in A$ y $\Gamma_+(\lambda) = \{\alpha \in A \mid \lambda_\alpha > 0\}$ entonces

$$S_1(\lambda) = \{x \in S \mid I_\alpha(x) = 0 \quad (\alpha \in \Gamma_+(\lambda))\}.$$

Notemos que x_0 es normal relativo a $S_1(\lambda)$ si $\mu = 0$ es la única solución de

1. $\mu_\alpha \geq 0$ y $\mu_\alpha I_\alpha(x_0) = 0$ $(\alpha \in \Gamma_0(\lambda))$,
2. $\sum_{\gamma=1}^m \mu_\gamma I'_\gamma(x_0; y) = 0$ para todo $y \in \mathbf{Y}$.

El siguiente teorema es nuestro resultado principal y puede ser visto como el análogo al Teorema 1.1.17 en dimensión infinita. El artículo original donde se plantea como conjetura es [5].

Teorema 3.3.7. *Supongamos que x_0 resuelve $P(S)$ localmente y existe $\lambda \in \mathbf{R}^q$ con $(x_0, \lambda) \in \mathcal{E}$. Entonces si x_0 es normal relativo a $S_1(\lambda)$ tendremos que*

$$J''(x_0; y) \geq 0 \textit{ para todo } y \in R_{S_1}(x_0).$$

Este teorema es consecuencia sencilla del Teorema 3.3.3 y del hecho que «normalidad implica regularidad», afirmación que se demostrará en el siguiente capítulo. Suponiendo esta afirmación verdadera, normalidad relativa a $S_1(\lambda)$ implicará que $R_{S_1}(x_0) = T_{S_1}(x_0)$, por lo que el Teorema 3.3.3 nos da la conclusión deseada.

3.4. Restricciones Mixtas de Igualdad y Desigualdad

Para acabar con este capítulo derivaremos condiciones de segundo orden para el problema de Lagrange con restricciones mixtas. Esto se hará bajo la suposición de que podemos insertar este problema en el isoperimétrico para usar los resultados de las secciones anteriores.

Supongamos que ahora el conjunto de restricciones está dado por

$$S = \{x \in \mathbf{X}_e \mid \phi_\alpha(t, x(t), \dot{x}(t)) \leq 0 \ (\alpha \in A), \ \phi_\beta(t, x(t), \dot{x}(t)) = 0 \ (\beta \in B) \text{ para todo } t \in [a, b]\},$$

donde los conjuntos de índices A, B, el conjunto $\mathbf{X}$ y los puntos extremos X_a y X_b son como antes.

Consideremos el problema $P(S)$ que consiste en minimizar

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

sobre el conjunto S. Supondremos adicionalmente que f, ϕ_γ son de clase C^2 y que la matriz de dimensión $m \times (n + q)$

$$\begin{pmatrix} \dfrac{\partial \phi_i}{\partial \dot{x}^k} & \delta_{i\alpha} \phi_\alpha \end{pmatrix} \quad (i = 1, \ldots, m; \ \alpha = 1, \ldots, q; \ k = 1, \ldots, n)$$

tiene rango m (aquí $\delta_{\alpha\alpha} = 1$ y $\delta_{\alpha\beta} = 0$ si $\alpha \neq \beta$) en el conjunto U definido por

$$U = \{(t, x, \dot{x}) \in [a, b] \times \mathbf{R}^n \times \mathbf{R}^n \mid \phi_\alpha(t, x, \dot{x}) \leq 0 \ (\alpha \in A), \ \phi_\beta(t, x, \dot{x}) = 0 \ (\beta \in B) \ \}.$$

Esta condición es equivalente a que, en cada punto $(t, x, \dot{x}) \in U$, la matriz

$$\left(\frac{\partial \phi_i}{\partial \dot{x}^k}\right) \quad (i = i_1, \ldots, i_p;\ k = 1, \ldots, n)$$

tiene rango p, donde $i_1, \ldots, i_p$ son los índices en A que satisfacen $\phi_i(t, x, \dot{x}) = 0$ (ver [10] y [18] para más detalles).

Las condiciones de primer orden para este problema se pueden resumir como sigue (ver por ejemplo [18, p. 265, Teorema 5.3]). Denotemos por $\mathcal{U}_m$ las funciones continuas a trozos que mapean $[a, b]$ en $\mathbf{R}^m$.

Teorema 3.4.1. *Supongamos que x_0 resuelve $P(S)$ localmente. Entonces existen $\lambda_0 \geq 0$ y $\mu \in \mathcal{U}_m$ que no se anulan simultáneamente de modo que*

1. $\mu_\alpha(t) \geq 0$ *y* $\mu_\alpha(t)\phi_\alpha(t, x_0(t), \dot{x}_0(t)) = 0$ $(\alpha \in A,\ t \in [a, b])$,
2. $\exists c \in \mathbf{R}^n$ *tal que*

$$F_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_a^t F_x(s, x_0(s), \dot{x}_0(s))ds + c \quad (t \in [a, b]),$$

donde $F(t, x, \dot{x}) = \lambda_0 f(t, x, \dot{x}) + \sum_{\gamma=1}^m \mu_\gamma(t)\phi_\gamma(t, x, \dot{x})$.

La condición (2) anterior, como ya hemos mencionado, es equivalente a que $J'(x_0; y) = 0$ para todo $y \in \mathbf{Y}$, donde

$$J(x) = \int_a^b F(\tilde{x}(t))dt = \lambda_0 I(x) + \sum_{\gamma=1}^m \int_a^b \mu_\gamma(t)\phi_\gamma(\tilde{x}(t))dt.$$

Definición 3.4.2. *Denotemos por $\mathcal{E}$ el conjunto de los pares $(x_0, \mu) \in S \times \mathcal{U}_m$ que satisfacen*

1. $\mu_\alpha(t) \geq 0$ *y* $\mu_\alpha(t)\phi_\alpha(t, x_0(t), \dot{x}_0(t)) = 0$ $(\alpha \in A,\ t \in [a, b])$,
2. $J'(x_0; y) = 0$ *para todo* $y \in \mathbf{Y}$, *donde*

$$J(x) = \int_a^b F_1(\tilde{x}_0(t))dt \quad y \quad F_1(t, x, \dot{x}) = f(t, x, \dot{x}) + \sum_{\gamma=1}^m \mu_\gamma(t)\phi_\gamma(t, x, \dot{x}).$$

Para este problema la definición de normalidad está dada como sigue.

Definición 3.4.3. *Un arco $x_0 \in S$ será llamado normal relativo a S si $\mu \equiv 0$ es la única solución de*

1. $\mu_\alpha(t) \geq 0$ *y* $\mu_\alpha(t)\phi_\alpha(\tilde{x}_0(t)) = 0$ $(\alpha \in A,\ t \in [a, b])$,

2. $J'(x_0; y) = 0$ *para todo* $y \in \mathbf{Y}$*, donde*

$$J(x) = \int_a^b F_0(\tilde{x}_0(t))dt \quad y \quad F_0(t, x, \dot{x}) = \sum_{\gamma=1}^{m} \mu_\gamma(t)\phi_\gamma(t, x, \dot{x}).$$

Por último tenemos que modificar el conjunto de restricciones tangenciales asociado a este problema.

Definición 3.4.4. *Definimos el conjunto* $R_S(x_0)$ *de restricciones tangenciales de* S *en* x_0 *por el conjunto de los elementos* $y \in \mathbf{Y}$ *que cumplen*

$$\phi_{\alpha x}(\tilde{x}_0(t))y(t) + \phi_{\alpha \dot{x}}(\tilde{x}_0(t))\dot{y}(t) \leq 0 \quad (\alpha \in \bar{I}_a(x_0(t)),\ t \in [a, b]),$$

$$\phi_{\beta x}(\tilde{x_0}(t))y(t) + \phi_{\beta \dot{x}}(\tilde{x}_0(t))\dot{y}(t) = 0 \quad (\beta \in B,\ t \in [a, b]),$$

donde $I_a(x_0(t)) = \{\alpha \in A \mid \phi_\alpha(\tilde{x}_0(t)) = 0)\}$*.*

Para este problema podemos plantear una conjetura similar a la del caso isoperimétrico dada en el Teorema 3.3.7. La escribimos a continuación.

Teorema 3.4.5. *Sea* $(x_0, \mu) \in \mathcal{E}$ *y* J *como en la definición de* $\mathcal{E}$*. Supongamos que* x_0 *resuelve* $P(S)$ *localmente y es normal relativo a* $S_1(\mu) = \{x \in S \mid I(x) = J(x)\}$*. Entonces* $J''(x_0; y) \geq 0$ *para todo* $y \in R_{S_1}(x_0)$*.*

Esta conjetura aún no se resuelve, sin embargo, veremos a continuación que un resultado similar se puede derivar de forma sencilla «extendiendo» al problema $P(\bar{S})$ que consiste en minimizar el funcional I sobre el conjunto $\bar{S}$ dado por

$$\bar{S} = \{x \in \mathbf{X}_e \mid I_\alpha(x) \leq 0\ (\alpha \in A),\ I_\beta(x) = 0\ (\beta \in B)\}$$

donde

$$I_\gamma(x) = \int_a^b \phi_\gamma(t, x(t), \dot{x}(t))dt \quad (\gamma \in A \cup B,\ x \in \mathbf{X}),$$

y suponiendo que nuestra solución x_0 resuelve el problema original y el extendido. Tendremos dos versiones de este resultado, la primera usando la noción de normalidad fuerte o cualquiera de sus equivalencias y la segunda usando normalidad relativa a S_1.

Teorema 3.4.6. *Sea* $(x_0, \mu) \in \mathcal{E}$ *y* J *como en la definición de* $\mathcal{E}$*. Supongamos que* x_0 *resuelve* $P(S)$ *localmente y* $\mu \equiv \lambda$ *es constante. Si además* x_0 *resuelve* $P(\bar{S})$ *y es un punto normal relativo a* $\bar{S}_0$*, entonces*

$$J''(x_0; y) \geq 0 \textit{ para todo } y \in \mathbf{Y} \textit{ con}$$

$$I'_\alpha(x_0; y) \leq 0 \quad (\alpha \in A \textit{ con } I_\alpha(x_0) = 0 \textit{ y } \lambda_\alpha = 0),$$

$$I'_\beta(x_0; y) = 0 \quad (\beta \in A \textit{ con } \lambda_\beta > 0 \textit{ ó } \beta \in B).$$

Teorema 3.4.7. *Si en el teorema anterior cambiamos la hipótesis de normalidad relativa a* $\bar{S}_0$ *por normalidad relativa a* $\bar{S}_1(\mu)$, *la conclusión sigue siendo verdadera.*

Estos resultados son inmediatos de los Teoremas 2.3.7 y 3.3.7 respectivamente. Es importante notar que el conjunto donde se satisface la no negatividad de J'' en los teoremas anteriores contiene al conjunto $R_{S_1}(x_0)$ del Teorema 3.4.5 (que se obtiene al aplicar la Definición 3.4.4 al conjunto $S_1(\mu)$).

Capítulo 4

Normalidad implica regularidad

En este capítulo completaremos la demostración del Teorema 3.3.7. Esto se hará a través de varios resultados auxiliares que son generalizaciones de las técnicas usadas en [19, C. 4, Sección 10].

4.1. Demostración de Normalidad Implica Regularidad

Recordemos que x_0 es normal relativo a S si $\lambda = 0 \in \mathbf{R}^m$ es la única solución de

1. $\lambda_\alpha \geq 0$ y $\lambda_\alpha I_\alpha(x_0) = 0 \quad (\alpha \in A)$,
2. $\sum_{\gamma=1}^{m} \lambda_\gamma I'_\gamma(x_0; y) = 0$ para todo $y \in \mathbf{Y}$.

Comencemos enunciando el siguiente resultado que podemos encontrar en [18, p. 70].

Lema 4.1.1. *Dados el funcional integral I y un arco admisible $x \in S$, existe un único elemento $z \in \mathbf{Y}$ de forma que*

$$I'(x; y) = (z; y) \quad \textit{para todo } y \in \mathbf{Y},$$

donde $(\cdot;\cdot)$ es el producto interno en $\mathbf{Y}$ definido por

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

y $\langle \cdot;\cdot \rangle$ el producto interno euclidiano. Al elemento z se le llamará el gradiente en I en x.

Denotemos por z_γ los gradientes de I_γ en x_0 para $\gamma \in A \cup B$, es decir, los únicos elementos $z_\gamma \in \mathbf{Y}$ de forma que

$$I'_\gamma(x_0; y) = (z_\gamma; y) = \int_{t_0}^{t_1} \langle \dot{z}_\gamma(t); \dot{y}(t) \rangle dt$$

para todo $y \in \mathbf{Y}$.

Lema 4.1.2. *Sea $\mathcal{P} \subset A \cup B$. Los funcionales lineales $I'_\gamma(x_0; \cdot)$ $(\gamma \in \mathcal{P})$ son linealmente independientes en $\mathbf{Y}$ si y solo si los gradientes $\{z_\gamma : \gamma \in \mathcal{P}\}$ son linealmente independientes en $\mathbf{Y}$.*

Demostración. De la igualdad

$$\sum_\gamma \lambda_\gamma I'_\gamma(x_0; y) = \sum_\gamma \lambda_\gamma (z_\gamma; y) = \sum_\gamma (\lambda_\gamma z_\gamma; y),$$

tendremos que los funcionales $I'_\gamma(x_0; \cdot)$ $(\gamma \in \mathcal{P})$ son linealmente independientes en $\mathbf{Y}$ si y solo si

$$\sum_\gamma (\lambda_\gamma z_\gamma; y) = 0 \ \text{ para todo } y \in \mathbf{Y}$$

implica que $\lambda_\gamma = 0$ $(\gamma \in \mathcal{P})$. Como $(\cdot;\cdot)$ es producto interno, la igualdad de arriba ocurre si y solo si

$$\sum_\gamma \lambda_\gamma z_\gamma = 0,$$

con lo que concluimos la demostración. $\square$

Sin pérdida de generalidad supondremos, hasta que se indique lo contrario, que todos los índices son activos, es decir $I_a(x_0) = A$. Las restricciones tangenciales de S en x_0 pueden ser reescritas de la siguiente manera.

$$R_S(x_0) = \{y \in \mathbf{Y} \mid (z_\alpha; y) \leq 0 \ (\alpha \in A), \ (z_\beta; y) = 0 \ (\beta \in B)\}.$$

Definamos el conjunto Γ por

$$\Gamma = \{\gamma \in A \mid \exists \ y \in R_S(x_0) \ni (z_\gamma; y) < 0\},$$

y el conjunto $\Delta = A - \Gamma$. Entonces las restricciones tangenciales de S en x_0 serán equivalentes a las restricciones tangenciales del conjunto S^* en x_0, donde

$$S^* = \{x \mid I_\alpha(x) \leq 0 \ (\alpha \in \Gamma), \ I_\beta(x) = 0 \ (\beta \in \Delta \cup B)\}.$$

Es decir, $R_S(x_0)$ es igual al conjunto

$$R_{S^*}(x_0) = \{y \in \mathbf{Y} \mid (z_\alpha; y) \leq 0 \ (\alpha \in \Gamma),\ (z_\beta; y) = 0 \ (\beta \in \Delta \cup B)\}.$$

Lema 4.1.3. *Si x_0 es regular de S^* entonces también es regular de S.*

Demostración. Como $S^* \subset S$, por el inciso 2 del Lema 3.1.4 tendremos que $T_{S^*}(x_0) \subset T_S(x_0)$. Por otro lado, si x_0 es regular de S^* entonces

$$T_{S^*}(x_0) = R_{S^*}(x_0) = R_S(x_0) \supset T_S(x_0),$$

por lo que

$$T_{S^*}(x_0) = R_{S^*}(x_0) = R_S(x_0) = T_S(x_0). \qquad \square$$

Definamos al conjunto H por

$$H = \{y \in \mathbf{Y} \mid (z_\alpha; y) < 0 \ (\alpha \in \Gamma),\ (z_\beta; y) = 0 \ (\beta \in \Delta \cup B)\} \cup \{0\}.$$

Lema 4.1.4. *H es un cono convexo cuya cerradura $\overline{H}$ (relativa a $\mathbf{Y}$) es $R_S(x_0)$.*

Demostración. Verificar que H es cono convexo es trivial. Supongamos que $h \in \overline{H}$, entonces existe una sucesión $\{h_q\} \subset H$ de forma que $\|h - h_q\|$ tiende a cero si q tiende a infinito. En particular $\dot{h}_q$ tiende uniformemente a $\dot{h}$. Como z_γ está en $\mathbf{Y}$, es acotada. Tomemos $M_\gamma > 0$ de forma que $|z_\gamma(t)| \leq M_\gamma$ para todo t. Sea $\epsilon > 0$, por la convergencia uniforme existe N de forma que si $q > N$ entonces $|\dot{h}_q(t) - \dot{h}(t)| < \epsilon / M_\gamma(b-a)$. De donde

$$\begin{aligned}
(z_\gamma; h - h_q) = &\left| \int_a^b \langle \dot{z}_\gamma(t); \dot{h}_q(t) - \dot{h}(t) \rangle dt \right| \\
\leq &\int_a^b |\langle \dot{z}_\gamma(t); \dot{h}_q(t) - \dot{h}_q(t) \rangle| dt \\
\leq &\int_a^b |\dot{z}_\gamma(t)| |\dot{h}_q(t) - \dot{h}_q(t)| dt \\
\leq &\int_a^b M_\gamma \frac{\epsilon}{M_\gamma(b-a)} dt = \epsilon
\end{aligned}$$

esto quiere decir que $\lim(z_\gamma; h_q) = (z_\gamma; h)$, por lo que si $\gamma \in A$ entonces

$$0 \geq \lim_q (z_\gamma; h_q) = (z_\gamma; h),$$

mientras que $0 = (z_\gamma; h)$ si $\gamma \in B$. Con esto obtenemos que $\overline{H} \subset R_S(x_0)$. Para verificar la otra contención notemos que, si $y \in R_S(x_0)$ y existe $0 \neq h \in H$, entonces $y + ch \in H$ para toda constante $c > 0$. Si definimos $h_q = y + h/q$ es claro que h_q y $\dot{h}_q$ tienden a y y $\dot{y}$ uniformemente por lo que $y \in \overline{H}$. En caso de que el único elemento en H es 0, entonces $\Gamma = \emptyset$ y $H = R_{S^*}(x_0)$, por lo que $\{0\} = \overline{H} = R_{S^*}(x_0) = R_S(x_0)$. □

Lema 4.1.5. *x_0 es regular de S si y sólo si $H \subset T_S(x_0)$. x_0 es regular de S^* si y sólo si $H \subset T_{S^*}(x_0)$.*

Demostración. Si x_0 es regular de S entonces $H \subset R_S(x_0) = T_S(x_0)$. Por otro lado si $H \subset T_S(x_0)$ entonces $R_S(x_0) = \overline{H} \subset T_S(x_0)$. Esta última contención ya que el cono tangente es cerrado. La demostración de la segunda conclusión es análoga. □

Nota 4.1.6. En [19, p. 238] se hace notar que las restricciones tangenciales en x_0 asociadas al conjunto

$$K = \{y \in \mathbf{Y} \mid I_\alpha(x) \leq 0 \ (\alpha \in \Delta),\ I_\beta(x) = 0 \ (\beta \in B)\}$$

son equivalentes a las restricciones tangenciales en x_0 asociadas al conjunto

$$K_0 = \{y \in \mathbf{Y} \mid I_\gamma(x) = 0 \ (\gamma \in \Delta \cup B)\},$$

en dimensión finita. Es fácil ver que podemos ajustar la técnica que se utiliza para concluir que $R_K(x_0) = R_{K_0}(x_0)$.

Lema 4.1.7. *Si $\Delta \neq \emptyset$, entonces los gradientes $\{z_\gamma \mid \gamma \in \Delta \cup B\}$ son linealmente dependientes.*

Demostración. Si $\Delta \neq \emptyset$, por la nota anterior podemos aplicar el Lema 1.3.4 a los funcionales lineales $(z_\gamma;\ \cdot\ )$ $(\gamma \in \Delta \cup B)$ para asegurar que existen multiplicadores λ_γ $(\gamma \in \Delta \cup B)$ de forma que $\lambda_\gamma > 0$ si $\gamma \in \Delta$ y

$$\sum_{\gamma \in \Delta \cup B} \lambda_\gamma z_\gamma = 0.$$ □

Proposición 4.1.8. *Si x_0 es regular de K entonces también es regular de S. Si x_0 es regular de K_0 también es regular de K, S^* y S.*

Demostración. Para demostrar la primer conclusión, por el Lema 4.1.5 basta ver que $H \subset T_S(x_0)$. Si $H = \{0\}$ la contención es trivial. Sea $0 \neq y \in H$, entonces como x_0 es regular de K y $y \in R_{K_0}(x_0) = R_K(x_0)$, tendremos que $y \in T_K(x_0)$ por lo que existe $\{x_q\} \subset K$ y $\{t_q\} \subset \mathbf{R}^+$ con $t_q \to 0$ y $x_q - x_0/t_q \to y$. Entonces

$$\lim_q \frac{I_\gamma(x_q)}{t_q} = \lim_q \frac{I_\gamma(x_q) - I_\gamma(x_0)}{t_q} = (z_\gamma, y).$$

Como $(z_\gamma; y) < 0$ para $\gamma \in \Gamma$, podemos escoger q_0 de forma que $I_\gamma(x_q) < 0$ si $q > q_0$. Por lo tanto $\{x_q\} \subset S$ si $q > q_0$ por lo que $y \in T_S(x_0)$. Si x_0 es regular de K_0, y $y \in H$ entonces $y \in R_{K_0}(x_0) = T_{K_0}(x_0)$ por lo que podemos proceder como antes para verificar que $H \subset T_{S^*}(x_0)$, y de nuevo el Lema 4.1.5 nos asegurará la regularidad de x_0 en S^* que a su vez, por el Lema 4.1.3, nos garantiza la normalidad de x_0 en S. Para verificar que la regularidad de K_0 implica la de K notemos que

$$R_K(x_0) = R_{K_0}(x_0) = T_{K_0}(x_0) \subset T_K(x_0) \subset R_K(x_0),$$

por lo que $R_K(x_0) = T_K(x_0)$, donde la primera igualdad es por la Nota 4.1.6, la segunda igualdad se debe a la regularidad de x_0 en K_0 y las últimas dos contenciones son por los Lemas 3.1.4 y 3.1.2 respectivamente. □

Nota 4.1.9. Como K_0 consta sólo de restricciones de igualdad, el Lema 4.1.2 con $\mathcal{P} = \Delta \cup B$ nos asegura que, si los gradientes $\{z_\beta \mid \beta \in \Delta \cup B\}$ son linealmente independientes, entonces x_0 es normal fuerte de K_0, y por la Proposición 3.2.4 regular de K_0.

Teorema 4.1.10. *Si los gradientes $\{z_\gamma \mid \gamma \in \Delta \cup B\}$ son linealmente independientes, entonces $\Delta = \emptyset$ y x_0 es regular de S. Si $\Delta = \emptyset$ y los gradientes $\{z_\beta \mid \beta \in B\}$ son linealmente independientes, entonces x_0 es regular de S.*

Demostración. Por el Lema 4.1.7 debemos tener $\Delta = \emptyset$, ya que de lo contrario los vectores $\{z_\gamma \mid \gamma \in \Delta \cup B\}$ no pueden ser linealmente independientes. Por la nota anterior, x_0 es regular de K_0, de forma que la Proposición 4.1.8 nos asegura que x_0 es regular de S. La segunda conclusión se sigue trivialmente de la primera. □

Por último para mostrar que normalidad de S implica regularidad de S, reescribiremos la normalidad de una forma equivalente a las hipótesis de la proposición anterior.

Teorema 4.1.11. *x_0 es normal de S si y sólo si los gradientes $\{z_\beta \mid \beta \in B\}$ son linealmente independientes y existe un vector $0 \neq h \in \mathbf{Y}$ con*

$$(z_\alpha; h) < 0 \ (\alpha \in A) \qquad (z_\beta; h) = 0 \ (\beta \in B).$$

Demostración. Primero supongamos que los gradientes $\{z_\beta \mid \beta \in B\}$ son linealmente independientes y existe un vector $0 \neq h \in \mathbf{Y}$ con las propiedades indicadas. Si $\lambda_1, \ldots, \lambda_m$ son multiplicadores como en la definición de normalidad entonces

$$0 = \sum_{\gamma=1}^{m} \lambda_\gamma I'_\gamma(x_0; h) = \sum_{\gamma=1}^{m} \lambda_\gamma (z_\gamma; h) = \sum_{\gamma=1}^{q} \lambda_\gamma (z_\gamma; h).$$

Como $(z_\gamma; h) < 0$ y $\lambda_\gamma \geq 0$ si $\gamma \in \Gamma = A$, debemos tener que $\lambda_\gamma = 0$ si $\gamma \in A$. De esto que

$$\sum_{\gamma=q+1}^{m} (\lambda_\gamma z_\gamma; y) = 0 \quad \text{para todo } y \in \mathbf{Y},$$

por lo que $\lambda_{q+1} z_{q+1} + \cdots + \lambda_m z_m = 0$, y como $z_{q+1}, \ldots, z_m$ son linealmente independientes esto implica que $\lambda_\gamma = 0$ para $\gamma \in B$.

Ahora supongamos que x_0 es normal de S. En particular haciendo $\lambda_\alpha = 0$ para $\alpha \in A$, la única solución de

$$\lambda_{q+1} I'_{q+1}(x_0; y) + \cdots + \lambda_m I'_m(x_0; y) = 0,$$

para todo $y \in \mathbf{Y}$ es $\lambda_\gamma = 0$ $(\gamma \in B)$. Pero

$$0 = \lambda_{q+1} I'_{q+1}(x_0; y) + \cdots + \lambda_m I'_m(x_0; y) = \left(\sum_{\gamma=q+1}^{m} \lambda_\gamma z_\gamma; y \right)$$

para todo $y \in \mathbf{Y}$ si y sólo si

$$\lambda_{q+1} z_{q+1} + \cdots + \lambda_m z_m = 0,$$

por lo que la única solución a esta última ecuación es $\lambda_\beta = 0$ $(\beta \in B)$. De esto que los vectores $\{z_\beta \mid \beta \in B\}$ son linealmente independientes. Más aún, $\Delta = \emptyset$ ya que de lo contrario existiría una solución de

$$\sum_{\gamma \in A \cup B} \lambda_\gamma z_\gamma = 0$$

con $\lambda_\alpha \geq 0$ si $\alpha \in A$ y distinta de cero, lo que por la hipótesis de normalidad no es posible. Entonces debe existir h_α con

$$(z_\alpha; h_\alpha) < 0$$

para cada $\alpha \in A$. El vector $h = h_1 + \cdots + h_q$ cumple con lo que necesitamos. □

Teorema 4.1.12. *Sea $x_0 \in S$. Si x_0 es normal de S, también es regular de S.*

Demostración. Notemos que la existencia de $0 \neq h \in \mathbf{Y}$ con las propiedades enunciadas en el teorema anterior, es justamente que $\Delta = \emptyset$. En vista de esto y del Teorema 4.1.11, la segunda conclusión del Teorema 4.1.10 puede ser reescrita como «normalidad implica regularidad». □

4.2. Ejemplo

En esta sección daremos un ejemplo donde la teoría clásica no es suficiente para decidir si un extremo es o no es solución. Consideremos el problema de minimizar $I(x) = \int_0^{\pi/2} [x^2(t) - \dot{x}^2(t)]dt$ sujeto a las restricciones

$$\begin{cases} I_1(x) = \int_0^{\pi/2} x(t)dt \leq 0, \\ I_2(x) = \int_0^{\pi/2} [x(t) + \dot{x}(t)]dt + 2 \leq 0, \\ x(0) = 1,\ x(\pi/2) = -1. \end{cases}$$

En este caso $n = 1$,

$$\begin{array}{cccccc} f = x^2 - \dot{x}^2 & f_x = 2x & f_{\dot{x}} = -2\dot{x} & f_{xx} = 2 & f_{x\dot{x}} = 0 & f_{\dot{x}\dot{x}} = -2 \\ f_1 = x & f_{1x} = 1 & f_{1\dot{x}} = 0 & f_{1xx} = 0 & f_{1x\dot{x}} = 0 & f_{1\dot{x}\dot{x}} = 0 \\ f_2 = x + \dot{x} + 4/\pi & f_{2x} = 1 & f_{2\dot{x}} = 1 & f_{2xx} = 0 & f_{2x\dot{x}} = 0 & f_{2\dot{x}\dot{x}} = 0 \end{array}$$

Definiendo la función

$$\begin{aligned} F &= f + \lambda_1 f_1 + \lambda_2 f_2 \\ &= x^2 - \dot{x}^2 + \lambda_1 x + \lambda_2(x + \dot{x} + 4/\pi), \end{aligned}$$

la ecuación de Euler-Lagrange

$$\frac{dF_{\dot{x}}}{dt} = F_x$$

toma la forma

$$\ddot{x}(t) + x(t) + \frac{\lambda_1 + \lambda_2}{2} = 0,$$

cuya solución general es

$$x(t) = c_1 \operatorname{sen}(t) + c_2 \cos(t) - \frac{\lambda_1 + \lambda_2}{2}.$$

Por lo tanto

$$1 = x(0) = c_2 - \frac{\lambda_1 + \lambda_2}{2}$$

$$-1 = x(\pi/2) = c_1 - \frac{\lambda_1 + \lambda_2}{2}$$

de donde

$$c_1 = \frac{\lambda_1 + \lambda_2}{2} - 1, \quad c_2 = 1 + \frac{\lambda_1 + \lambda_2}{2}.$$

Por otro lado

$$\begin{aligned} I_1(x) = & \int_0^{\pi/2} [c_1 \operatorname{sen}(t) + c_2 \cos(t) - \frac{\lambda_1 + \lambda_2}{2}]dt \\ &= c_1[\cos(0) - \cos(\pi/2)] + c_2[\operatorname{sen}(\pi/2) - \operatorname{sen}(0)] - \frac{\pi}{2}\frac{\lambda_1 + \lambda_2}{2} \\ &= c_1 + c_2 - \frac{\pi}{2}\frac{\lambda_1 + \lambda_2}{2} \end{aligned}$$

e

$$\begin{aligned} I_2(x) = I_1(x) + & \int_0^{\pi/2} [c_1 \cos(t) - c_2 \operatorname{sen}(t)]dt + 2 \\ &= c_1[\operatorname{sen}(\pi/2) - \operatorname{sen}(0)] + c_2[\cos(\pi/2) - \cos(0)] + I_1(x) + 2 \\ &= c_1 - c_2 + 2 + c_1 + c_2 - \frac{\pi}{2}\frac{\lambda_1 + \lambda_2}{2}. \end{aligned}$$

Notemos que, si $\lambda_1 = 0 = \lambda_2$ entonces $c_1 = -1$ y $c_2 = 1$ y el arco admisible

$$x_0(t) = -\operatorname{sen}(t) + \cos(t)$$

satisface $I_1(x_0) = 0 = I_2(x_0)$. En este caso $S_1(0,0) = S$.

Veamos que x_0 es normal de S. Supongamos que $\lambda_1 I_1'(x_0; y) + \lambda_2 I_2'(x_0; y) = 0$ para cada $y \in \mathbf{Y}$ y $\lambda_1, \lambda_2 \geq 0$, entonces

$$\begin{aligned} \lambda_1 \int_0^{\pi/2} y(t)dt + \lambda_2 \int_0^{\pi/2} [y(t) + \dot{y}(t)]dt &= (\lambda_1 + \lambda_2) \int_0^{\pi/2} y(t)dt + \lambda_2 \int_0^{\pi/2} \dot{y}(t)dt \\ &= (\lambda_1 + \lambda_2) \int_0^{\pi/2} y(t)dt = 0. \end{aligned}$$

La segunda igualdad es consecuencia de que $\int_0^{\pi/2} \dot{y}(t)dt$ se anula en $\mathbf{Y}$. Como requerimos que la igualdad se satisfaga para todo $y \in \mathbf{Y}$, $\lambda_1 + \lambda_2$ debe ser cero. Combinando este hecho con la condición $\lambda_1, \lambda_2 \geq 0$, obtenemos que $\lambda_1 = \lambda_2 = 0$. Por otro lado, si eliminamos la condición $\lambda_1, \lambda_2 \geq 0$, obtenemos que $\lambda_1 = -\lambda_2$ son posibles soluciones no triviales, lo que quiere decir que x_0 no es fuertemente normal de S.

Finalmente notemos que la función y definida por

$$y(t) = \begin{cases} -t, \quad t \in [0, \pi/4], \\ -\pi/2 + t, \quad t \in [\pi/4, \pi/2] \end{cases}$$

es continua, satisface $y(0) = 0 = y(\pi/2)$, $I_1'(x_0; y) \leq 0$, $I_2'(x_0; y) \leq 0$ y si definimos $J(x) := \int_0^{\pi/2} F(\tilde{x}(t))dt$, entonces

$$\begin{aligned} J''(x_0, y) &= \int_0^{\pi/2} \{y^2(t) - \dot{y}^2(t)\}dt \\ &= \int_0^{\pi/4} (t^2 - 1)dt + \int_{\pi/4}^{\pi/2} \left\{(t - \pi/2)^2 - 1\right\} dt \\ &= \frac{(\pi/2)^3}{3} - \pi/2 < 0. \end{aligned}$$

por lo que x_0 no es solución.

Capítulo 5

Generalización a Control Óptimo

En este capítulo daremos un breve resumen sobre los resultados principales que involucran normalidad y regularidad, para después abordar el problema de cómo trasladar nuestra nueva versión de normalidad al contexto de control óptimo.

5.1. Resumen

Hagamos un recuento de los resultados más importantes que hemos discutido hasta este momento que involucran normalidad y regularidad. Denotamos por $\mathcal{E}$ el conjunto de elementos $(x_0, \lambda) \in S \times \mathbf{R}^m$ que satisfacen

1. $\lambda_\alpha \geq 0,\ \lambda_\alpha I_\alpha(x_0) = 0\ (\alpha \in A)$,
2. Si $J := I + \sum_{\gamma=1}^{m} \lambda_\gamma I_\gamma$ entonces $J'(x_0; y) = 0$ para todo $y \in \mathbf{Y}$.

Para $(x_0, \lambda) \in \mathbf{X} \times \mathbf{R}^m$ consideremos los siguientes conjuntos de restricciones

$$
\begin{aligned}
S_0(x_0) &= \{x \in \mathbf{X}_e \mid I_\alpha(x) = 0\ (\alpha \in I_a(x_0) \cup B)\},\\
S_1(\lambda) &= \{x \in \mathbf{X}_e \mid I_\alpha(x) \leq 0\ (\alpha \in \Gamma_0(\lambda)),\ I_\beta(x) = 0\ (\beta \in \Gamma_+(\lambda) \cup B)\},\\
S &= \{x \in \mathbf{X}_e \mid I_\alpha(x) \leq 0\ (\alpha \in A),\ I_\beta(x) = 0\ (\beta \in B)\},
\end{aligned}
$$

junto con sus respectivos conjuntos de restricciones tangenciales en x_0

$$R_{S_0}(x_0) = \{y \in \mathbf{Y} \mid I'_\alpha(x_0; y) = 0 \ (\alpha \in I_a(x_0) \cup B)\},$$
$$R_{S_1}(x_0) = \{y \in \mathbf{Y} \mid I'_\alpha(x_0; y) \leq 0 \ (\alpha \in I_a(x_0) \cap \Gamma_0(\lambda)), \ I'_\beta(x_0; y) = 0 \ (\beta \in \Gamma_+(\lambda) \cup B)\},$$
$$R_S(x_0) = \{y \in \mathbf{Y} \mid I'_\alpha(x_0; y) \leq 0 \ (\alpha \in I_a(x_0)), \ I'_\beta(x_0; y) = 0 \ (\beta \in B)\},$$

donde $I_a(x_0) = \{\alpha \in A \mid I_\alpha(x_0) = 0\}$, $\Gamma_0(\lambda) = \{\alpha \in A \mid \lambda_\alpha = 0\}$ y $\Gamma_+(\lambda) = \{\alpha \in A \mid \lambda_\alpha > 0\}$. Notemos que $R_{S_0}(x_0) \subset R_{S_1}(x_0) \subset R_S(x_0)$.

Las condiciones necesarias basadas en la regularidad con respecto de $S_0, S_1(\lambda)$ y S se pueden resumir de la siguiente manera.

Teorema 5.1.1. *Supongamos que $(x_0, \lambda) \in \mathcal{E}$ y x_0 es solución local de $P(S)$. Entonces los siguientes enunciados son verdaderos.*

- **a.** *Si x_0 es regular de S, entonces $\exists \lambda \in \mathbf{R}^m$ de modo que $(x_0, \lambda) \in \mathcal{E}$.*
- **b.** *Si $(x_0, \lambda) \in \mathcal{E}$ y x_0 es regular de $S_0(x_0)$, entonces $J''(x_0, y) \geq 0$ para todo $y \in R_{S_0}(x_0)$.*
- **c.** *Si $(x_0, \lambda) \in \mathcal{E}$ y x_0 es regular de $S_1(\lambda)$, entonces $J''(x_0; y) \geq 0$ para todo $y \in R_{S_1}(x_0)$.*

Con respecto a normalidad tenemos las siguientes definiciones.

- $x_0 \in S$ es un arco normal relativo a $S_0(x_0)$ si $\lambda_\alpha I_\alpha(x_0) = 0 \ (\alpha \in A)$ y $\sum_{\gamma=1}^m \lambda_\gamma I'_\gamma(x_0; y) = 0$ para todo $y \in \mathbf{Y}$ implica que $\lambda = 0$.
- $x_0 \in S$ es un arco normal relativo a $S_1(\lambda)$, si $\mu_\alpha \geq 0, \ \mu_\alpha I_\alpha(x_0) = 0 \ (\alpha \in \Gamma_0(\lambda))$ y $\sum_{\gamma=1}^m \mu_\gamma I'_\gamma(x_0; y) = 0$ para todo $y \in \mathbf{Y}$ implica que $\mu = 0$.
- $x_0 \in S$ es un arco normal relativo a S, si $\lambda_\alpha \geq 0, \lambda_\alpha I_\alpha(x_0) = 0 \ (\alpha \in A)$ y $\sum_{\gamma=1}^m \lambda_\gamma I'_\gamma(x_0; y) = 0$ para todo $y \in \mathbf{Y}$ implica que $\lambda = 0$.

Notemos que, si x_0 es normal relativo a S_0, entonces es normal relativo a $S_1(\lambda)$ con $(x_0, \lambda) \in \mathcal{E}$, y por lo tanto normal relativo a S. Además, si x_0 es normal relativo a S_0, es decir s-normal relativo a S, entonces es p-regular y por lo tanto regular de $S_0, S_1(\lambda)$ y S. En particular esto implica el siguiente resultado.

Teorema 5.1.2. *Supongamos que $(x_0, \lambda) \in \mathcal{E}$. Si x_0 es solución local de $P(S)$ y es normal relativo a S_0, entonces $J''(x_0; y) \geq 0$ para todo $y \in R_{S_1}(x_0)$.*

El resultado principal sobre condiciones necesarias de segundo orden es el Teorema 3.3.7, que puede ser escrito como sigue.

Teorema 5.1.3. *Supongamos que* $(x_0, \lambda) \in \mathcal{E}$. *Si* x_0 *es solución local de* $P(S)$ *y es normal relativo a* $S_1(\lambda)$, *entonces* $J''(x_0; y) \geq 0$ *para todo* $y \in R_{S_1}(x_0)$.

5.2. El Caso en Control Óptimo

El problema isoperimétrico de Lagrange con puntos extremos fijos en el contexto de control óptimo se plantea como sigue. Denotemos por $\mathcal{U}_p$ el conjunto de las funciones $u : [a, b] \to \mathbf{R}^p$ continuas a trozos y $Z = \mathbf{X} \times \mathcal{U}_p$. Supongamos que tenemos funciones $L, L_\gamma : \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^p \to \mathbf{R}$ y $f : \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^p \to \mathbf{R}^n$ de clase C^2 en las variables independientes (t, x, u). Definimos los conjuntos

$$D = \{(x, u) \in Z \mid \dot{x}(t) = f(t, x(t), u(t)),\ x(a) = X_a,\ x(b) = X_b\},$$

y

$$S = \{(x, u) \in D \mid I_\alpha(x, u) \leq 0\ (\alpha \in A),\quad I_\beta(x, u) = 0\ (\beta \in B)\},$$

donde X_a y X_b son puntos fijos en $\mathbf{R}^n$, A y B son conjuntos de índices como en los capítulos anteriores y los funcionales I_γ están definidos por

$$I_\gamma(x, u) = \int_a^b L_\gamma(t, x(t), u(t))dt.$$

El nuevo problema, que denotaremos por $CP(S)$, es el de minimizar el funcional $I(x, u) = \int_a^b L(t, x(t), u(t))dt$ en el nuevo conjunto de restricciones S. Para cada pareja $(x_0, u_0) \in Z$ denotaremos por $\tilde{x}_0(t)$ a la terna $(t, x_0(t), u_0(t)) \in [a, b] \times \mathbf{R}^n \times \mathbf{R}^p$. Ahora definimos los conjuntos

$$\mathbf{Y} = \{(y, v) \in Z \mid \dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t),\ y(a) = 0 = y(b)\}$$

y

$$R_S(x_0, u_0) = \{(y, v) \in \mathbf{Y} \mid I'_\alpha(x_0, u_0; y, v) \leq 0\ (\alpha \in I_a(x_0, u_0)),\quad I'_\beta(x_0, u_0; y, v) = 0\ (\beta \in B)\},$$

donde $I_a(x_0, u_0)$ tiene el mismo significado que en las secciones anteriores, y para $(x_0, u_0) \in Z$ y $(y, v) \in \mathbf{Y}$

$$I'_\gamma(x_0, u_0; y, v) = \int_a^b \{L_{\gamma x}(\tilde{x}_0(t))y(t) + L_{\gamma u}(\tilde{x}_0(t))v(t)\}dt\quad (\gamma \in A \cup B).$$

Condiciones necesarias de primer orden para $CP(S)$ se enuncian como sigue (ver [18, p. 254, Teorema 2.1] por ejemplo).

Teorema 5.2.1. *Supongamos que $(x_0, u_0) \in S$ resuelve $CP(S)$. Entonces $\exists\ \lambda_0 \geq 0,\ \lambda \in \mathbf{R}^m$ y $p \in \mathbf{X}$, que no se anulan simultáneamente para todo $t \in [a, b]$, y de forma que, junto con la función*

$$H(t, x, u, p) = \langle p; f(t, x, u)\rangle - \lambda_0 L(t, x, u) - \sum_{\gamma=1}^{m} \lambda_\gamma L_\gamma(t, x, u),$$

satisfacen

1. $\lambda_\alpha \geq 0,\ \lambda_\alpha I_\alpha(x_0, u_0) = 0\ (\alpha \in A)$.

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

$$\dot{x}^i = H_{p^i} = f^i, \quad \dot{p}_i = -H_{x^i},$$

en los puntos $(t, x_0(t), u_0(t), p(t))$ para t en los intervalos de continuidad de u_0.

3. *Se tiene la desigualdad*

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

para todo u en una vecindad de $u_0(t)$.

4. *La función $H(t, x_0(t), u_0(t), p(t))$ es continua y satisface*

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

en los intervalos de continuidad de u_0.

5. $H_u(t, x_0(t), u_0(t), p(t)) = 0 \quad (t \in [a, b])$.

De nuevo, sin pérdida de generalidad supongamos que todos los índices son activos. Normalidad para este problema se puede definir como sigue (ver [18, p. 273], comparar con v-normalidad de la Sección 2.4).

Definición 5.2.2. *Llamaremos a $(x_0, u_0) \in S$ normal fuerte si existen $(y_\sigma, v_\sigma) \in Z$ $(\sigma = 1, \ldots, N = m + n)$, de forma que $y(a) = 0,\ \dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t)$ y el determinante*

$$\begin{vmatrix} I'_\gamma(x_0, u_0; y_\sigma, v_\sigma) \\ y^i_\sigma(b) \end{vmatrix} \qquad \begin{matrix} (\gamma = 1, \ldots, m; \quad i = 1, \ldots, n), \\ (\sigma = 1, \ldots, N), \end{matrix}$$

es distinto de cero.

Si en el Teorema 5.2.1 además suponemos que (x_0, u_0) es normal fuerte, podremos tomar $\lambda_0 = 1$, y en este caso los multiplicadores serán únicos (ver [18, p. 275, Teorema 7.2]).

Definamos ahora la segunda variación del funcional I como

$$I''(x_0, u_0; y, v) = \int_a^b 2\Omega(t, y(t), v(t))dt,$$

donde

$$2\Omega(t, y, v) := \langle y; L_{xx}(\tilde{x}_0(t))y\rangle + 2\langle y; L_{xu}(\tilde{x}_0(t))v\rangle + \langle v; L_{uu}(\tilde{x}_0(t))v\rangle,$$

y las segundas variaciones de los funcionales I_γ están dadas de manera análoga.

Definimos el nuevo conjunto de extremos $\mathcal{E}$, como los elementos $(x_0, u_0, \lambda, p) \in S \times \mathbf{R}^m \times \mathbf{X}$ que satisfacen las condiciones (1), (2) y (5) del Teorema 5.2.1 con $\lambda_0 = 1$ (se define así ya que la condición 5 es consecuencia de las condiciones 3 y 4). Al igual que en los capítulos anteriores, para $(x_0, u_0, \lambda) \in S \times \mathbf{R}^m$ definimos el conjunto

$$S_1(\lambda) = \{(x, u) \in D \mid I_\alpha(x, u) \leq 0 \ (\alpha \in \Gamma_0(\lambda)), \quad I_\beta(x, u) = 0 \ (\beta \in \Gamma_+(\lambda) \cup B)\},$$

y su conjunto de restricciones tangenciales asociado

$$\begin{aligned} R_{S_1}(x_0, u_0) = \{(y, v) \in \mathbf{Y} \mid\ & I'_\alpha(x_0, u_0; y, v) \leq 0 \ (\alpha \in \Gamma_0(\lambda) \cap I_a(x_0, u_0)), \quad I'_\beta(x_0, u_0; y, v) = 0 \\ & (\beta \in \Gamma_+(\lambda) \cup B)\}, \end{aligned}$$

donde $I_a(x_0, u_0) = \{\alpha \in A \mid I_\alpha(x_0, u_0) = 0\}$, $\Gamma_0(\lambda) = \{\alpha \in A \mid \lambda_\alpha = 0\}$ y $\Gamma_+(\lambda) = \{\alpha \in A \mid \lambda_\alpha > 0\}$.

La condición clásica de segundo orden para $CP(S)$ que aparece en [18, p. 285, Teorema 9.1] puede ser reescrita de la siguiente manera.

Teorema 5.2.3. *Supongamos que (x_0, u_0) es solución local de $CP(S)$ y es normal fuerte. Si λ y p son los únicos multiplicadores con $(x_0, u_0, \lambda, p) \in \mathcal{E}$ y definimos la función K y el funcional J por*

$$K(t, x, u) := -H(t, x, u, p(t)) - \langle \dot{p}(t); x\rangle, \quad J(x, u) := \int_a^b K(t, x(t), u(t))dt + C,$$

donde H es como en el Teorema 5.2.1 con $\lambda_0 = 1$ y $C = \langle p(b), X_b\rangle - \langle p(a); X_a\rangle$, tendremos que $J''(x_0, u_0; y, v) \geq$

0 para todo $(y,v) \in R_{S_1}(x_0,u_0)$.

Para definir normalidad relativa a S de manera semejante a la Definición 3.3.5, fijemos $(x_0,u_0) \in S$ y definamos los siguientes funcionales lineales en Z

$$F_\gamma(y,v) = I'_\gamma(x_0,u_0;y,v) \quad (\gamma = 1,\ldots,m),$$

$$F_{\gamma+i}(y,v) = y^i(b) \quad (i = 1,\ldots,n).$$

Notemos que la Definición 5.2.2, por el Lema 1.3.1, es equivalente a la independencia lineal de estos funcionales en el espacio vectorial real definido por

$$\mathbf{Y}_a = \{(y,v) \in Z \mid \dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t),\ y(a) = 0\},$$

por lo tanto, permitiendo de nuevo índices inactivos, una forma natural de extender la Definición 3.3.5 es la siguiente.

Definición 5.2.4. *Diremos que* $(x_0,u_0) \in S$ *es normal relativo a* S *si la única solución* $\lambda \in \mathbf{R}^{m+n}$ *al sistema*

S-1) $\lambda_\alpha \geq 0,\ \lambda_\alpha I_\alpha(x_0,u_0) = 0\ (\alpha \in A)$,

S-2) $\sum_{\gamma=1}^{m+n} \lambda_\gamma F_\gamma(y,v) = 0$ *para todo* $(y,v) \in \mathbf{Y}_a$,

es $\lambda = 0$.

Con esta definición, es natural preguntarnos si podemos extender el Teorema 3.3.7 al caso de control. En las siguientes secciones demostraremos que es posible llevar acabo esta tarea. Explícitamente, el resultado que queremos demostrar es el siguiente.

Teorema 5.2.5. *Supongamos que* (x_0,u_0) *es solución local de* $CP(S)$. *Si* λ *y* p *son multiplicadores con* $(x_0,u_0,\lambda,p) \in \mathcal{E}$, (x_0,u_0) *es normal relativo a* $S_1(\lambda)$ *y definimos* J *como en el teorema anterior, entonces tendremos que* $J''(x_0,u_0;y,v) \geq 0$ *para todo* $(y,v) \in R_{S_1}(x_0,u_0)$.

5.3. Dos Nociones Equivalentes de Normalidad Débil

Recordemos que los elementos del conjunto de extremos $\mathcal{E}$ son las ternas $(x,u,\lambda,p) \in S \times \mathbf{R}^m \times \mathbf{X}$ que satisfacen los incisos (1), (2) y (5) del Teorema 5.2.1 con $\lambda_0 = 1$. Explícitamente estas condiciones son

$\mathcal{E}$-1) $\lambda_\alpha \geq 0,\ \lambda_\alpha I_\alpha(x,u) = 0\ (\alpha \in A)$,

$\mathcal{E}$-2) $\dot{p}(t) = -f_x^*(\tilde{x}(t))p(t) + L_x^*(\tilde{x}(t)) + \sum_{\gamma=1}^m \lambda_\gamma L_{\gamma x}^*(\tilde{x}(t))$,

$\mathcal{E}$-3) $f_u^*(\tilde{x}(t))p(t) = L_u^*(\tilde{x}(t)) + \sum_{\gamma=1}^m \lambda_\gamma L_{\gamma u}^*(\tilde{x}(t))$.

En gran parte de la literatura de suficiencia para distintos problemas de control ([10, 17, 32] por mencionar algunos), la normalidad se define en términos de la solución de un sistema equivalente al de arriba cuando el multiplicador de la función de costo λ_0 es igual a cero. Una definición análoga a estas sería la siguiente.

Definición 5.3.1. *Diremos que* $(x_0, u_0) \in S$ *es d-normal si la única solución* $(\lambda, p) \in \mathbf{R}^m \times \mathbf{X}$ *al sistema*

d-1) $\lambda_\alpha \geq 0,\ \lambda_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha \in A)$,

d-2) $\dot{p}(t) = -f_x^*(\tilde{x}_0(t))p(t) + \sum_{\gamma=1}^m \lambda_\gamma L_{\gamma x}^*(\tilde{x}_0(t))$,

d-3) $f_u^*(\tilde{x}_0(t))p(t) = \sum_{\gamma=1}^m \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t))$,

es $p \equiv 0$ *y* $\lambda = 0$.

Por lo tanto nos preguntamos cómo se relaciona esta nueva definición con la Definición 5.2.4 de normalidad relativa a S. A continuación veremos que, en efecto, son equivalentes.

Teorema 5.3.2. *Sea* $(x_0, u_0) \in S$*. Los siguientes enunciados son equivalentes.*

1. (x_0, u_0) *es normal relativo a* S.

2. (x_0, u_0) *es d-normal.*

Demostración. Supongamos primero que (x_0, u_0) es normal relativo a S y tomemos $(\lambda, p) \in \mathbf{R}^m \times \mathbf{X}$ que satisface (d-1), (d-2) y (d-3). Queremos probar que $(\lambda, p) = (0, 0)$. Sea $(y, v) \in \mathbf{Y}_a$ arbitrario, al multiplicar por la izquierda a (d-2) por $y^*(t)$, a (d-3) por $v^*(t)$, y despejando la sumatoria en (d-2) llegamos a las ecuaciones

$$y^*(t) \sum_{\gamma=1}^m \lambda_\gamma L_{\gamma x}^*(\tilde{x}_0(t)) = y^*(t)\dot{p}(t) + y^*(t) f_x^*(\tilde{x}_0(t))p(t), \tag{5.3.1}$$

$$v^*(t) \sum_{\gamma=1}^m \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t)) = v^*(t) f_u^*(\tilde{x}_0(t))p(t). \tag{5.3.2}$$

Integrando ambas ecuaciones en $[a,b]$ y luego sumándolas obtenemos que

$$
\begin{aligned}
&\int_a^b \{y^*(t)\dot{p}(t) + y^*(t)f_x^*(\tilde{x}_0(t))p(t) + v^*(t)f_u^*(\tilde{x}_0(t))p(t)\}dt \\
&\qquad = \int_a^b \{y^*(t)\sum_{\gamma=1}^m \lambda_\gamma L_{\gamma x}^*(\tilde{x}_0(t)) + v^*(t)\sum_{\gamma=1}^m \lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t))\}dt \\
&\qquad = \sum_{\gamma=1}^m \int_a^b \{y^*(t)\lambda_\gamma L_{\gamma x}^*(\tilde{x}_0(t)) + v^*(t)\lambda_\gamma L_{\gamma u}^*(\tilde{x}_0(t))\}dt \\
&\qquad = \sum_{\gamma=1}^m \lambda_\gamma \int_a^b \{L_{\gamma x}(\tilde{x}_0(t))y(t) + L_{\gamma u}(\tilde{x}_0(t))v(t)\}dt \\
&\qquad = \sum_{\gamma=1}^m \lambda_\gamma F_\gamma(y,v).
\end{aligned}
$$

Por otro lado podemos reescribir la primera integral como

$$
\begin{aligned}
&\int_a^b \{y^*(t)\dot{p}(t) + y^*(t)f_x^*(\tilde{x}_0(t))p(t) + v^*(t)f_u^*(\tilde{x}_0(t))p(t)\}dt \\
&\qquad = \int_a^b \{y^*(t)\dot{p}(t) + [y^*(t)f_x^*(\tilde{x}_0(t)) + v^*(t)f_u^*(\tilde{x}_0(t))]p(t)\}dt \\
&\qquad = \int_a^b \{y^*(t)\dot{p}(t) + \dot{y}^*(t)p(t)\}dt \\
&\qquad = \int_a^b \frac{d}{dt}\langle y(t); p(t)\rangle dt \\
&\qquad = \sum_{i=1}^n \{y^i(b)p^i(b) - y^i(a)p^i(a)\} = \sum_{i=1}^n y^i(b)p^i(b),
\end{aligned}
$$

ya que, como $(y,v) \in \mathbf{Y}_a$, $y^i(a) = 0$ para $i = 1,\dots,n$. Entonces, si definimos $\lambda_{m+i} := -p^i(b)$ $(i = 1,\dots,n)$, tendremos por lo anterior que

$$
0 = \sum_{\gamma=1}^m \lambda_\gamma F_\gamma(y,v) - \sum_{i=1}^n y^i(b)p^i(b) = \sum_{\gamma=1}^{m+n} \lambda_\gamma F_\gamma(y,v)
$$

para todo $(y,v) \in \mathbf{Y}_a$, por lo que $\lambda \in \mathbf{R}^{m+n}$ satisface (S-2). Como (d-1) es exactamente (S-1) y (x_0,u_0) es normal relativo a S, se tiene que $\lambda = 0$. Es decir, $\lambda_\gamma = 0$ $(\gamma = 1,\dots,m)$ y $p^i(b) = 0$ $(i = 1,\dots,n)$. Como esto transforma la ecuación (d-2) en una ecuación lineal homogénea de la cual p es solución, el inciso (2) del Teorema 1.3.7 nos garantiza que $p \equiv 0$.

Supongamos ahora que (x_0,u_0) es d-normal y $\lambda \in \mathbf{R}^{m+n}$ satisface (S-1) y (S-2). Definamos p como la única solución de la ecuación diferencial lineal (d-2) con condición final $p^i(b) = -\lambda_{m+i}$ $(i = 1, \ldots, n)$. De nuevo, multiplicando por la izquierda a (d-2) por $y^*(t)$ e integrando llegamos a la ecuación (5.3.1). Si $(y, v) \in \mathbf{Y}_a$ por (S-2) y lo anterior tendremos que

$$\begin{aligned}
0 = \sum_{\gamma=1}^{m+n} \lambda_\gamma F_\gamma(y,v) &= \int_a^b \sum_{\gamma=1}^{m} \{y^*(t)\lambda_\gamma L^*_{\gamma x}(\tilde{x}_0(t)) + v^*(t)\lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t))\}dt + \sum_{i=1}^{n} \lambda_{m+i} y^i(b) \\
&= \int_a^b \{y^*(t)\dot{p}(t) + y^*(t) f^*_x(\tilde{x}_0(t))p(t)\}dt + \int_a^b v^*(t) \sum_{\gamma=1}^{m} \lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t))dt - \sum_{i=1}^{n} p^i(b)y^i(b) \\
&= \int_a^b \{y^*(t)\dot{p}(t) + [\dot{y}^*(t) - v^*(t) f^*_u(\tilde{x}_0(t))]p(t)\}dt + \int_a^b v^*(t) \sum_{\gamma=1}^{m} \lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t))dt - \sum_{i=1}^{n} p^i(b)y^i(b) \\
&= \int_a^b \{y^*(t)\dot{p}(t) + \dot{y}^*(t)p(t)\}dt - \int_a^b v^*(t) f^*_u(\tilde{x}_0(t))p(t)dt + \int_a^b v^*(t) \sum_{\gamma=1}^{m} \lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t))dt - \sum_{i=1}^{n} p^i(b)y^i(b) \\
&= \int_a^b v^*(t) \left[\sum_{\gamma=1}^{m} \lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t)) - f^*_u(\tilde{x}_0(t))p(t)\right] dt.
\end{aligned}$$

Como para cada $v \in \mathcal{U}_p$ existe $y \in \mathbf{X}$ de forma que $(y, v) \in \mathbf{Y}_a$, tendremos que

$$0 = \int_a^b v^*(t) \left[\sum_{\gamma=1}^{m} \lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t)) - f^*_u(\tilde{x}_0(t))p(t)\right] dt$$

para todo $v \in \mathcal{U}_p$, lo que implica la igualdad

$$\sum_{\gamma=1}^{m} \lambda_\gamma L^*_{\gamma u}(\tilde{x}_0(t)) = f^*_u(\tilde{x}_0(t))p(t)$$

para todo $t \in [a, b]$. Esto es precisamente (d-3) y, como $(\lambda_1, \ldots, \lambda_m)$ y p también satisfacen (d-1) y (d-2), tendremos que $\lambda_\gamma = 0$ $(\gamma = 1, \ldots, m)$ y $p \equiv 0$, lo que a su vez implica que $\lambda = (\lambda_1, \ldots, \lambda_{m+n}) = 0$. □

Notemos que la definición del conjunto de extremos $\mathcal{E}$, en principio, no es análoga a la definición del conjunto de extremos para el caso en cálculo de variaciones. Una extensión natural de esta última al caso de control sería la siguiente.

Definición 5.3.3. *Definamos al conjunto $\bar{\mathcal{E}}$ como las parejas $(x_0, u_0, \lambda) \in S \times \mathbf{R}^{m+n}$ de forma que*

$\bar{\mathcal{E}}$-1) $\lambda_\alpha \geq 0$, $\lambda_\alpha I_\alpha(x_0, u_0) = 0$ $(\alpha \in A)$,

$\bar{\mathcal{E}}$-2) Si $G := I'(x_0, u_0; \cdot) + \sum_{\gamma=1}^{m+n} \lambda_\gamma F_\gamma$, entonces $G(y, v) = 0$ para todo $(y, v) \in \mathbf{Y}_a$.

Cambiar el sistema de ecuaciones (d) por $(\mathcal{E})$ y el sistema (S) por $(\bar{\mathcal{E}})$ en la demostración del Teorema 5.3.2 probará el siguiente resultado.

Teorema 5.3.4. *Si $(x_0, u_0) \in S$ entonces los siguientes enunciados son verdaderos.*

1. *Si $(x_0, u_0, \lambda, p) \in \mathcal{E}$ entonces $(x_0, u_0, (\lambda, -p^1(b), \ldots, -p^n(b))) \in \bar{\mathcal{E}}$.*
2. *Si $(x_0, u_0, \lambda) \in \bar{\mathcal{E}}$ entonces existe $p \in \mathbf{X}$ con $p^i(b) = -\lambda_{m+i}$ $(i = 1, \ldots, n)$ y de forma que $(x_0, u_0, (\lambda_1, \ldots, \lambda_m), p) \in \mathcal{E}$.*

El Teorema 5.3.4 nos permite trabajar indistintamente con cualquiera de los conjuntos de extremos.

Recordemos que, para $(x_0, u_0, \lambda, p) \in \mathcal{E}$, el funcional J definido en $\mathbf{X}$ que satisface la condición clásica de segundo orden (ver Teorema 5.2.3) está definido por

$$\begin{aligned}
J(x,u) &= \int_a^b K(t, x(t), u(t))dt + C \\
&= \int_a^b \left\{ L(t, x(t), u(t)) + \sum_{\gamma=1}^{m} \lambda_\gamma L_\gamma(t, x(t), u(t)) - \langle p(t); f(t, x(t), u(t))\rangle - \langle \dot{p}(t); x(t)\rangle \right\} dt + C \\
&= \int_a^b \{-H(t, x(t), u(t), \lambda, p(t)) + \langle \dot{p}(t), x\rangle\}dt + C,
\end{aligned}$$

donde $C = \langle p(b); X_b\rangle - \langle p(a); X_a\rangle$. Si $(x, u) \in S$ se puede verificar fácilmente que

$$J(x,u) = I(x,u) + \sum_{\gamma=1}^{m} \lambda_\gamma I_\gamma(x,u). \tag{5.3.3}$$

Tenemos la siguiente proposición que será fundamental más adelante.

Proposición 5.3.5. *Sea $(x_0, u_0, \lambda, p) \in \mathcal{E}$. Entonces $J'(x_0, u_0; y, v) = 0$ para todo $(y, v) \in \mathbf{X} \times \mathcal{U}_p$.*

Demostración. De la definición de J tenemos que

$$J'(x_0, u_0; y, v) = \int_a^b \{K_x(\tilde{x}_0(t))y(t) + K_u(\tilde{x}_0(t))v(t)\}dt,$$

pero como $(x_0, u_0, \lambda, p) \in \mathcal{E}$ se dan las igualdades

$$K_x(\tilde{x}_0(t)) = -H_x(t, x_0(t), u_0(t), \lambda, p(t)) - \dot{p}(t) = 0, \quad K_u(\tilde{x}_0(t)) = -H_u(t, x_0(t), u_0(t), \lambda, p(t)) = 0$$

con lo que el resultado queda probado. □

5.4. Las Variaciones Como Diferenciales

En esta sección demostraremos que la primera y segunda variación definidas en la sección anterior son las diferenciales de primer y segundo orden de los funcionales I, I_γ para una norma adecuada.

Consideremos a $Z = \mathbf{X} \times \mathcal{U}_p$ como espacio vectorial real normado, con la norma definida por

$$\|(x,u)\| = \|x\| + \|u\|,$$

donde la norma $\|x\|$ en $\mathbf{X}$ es la norma débil de la Definición 1.2.1, y la norma $\|u\|$ en $\mathcal{U}_p$ es la norma del supremo

$$\|u\| = \sup_{a \le t \le b} |u(t)|.$$

Nota 5.4.1. Es claro que con esta norma una sucesión $\{(x_k, u_k)\}_{k=1}^{\infty} \subset Z$ converge a un elemento $(x,u) \in Z$ si y solo si x_k converge a x en la norma de $\mathbf{X}$ y u_k converge a u en la norma de $\mathcal{U}_p$.

La siguiente proposición es fundamental para extender lo que hemos hecho en los capítulos anteriores.

Teorema 5.4.2. *Sea I el funcional integral en Z definido por $I(x,u) = \int_a^b L(t, x(t), u(t))dt$. Entonces para $(x_0, u_0) \in Z$, los funcionales lineales $I'(x_0, u_0; \cdot)$ e $I''(x_0, u_0; \cdot)$ definidos anteriormente, son la primera y segunda diferencial del funcional I con respecto de la norma en Z, es decir, se cumplen las fórmulas de Taylor de primer y segundo orden*

$$I(x,u) = I(x_0,u_0) + I'(x_0,u_0; x - x_0, u - u_0) + R_1(x - x_0, u - u_0)$$

e

$$I(x,u) = I(x_0,u_0) + I'(x_0,u_0; x - x_0, u - u_0) + \frac{1}{2} I''(x_0,u_0; x - x_0, u - u_0) + R_2(x - x_0, u - u_0),$$

donde R_1 y R_2 satisfacen

$$\lim_{(x,u)\to(x_0,u_0)} \frac{R_1(x - x_0, u - u_0)}{\|(x - x_0, u - u_0)\|} = 0 \quad y \quad \lim_{(x,u)\to(x_0,u_0)} \frac{R_2(x - x_0, u - u_0)}{\|(x - x_0, u - u_0)\|^2} = 0.$$

Demostraremos esta proposición a través de varios lemas, generalizando el método de [18, Capítulo 2, Sección 6] para mostrar lo mismo en el caso de cálculo de variaciones. Para esto fijamos un punto $(x_0, u_0) \in Z$ y definimos el conjunto $\mathcal{F}$ por

$$\mathcal{F} := \{(t, x_0(t), u_0(t)) \mid t \in [a,b]\}.$$

Nota 5.4.3. Recordemos que seguimos usando la convención de representar a los puntos $u_0(t+)$ y $u_0(t-)$ únicamente por $u_0(t)$ en un punto esquina t de u_0. Esto nos garantiza que $\mathcal{F}$ es un conjunto cerrado. Además, como $(x_0, u_0) \in Z$, $\mathcal{F}$ también será acotado, por lo que es un conjunto compacto.

Lema 5.4.4. *Sea $\mathcal{S}$ vecindad abierta y acotada de $\mathcal{F}$. Entonces existe $\rho > 0$ de forma que la ρ-vecindad de $\mathcal{F}$ está contenida en $\mathcal{S}$. En particular si $\|(x, u) - (x_0, u_0)\| < \rho$ entonces $(t, x(t), u(t)) \in \mathcal{S}$.*

Demostración. Supongamos lo contrario, es decir, para cada $\rho > 0$ existe $t_\rho \in [a, b]$ de forma que la ρ-vecindad de $\tilde{x}_0(t)$ no está contenida en $\mathcal{S}$. De esto podemos construir una sucesión $\{\tilde{x}_0(t_k)\}_{k=1}^{\infty} \subset \mathcal{F}$ de forma que existe otra sucesión $\{h_k\}_{k=1}^{\infty} \subset \mathcal{S}^c$ con

$$|\tilde{x}_0(t_k) - h_k| < \frac{1}{k}.$$

Como $\{\tilde{x}_0(t_k)\}_{k=1}^{\infty}$ es una sucesión en un compacto, tiene una subsucesión que converge a un elemento $f \in \mathcal{F}$, a la que por simplicidad llamaremos del mismo modo. De esto obtenemos que

$$f = \lim_{k\to\infty} \tilde{x}_0(t_k) = \lim_{k\to\infty} h_k,$$

pero como $\{h_k\}_{k=1}^{\infty}$ es una sucesión en el conjunto cerrado $\mathcal{S}^c$, el límite debe de estar también en ese conjunto, lo que contradice el hecho que $f \in \mathcal{F}$. □

Lema 5.4.5. *Representemos por P a L o a alguna de sus derivadas parciales de orden menor o igual que 2. Entonces*

$$\lim_{(x,u)\to(x_0,u_0)} P(t, x(t), u(t)) = P(t, x_0(t), u_0(t))$$

uniformemente en $a \leq t \leq b$.

Demostración. Tomemos al conjunto $\mathcal{S}$ como una vecindad abierta y acotada de $\mathcal{F}$. Entonces $\bar{\mathcal{S}}$, la cerradura de $\mathcal{S}$, es un conjunto compacto. Como P es continua, es uniformemente continua en $\bar{\mathcal{S}}$, es decir, para todo $\epsilon > 0$ existe $\delta > 0$ de forma que

$$|P(t, x, u) - P(s, v, w)| < \epsilon \quad \text{si} \quad |(t, x, u) - (s, v, w)| < \delta \;\text{ y }\; (t, x, u), (s, v, w) \in \bar{\mathcal{S}}.$$

Sean ϵ y δ como arriba y ρ como en el lema anterior. Tomemos $\eta := \min\{\delta, \rho\}$ y $(x, u) \in Z$ de forma que

$\|(x,u)-(x_0,u_0)\|<\eta$. Entonces para todo $t\in[a,b]$ tendremos que

$$\begin{aligned}|(t,x(t),u(t))-(t,x_0(t),u_0(t))| &= |(x(t)-x_0(t),u(t)-u_0(t))|\\ &\le |x(t)-x_0(t)|+|u(t)-u_0(t)|\\ &\le \|x-x_0\|+\|u-u_0\|\\ &= \|(x,u)-(x_0,u_0)\|<\delta\end{aligned}$$

y $(t,x(t),u(t))\in\bar{\mathcal{S}}$ para todo $t\in[a,b]$. Por lo tanto para todo $t\in[a,b]$ se cumple la desigualdad

$$|P(t,x(t),u(t))-P(t,x_0(t),u_0(t))|<\epsilon,$$

que es lo que queríamos demostrar. □

Lema 5.4.6. *Para todo $\epsilon>0$ existe $\delta>0$ de forma que*

$$|I'(x_0,u_0;y,v)-I'(x,u;y,v)|<\epsilon\|(y,v)\|$$

y

$$|I''(x_0,u_0;y,v)-I''(x,u;y,v)|<\epsilon\|(y,v)\|^2,$$

para todo $(y,v)\in Z$ y (x,u) con $\|(x,u)-(x_0,u_0)\|<\delta$.

Demostración. Por el lema anterior, dado $\epsilon>0$ existen $\delta_i>0$ $(i=1,\dots,n)$, de forma que, para todo $t\in[a,b]$ se cumple la desigualdad

$$|f_{x^i}(t,x(t),u(t))-f_{x^i}(t,x_0(t),u_0(t))|<\frac{\epsilon}{n(b-a)},$$

si $\|(x,u)-(x_0,u_0)\|<\delta_i$. Similarmente existen $\delta_j>0$ $(j=1,\dots,p)$ de forma que, para todo $t\in[a,b]$, se cumple la desigualdad

$$|f_{u^j}(t,x(t),u(t))-f_{u^j}(t,x_0(t),u_0(t))|<\frac{\epsilon}{p(b-a)},$$

si $\|(x,u)-(x_0,u_0)\|<\delta_j$. Definamos $\delta:=\min\{\delta_i,\delta_j\mid i=1,\dots,n;\ j=1,\dots,p\}$ y tomemos a $(x,u)\in Z$ con $\|(x,u)-(x_0,u_0)\|<\delta$ y a $(y,v)\in Z$ arbitrario. Recordando que $\tilde{x}(t)$ y $\tilde{x}_0(t)$ representan las ternas $(t,x(t),u(t))$ y $(t,x_0(t),u_0(t))$ respectivamente tendremos que

$$
\begin{aligned}
|I'(x_0,u_0;y,v)-I'(x,u;y,v)| &= \left|\int_a^b \{[L_x(\tilde{x}(t))-L_x(\tilde{x}_0(t))]y(t)+[L_u(\tilde{x}(t))-L_u(\tilde{x}_0(t))]v(t)\}dt\right| \\
&\le \sum_{i=1}^{n}\int_a^b |L_{x^i}(\tilde{x}(t))-L_{x^i}(\tilde{x}_0(t))||y_i(t)|dt \\
&\quad +\sum_{j=1}^{p}\int_a^b |L_{u^j}(\tilde{x}(t))-L_{u^j}(\tilde{x}_0(t))||v_j(t)|dt \\
&< \sum_{i=1}^{n}\int_a^b \frac{\epsilon}{n(b-a)}\|y\|dt+\sum_{j=1}^{p}\int_a^b \frac{\epsilon}{p(b-a)}\|v\|dt \\
&= n\frac{\epsilon}{n(b-a)}\|y\|(b-a)+p\frac{\epsilon}{p(b-a)}\|v\|(b-a) \\
&= \epsilon(\|y\|+\|v\|)=\epsilon\|(y,v)\|,
\end{aligned}
$$

con lo que demostramos la primera desigualdad.

Para mostrar la segunda desigualdad, definamos la norma de una matriz $A=(a_{ij})$ por

$$|A|=\sum_{i,j}|a_{ij}|.$$

Con esta definición es claro que si b es un vector que se puede multiplicar por la derecha, entonces $|Ab|\le|A||b|$. Por el lema anterior existen δ_1,δ_2 y δ_3 números positivos de forma que, para todo $t\in[a,b]$

$$|f_{xx}(\tilde{x}(t))-f_{xx}(\tilde{x}_0(t))|<\frac{\epsilon}{b-a}\quad\text{si}\quad \|(x,u)-(x_0,u_0)\|<\delta_1,$$

$$|f_{xu}(\tilde{x}(t))-f_{xu}(\tilde{x}_0(t))|<\frac{\epsilon}{b-a}\quad\text{si}\quad \|(x,u)-(x_0,u_0)\|<\delta_2$$

y

$$|f_{uu}(\tilde{x}(t))-f_{uu}(\tilde{x}_0(t))|<\frac{\epsilon}{b-a}\quad\text{si}\quad \|(x,u)-(x_0,u_0)\|<\delta_3.$$

Para ver que existe δ_1 con estas propiedades, usamos nuevamente el lema anterior para asegurar la existencia de $\delta_{ij}>0$ $(i,j=1,\dots,n)$ de forma que

$$|f_{x^ix^j}(\tilde{x}(t))-f_{x^ix^j}(\tilde{x}_0(t))|<\frac{\epsilon}{n^2(b-a)}\quad\text{si}\quad \|(x,u)-(x_0,u_0)\|<\delta_{ij},$$

por lo que, si definimos $\delta_1:=\min\{\delta_{ij}\mid i,j=1,\dots,n\}$ y tomamos $\|(x,u)-(x_0,u_0)\|<\delta_1$, tendremos que para todo $t \in [a, b]$

$$|f_{xx}(\tilde{x}(t)) - f_{xx}(\tilde{x}_0(t))| = \sum_{i,j}^{n} |f_{x^i x^j}(\tilde{x}(t)) - f_{x^i x^j}(\tilde{x}_0(t))| < \frac{\epsilon}{b-a}.$$

La existencia de δ_2 y δ_3 se pueden verificar de forma similar. Ahora definamos $\delta := \text{mín}\{\delta_1, \delta_2, \delta_3\}$, $M_1(t) := f_{xx}(\tilde{x}(t)) - f_{xx}(\tilde{x}_0(t))$, $M_2(t) := f_{xu}(\tilde{x}(t)) - f_{xu}(\tilde{x}_0(t))$, $M_3(t) := f_{uu}(\tilde{x}(t)) - f_{uu}(\tilde{x}_0(t))$ y tomemos (x, u) con $\|(x, u) - (x_0, u_0)\| < 0$ y $(y, v) \in Z$ arbitrario. Entonces para todo $t \in [a, b]$ tendremos que

$$\begin{aligned}
|I''(x, u; y, v) - I''(x_0, u_0; y, v)| &= \left| \int_a^b \{\langle y(t); M_1(t)y(t)\rangle + 2\langle y(t); M_2(t)v(t)\rangle + \langle v(t); M_3(t)v(t)\rangle\} dt \right| \\
&\leq \int_a^b |\langle y(t); M_1(t)y(t)\rangle|\, dt + 2\int_a^b |\langle y(t); M_2(t)v(t)\rangle| dt \\
&\quad + \int_a^b |\langle v(t); M_3(t)v(t)\rangle| dt \\
&\leq \int_a^b |y(t)||M_1(t)y(t)| dt + 2\int_a^b |y(t)||M_2(t)v(t)| dt \\
&\quad + \int_a^b |v(t)||M_3(t)v(t)| dt \\
&\leq \int_a^b |M_1(t)||y(t)|^2 dt + 2\int_a^b |M_2(t)||v(t)||y(t)| dt \\
&\quad + \int_a^b |M_3(t)||v(t)|^2 dt \\
&< \int_a^b \frac{\epsilon}{b-a}\|y\|^2 dt + 2\int_a^b \frac{\epsilon}{b-a}\|y\|\|v\| dt + \int_a^b \frac{\epsilon}{b-a}\|v\|^2 dt \\
&= \epsilon(\|y\|^2 + 2\|y\|\|v\| + \|v\|^2) = \epsilon\|(y, v)\|^2,
\end{aligned}$$

con lo que demostramos la segunda desigualdad. □

Estamos listos para demostrar el Teorema 5.4.2.

Demostración del Teorema 5.4.2. Definamos la función G por

$$G(\epsilon) := I((x_0, u_0) + \epsilon(x - x_0, u - u_0))$$

para $0 \leq \epsilon \leq 1$. Usando la fórmula de Taylor de primer y segundo grado para G alrededor de 0 y haciendo $(z, w) = (x - x_0, u - u_0)$ obtenemos que

$$I(x, u) = I(x_0, u_0) + I'(x_0, u_0; z, w) + R_1(z, w),$$

$$I(x,u) = I(x_0,u_0) + I'(x_0,u_0;z,w) + \frac{1}{2}I''(x_0,u_0;z,w) + R_2(z,w)$$

y, como f es de clase C^2 podemos tomar a los residuos de la forma

$$R_1(z,w) = \int_0^1 \{G'(\theta) - G'(0)\}d\theta$$

y

$$R_2(z,w) = \int_0^1 (1-\theta)\{G''(\theta) - G''(0)\}d\theta$$

que se convierten en

$$R_1(z,w) = \int_0^1 \{I'((x_0,u_0) + \theta(z,w);(z,w)) - I'((x_0,u_0);(z,w))\}d\theta$$

y

$$R_2(z,w) = \int_0^1 (1-\theta)\{I''((x_0,u_0) + \theta(z,w);(z,w)) - I''((x_0,u_0);(z,w))\}d\theta.$$

Por lo tanto, tomando $\epsilon > 0$ arbitrario y $\delta > 0$ como en el Lema 5.4.6 tendremos que

$$\left|\frac{|R_1(z,w)}{\|(z,w)\|}\right| < \frac{\epsilon\|(z,w)\|}{\|(z,w)\|} = \epsilon$$

y

$$\left|\frac{R_2(z,w)}{\|(z,w)\|^2}\right| < \frac{\epsilon\|(z,w)\|^2}{\|(z,w)\|^2} = \epsilon,$$

si $\|(z,w)\| = \|(x-x_0,u-u_0)\| < \delta$, lo que prueba las igualdades

$$\lim_{(x,u)\to(x_0,u_0)} \frac{R_1(x-x_0,u-u_0)}{\|(x-x_0,u-u_0)\|} = 0 \quad \text{y} \quad \lim_{(x,u)\to(x_0,u_0)} \frac{R_2(x-x_0,u-u_0)}{\|(x-x_0,u-u_0)\|^2} = 0,$$

que es lo que queríamos demostrar. □

5.5. Extensión de Resultados Previos

En esta sección veremos que algunos resultados críticos para la demostración del Teorema 3.3.7 tienen su análogo en el contexto de control.

Comencemos definiendo los conos curvilíneos y secuenciales como en la Sección 3. Aunque estas definiciones son aplicables a conjuntos arbitrarios, nosotros estamos interesados únicamente en estudiar las propiedades de los conos tangentes del conjunto de restricciones S.

Definición 5.5.1. *Sea $\mathcal{P} \subset Z$.*

1. El cono tangente curvilíneo $C_{\mathcal{P}}(x_0, u_0)$ del conjunto $\mathcal{P}$ en el punto $(x_0, u_0) \in \mathcal{P}$ está definido por

$$\begin{aligned} C_{\mathcal{P}}(x_0, u_0) = \{(y, v) \in Z \mid \exists[\delta > 0, (x(\cdot, \epsilon), u(\cdot, \epsilon)) \in \mathcal{P}\ (0 \leq \epsilon \leq \delta)] \ni x_t(\cdot,\cdot),\ x_\epsilon(\cdot,\cdot),\ x_{t\epsilon}(\cdot,\cdot),\ u(\cdot,\cdot) \text{ y} \\ u_\epsilon(\cdot,\cdot) \text{ son continuas a trozos en } t,\ (x(\cdot, 0), u(\cdot, 0)) = (x_0, u_0) \text{ y } (x_\epsilon(\cdot, 0), u_\epsilon(\cdot, 0)) = (y, v)\}. \end{aligned}$$

2. El cono tangente secuencial $T_{\mathcal{P}}(x_0, u_0)$ (o simplemente cono tangente) del conjunto $\mathcal{P}$ en el punto $(x_0, u_0) \in \mathcal{P}$ está definido por

$$T_{\mathcal{P}}(x_0, u_0) = \Bigg\{(y, v) \mid \exists[\{t_k\}_{k=1}^{\infty} \subset \mathbf{R}^+,\ \{(x_k, u_k)\}_{k=1}^{\infty} \subset \mathcal{P}] \ni t_k \to 0 \text{ y } \frac{(x_k, u_k) - (x_0, u_0)}{t_k} \to (y, v)\Bigg\}.$$

Proposición 5.5.2. *Si $(x_0, u_0) \in S$ entonces $C_S(x_0, u_0) \subset R_S(x_0, u_0)$ y $T_S(x_0, u_0) \subset R_S(x_0, u_0)$.*

Demostración. Demostremos la primera contención. Sean $(y, v) \in C_S(x_0, u_0)$, $\delta > 0$ y $(x(\cdot, \epsilon), u(\cdot, \epsilon)) \in S$ como en el inciso (1) de la definición anterior. Primero notemos que, como $(x(\cdot, \epsilon), u(\cdot, \epsilon)) \in S$, las funciones $x(a, \epsilon)$ y $x(b, \epsilon)$ son constantes para todo $0 \leq \epsilon \leq \delta$, por lo que $y(a) = x_\epsilon(a, 0) = 0 = x_\epsilon(b, 0) = y(b)$. Además, al derivar con respecto de ϵ y evaluar en $\epsilon = 0$ la igualdad $\dot{x}(t, \epsilon) = f(t, x(t, \epsilon), u(t, \epsilon))$ obtenemos que

$$\dot{x}_\epsilon(t, 0) = f_x(\tilde{x}_0(t))x_\epsilon(t, 0) + f_u(\tilde{x}_0(t))u_\epsilon(t, 0),$$

y, como nuestras hipótesis en $x(\cdot,\cdot)$ y sus parciales implican la igualdad de las parciales cruzadas (ver por ejemlo [11] o [25]), obtenemos que

$$\dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t),$$

por lo que $(y, v) \in \mathbf{Y}$. Ahora definamos las funciones $g_\gamma(\epsilon) := I_\gamma(x(\cdot, \epsilon), u(\cdot, \epsilon))$ $(\gamma \in I_a(x_0, u_0) \cup B)$. Si $\gamma \in I_a(x_0, u_0)$ entonces g_γ tiene un máximo local en $\epsilon = 0$, por lo que

$$0 \geq g'_\gamma(0) = I'_\gamma(x_0, u_0; y, v).$$

Si $\gamma \in B$ la función g_γ es constante en $0 \leq \epsilon \leq \delta$, por lo que

$$0 = g'_\gamma(0) = I'_\gamma(x_0, u_0; y, v).$$

Con esto terminamos de probar que $C_S(x_0, u_0) \subset R_S(x_0, u_0)$. Para probar la segunda contención tomemos $(y, v) \in T_S(x_0, u_0)$, $\{t_k\}_{k=1}^{\infty} \subset \mathbf{R}^+$ y $\{(x_k, u_k)\}_{k=1}^{\infty} \subset S$ como en el inciso (2) de la definición anterior. Como $(x_k - x_0, u_k - u_0)/t_k$ tiende a (y, v) en la norma de Z, en particular tendremos que

$$\frac{x_k(a) - x_0(a)}{t_k} \to y(a) \ \text{ y } \ \frac{x_k(b) - x_0(b)}{t_k} \to y(b)$$

y, como x_k y x_0 tienen los mismos puntos extremos para todo k, $y(a) = 0 = y(b)$. Fijando $(t, x_0(t), u_0(t))$, por el desarrollo de Taylor de primer orden de f alrededor de $(x_0(t), u_0(t))$, existe R_1 de forma que

$$\begin{aligned} f(t, x_k(t), u_k(t)) = f(\tilde{x}_0(t)) + f_x(\tilde{x}_0(t))(x_k(t) - x_0(t)) + f_u(\tilde{x}_0(t))(u_k(t) - u_0(t)) \\ + R_1(x_k(t) - x_0(t), u_k(t) - u_0(t)), \end{aligned}$$

y R_1 satisface

$$\lim_{(x,u)\to(x_0(t),u_0(t))} \frac{R_1(x - x_0(t), u - u_0(t))}{|(x - x_0(t), u - u_0(t))|} = 0.$$

Entonces podemos escribir

$$\begin{aligned} \frac{\dot{x}_k(t) - \dot{x}_0(t)}{t_k} &= \frac{f(t, x_k(t), u_k(t)) - f(t, x_0(t), u_0(t))}{t_k} \\ &= f_x(\tilde{x}_0(t)) \left[\frac{x_k(t) - x_0(t)}{t_k}\right] + f_u(\tilde{x}_0(t)) \left[\frac{u_k(t) - u_0(t)}{t_k}\right] \\ &\quad + \frac{R_1(x_k(t) - x_0(t), u_k(t) - u_0(t))}{t_k} \end{aligned}$$

y, tomando el límite cuando k tiende a infinito en ambos lados de la ecuación obtenemos que

$$\dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t),$$

ya que por definición de convergencia en Z, $\dfrac{x_k - x_0}{t_k} \to y$, $\dfrac{\dot{x}_k - \dot{x}_0}{t_k} \to \dot{y}$ y $\dfrac{u_k - u_0}{t_k} \to v$ de manera uniforme en $[a,b]$, además

$$\lim_{k\to\infty} \frac{R_1(x_k(t)-x_0(t),u_k(t)-u_0(t))}{t_k} = \lim_{k\to\infty} \frac{R_1(x_k(t)-x_0(t),u_k(t)-u_0(t))}{|(x_k(t)-x_0(t),u_k(t)-u_0(t))|}\frac{|(x_k(t)-x_0(t),u_k(t)-u_0(t))|}{t_k}$$
$$= 0\cdot|(y(t),v(t))| = 0,$$

por lo que $(y,v)\in\mathbf{Y}$. Por otro lado, dividiendo la fórmula de Taylor de primer orden del Teorema 5.4.2 entre t_k tendremos que, para cada $\gamma\in I_a(x_0,u_0)\cup B$

$$\frac{I_\gamma(x_k,u_k)-I_\gamma(x_0,u_0)}{t_k} = I'_\gamma\left(x_0,u_0;\frac{x_k-x_0}{t_k},\frac{u_k-u_0}{t_k}\right) + \frac{R_1(x_k-x_0,u_k-u_0)}{t_k}$$

y, como $(x_k,u_k)\in S$ para toda k, tomando el límite cuando k tiende a infinito de ambos lados de la igualdad obtenemos que

$$0 \geq \lim_{k\to\infty}\frac{I_\gamma(x_k,u_k)}{t_k} = I'_\gamma(x_0,u_0,y,v)$$

si $\gamma\in I_a(x_0,u_0)$ y

$$0 = \lim_{k\to\infty}\frac{I_\gamma(x_k,u_k)}{t_k} = I'_\gamma(x_0,u_0,y,v)$$

si $\gamma\in B$. Esto completa la demostración. □

Nota 5.5.3. Para llegar a las igualdades del lado derecho de los dos límites de arriba se usó la continuidad de los funcionales lineales $I'_\gamma(x_0,u_0;\cdot)$ en el espacio Z. Este hecho no es difícil de demostrar.

Esta proposición nos permite definir nociones de regularidad similares a las dadas en las Secciones 2 y 3.

Definición 5.5.4. *Sea $(x_0,u_0)\in S$.*

1. *Diremos que (x_0,u_0) es p-regular si $C_S(x_0,u_0)=R_S(x_0,u_0)$.*
2. *Diremos que (x_0,u_0) es regular si $T_S(x_0,u_0)=R_S(x_0,u_0)$.*

Veamos ahora que podemos extender inmediatamente los Teoremas 3.3.2 y 3.3.3 (Teoremas 5.5.5 y 5.5.6 respectivamente) mediante el Teorema 5.4.2.

Teorema 5.5.5. *Supongamos que $(x_0,u_0)\in S$ es regular de S y resuelve $CP(S)$ localmente. Entonces existe $\lambda\in\mathbf{R}^{m+n}$ con $(x_0,u_0,\lambda)\in\bar{\mathcal{E}}$. En particular, por el Toerema 5.3.4, existe $p\in\mathbf{X}$ de forma que $(x_0,u_0,(\lambda_1,\ldots,\lambda_m),p)\in\mathcal{E}$.*

Demostración. Tomemos $(y,v)\in T_S(x_0,u_0)$, $\{t_k\}_{k=1}^\infty\subset\mathbf{R}^+$ y $\{(x_k,u_k)\}_{k=1}^\infty\subset S$ como en el inciso (2) de la Definición 5.5.1. Para valores grandes de k tenemos que $I(x_k.u_k)\geq I(x_0,u_0)$. Por el Teorema 5.4.2

$$0 \leq \lim_{i\to\infty} \frac{I(x_k,u_k)-I(x_0,u_0)}{t_k} = I'(x_0,u_0;y,v),$$

por lo que la regularidad nos asegura que $I'(x_0,u_0;y,v) \geq 0$ para todo $(y,v) \in R_S(x_0,u_0)$. Pero $R_S(x_0,u_0)$ se puede reescribir como

$$R_S(x_0,u_0) = \{(y,v) \in \mathbf{Y}_a \mid F_\alpha(y,v) \leq 0 \ (\alpha \in A),\ F_\beta(y,v) = 0 \ (\beta = q+1,\ldots,m+n)\},$$

por lo que el Lema 1.3.3 nos da el resultado deseado. □

Nota 5.5.6. De la demostración del teorema anterior podemos concluir que, si (x_0,u_0) es una solución local de $CP(S)$ entonces

$$T_S(x_0,u_0) \subset R_S(x_0,u_0) \cap \{(y,v) \in \mathbf{Y} \mid I'(x_0,u_0;y,v) \geq 0\}.$$

El siguiente resultado es una generalización directa del Teorema 3.3.3 ya que este último es el caso particular cuando $f(t,x,u) = u$.

Teorema 5.5.7. *Sea $(x_0,u_0) \in S$ solución local de $CP(S)$ y $(\lambda,p) \in \mathbf{R}^m \times \mathbf{X}$ con $(x_0,u_0,\lambda,p) \in \mathcal{E}$. Entonces para J definida como en la Sección 5.3 tendremos que*

$$J''(x_0,u_0;y,v) \geq 0 \quad \textit{para todo } (y,v) \in T_{S_1}(x_0,u_0).$$

Demostración. Sean $(y,v) \in T_{S_1}(x_0,u_0), \{t_k\}_{k=1}^\infty \subset \mathbf{R}^+$ y $\{(x_k,u_k)\}_{k=1}^\infty \subset S_1(\lambda)$ como en el inciso (2) de la Definición 5.5.1. De la igualdad (5.3.3) tenemos que $I(x,u) = J(x,u)$ para todo $(x,u) \in S_1(\lambda)$. Por otro lado, del desarrollo de Taylor de segundo orden del Teorema 5.4.2 obtenemos

$$\begin{aligned}\frac{1}{2}J''(x_0,u_0;y,v) &= \lim_{k\to\infty} \frac{J(x_k,u_k)-J(x_0,u_0)-J'(x_0,u_0;x_k-x_0,u_k-u_0)}{t_k^2} \\ &= \lim_{k\to\infty} \frac{I(x_k,u_k)-I(x_0,u_0)}{t_k^2} \geq 0\end{aligned}$$

ya que, por la Proposición 5.3.5, $J'(x_0,u_0;x_k-x_0,u_k-u_0) = 0$ para todo k, mientras que para k suficientemente grande $I(x_k,u_k) \geq I(x_0,u_0)$ por ser (x_0,u_0) solución local. □

Como ocurrió en el caso en cálculo de variaciones, el teorema anterior y la afirmación «normalidad de S implica regularidad de S» aplicada a $S_1(\lambda)$ probarán el Teorema 5.2.5. La siguiente sección se encargará de probar la afirmación anterior con nuestras nuevas definiciones en control, mientras que en la última sección se da una demostración alternativa del Teorema 5.2.5 sin recurrir al Teorema 5.5.7.

Recordemos que un resultado crítico en la demostración de que normalidad implica regularidad en el caso del cálculo de variaciones fue el Teorema 3.2.4. A continuación veremos que este resultado sigue siendo cierto en control, aunque la demostración se torna más complicada.

Teorema 5.5.8. *Supongamos que S está dado únicamente por restricciones de igualdad y $(x_0, u_0) \in S$ es un arco fuertemente normal de S. Entonces $C_S(x_0, u_0) \subset T_S(x_0, u_0)$. En particular, para S definido únicamente con restricciones de igualdad, p-regularidad implica regularidad.*

Demostración. Sea (x_0, u_0) normal de

$$S = \{(x, u) \mid I_\beta(x, u) = 0 \ (\beta \in B)\}$$

con $B = \{1, \ldots, m\}$ y $(y, v) \in C_S(x_0, u_0)$. Por la demostración de [18, Teorema 7.1], podemos tomar a $(x(\cdot, \epsilon), u(\cdot, \epsilon)) \in S$ $(0 \leq \epsilon \leq \delta)$ como en la definición de $C_S(x_0, u_0)$ por

$$u(t, \epsilon) = u_0(t) + \epsilon v(t) + \sum_{\sigma=1}^{m+n} B_\sigma(\epsilon) v_\sigma(t),$$

donde los elementos $(y_\sigma, v_\sigma) \in \mathbf{Y}_a$ son como en la Definición 5.2.2 y $x(t, \epsilon)$ es la única respuesta de $u(t, \epsilon)$ (es decir la única solución de la ecuación diferencial $\dot{x}(t, \epsilon) = f(t, x(t, \epsilon), u(t, \epsilon)))$ que une los puntos extremos de x_0. Además B es de clase C^2 en $0 \leq \epsilon \leq \delta$ y $B_\sigma(0) = 0 = B'_\sigma(0)$.
Definamos las siguientes funciones con dominio $[a, b]$ para k suficientemente grande,

$$u_k := u(\cdot, 1/k) \qquad x_k := x(\cdot, 1/k)$$
$$w(\cdot, \epsilon) := \frac{u(\cdot, \epsilon) - u_0(\cdot)}{\epsilon} \qquad z(\cdot, \epsilon) := \frac{x(\cdot, \epsilon) - x_0(\cdot)}{\epsilon},$$

las constantes $t_k := 1/k$ y la sucesión $\{(z_k, w_k)\}_{k=1}^\infty \subset Z$, donde $z_k := z(\cdot, 1/k)$ y $w_k := w(\cdot, 1/k)$. Claramente (x_k, u_k) está en S para todo k, por lo que si verificamos que z_k tiende a y en la norma de $\mathbf{X}$ y w_k tiende a v en la norma de $\mathcal{U}_p$ tendremos que (y, v) estará en $T_S(x_0, u_0)$. Para probar esto se tiene que mostrar que las convergencias $z_k \to y$, $\dot{z}_k \to \dot{y}$ y $w_k \to v$ son uniformes en $[a, b]$.

Primero veamos que $w_k \to v$ uniformemente en $[a, b]$. Para esto notemos que, como $v_\sigma \in U_p$, existe $M > 0$ de forma que $|v_\sigma(t)| < M$ para todo $\sigma = 1, \ldots, m + n$ y $t \in [a, b]$. Entonces

$$\left|\frac{u_k(t)-u_0(t)}{t_k}-v(t)\right| = \left|\frac{\sum_{\sigma=1}^{m+n} B_\sigma(1/k)v_\sigma(t)}{1/k}\right| \leq \sum_{\sigma=1}^{m+n}\left|\frac{B_\sigma(1/k)}{1/k}\right||v_\sigma(t)| < \sum_{\sigma=1}^{m+n}\left|\frac{B_\sigma(1/k)}{1/k}\right| M,$$

y como

$$\lim_{k\to\infty}\left|\frac{B_\sigma(1/k)}{1/k}\right| = \lim_{k\to\infty}\left|\frac{B_\sigma(1/k)-B_\sigma(0)}{1/k}\right| = B'_\sigma(0) = 0,$$

para cada $\eta > 0$ existe $N > 0$ de forma que, si $k > N$,

$$\left|\frac{B_\sigma(1/k)}{1/k}\right| < \frac{\eta}{(m+n)M}$$

para todo $\sigma = 1, \ldots, n+q$. Entonces si $k > N$

$$|w_k(t) - v(t)| < \sum_{\sigma=1}^{m+n} \frac{\eta}{(m+n)M} M = \eta$$

para todo $t \in [a, b]$, que es lo que queríamos demostrar. Desgraciadamente a diferencia de $u(\cdot,\cdot)$, no conocemos explícitamente a $x(\cdot,\cdot)$, por lo que para demostrar que z_k tiende uniformemente a y tendremos que basarnos en teoremas de dependencia continua de parámetros para ecuaciones diferenciales ordinarias.

Restrinjamos el dominio de f al conjunto $\bar{\mathcal{S}}$, donde $\mathcal{S}$ es una ρ-vecindad del conjunto $\mathcal{F}$ definido en la sección anterior. Como $f \in C^2$ existe $M_1 > 0$ con $|f''(t, x, u)| < M_1$ para todo $(t, x, u) \in \bar{\mathcal{S}}$, ya que la segunda derivada de f

$$f'' : \bar{\mathcal{S}} \to \mathcal{L}_2(\mathbf{R}^{n+p+1} \times \mathbf{R}^{n+p+1}, \mathbf{R}^n)$$

es una función continua del compacto $\bar{\mathcal{S}}$ al espacio normado de las funciones bilineales continuas con dominio $\mathbf{R}^{n+p+1} \times \mathbf{R}^{n+p+1}$ y rango $\mathbf{R}^n$. Gracias a esto podemos usar el Teorema de Taylor con resto de Lagrange [8, p. 263, Teorema 7.24] para obtener que, para todo $(t_0, x_0, u_0) \in \mathcal{S}$ existe una función $R(t_0, x_0, u_0; \cdot)$ que satisface

$$\begin{aligned} f(t, x, u) = {} & f(t_0, x_0, u_0) + f_t(t_0, x_0, u_0)(t - t_0) + f_x(t_0, x_0, u_0)(x - x_0) + f_u(t_0, x_0, u_0)(u - u_0) \\ & + R(t_0, x_0, u_0; t - t_0, x - x_0, u - u_0) \end{aligned} \tag{5.5.1}$$

para todo $(t, x, u) \in \mathbf{R}^{n+p+1}$ y además, se satisface la desigualdad

$$|R(t_0, x_0, u_0; t - t_0, x - x_0, u - u_0)| \leq \frac{M_1}{2}|(t - t_0, x - x_0, u - u_0)|^2 \tag{5.5.2}$$

para todo $(t_0, x_0, u_0) \in \mathcal{S}, (t, x, u) \in \mathbf{R}^{n+p+1}$.

Despejando R de la igualdad (5.5.1) podemos ver que, como resulta suma y productos de funciones de clase C^1, R también lo será vista como función de $\mathcal{S} \times \mathbf{R}^{n+p+1}$ a $\mathbf{R}^n$. En particular si restringimos R al conjunto $\mathcal{S} \times \mathcal{B}$, donde $\mathcal{B} \subset \mathbf{R}^{n+p+1}$ es una ρ-vecindad abierta de 0 ($\rho > 0$), la diferencial de R estará acotada en el conjunto $\bar{\mathcal{S}} \times \bar{\mathcal{B}}$, por lo que usando [8, p. 197, Corolario 5.14] podemos concluir que R es uniformemente Lipschitz continua en $\mathcal{S} \times \mathcal{B}$, es decir existe $M_2 > 0$ con

$$|R(t_0, x_0, u_0; s_0, y_0, v_0) - R(t, x, u; s, y, v)| \leq M_2 |(t_0, x_0, u_0; s_0, y_0, v_0) - (t, x, u; s, y, v)| \tag{5.5.3}$$

para todo $(t_0, x_0, u_0; s_0, y_0, v_0)$, $(t, x, u; s, y, v) \in \mathcal{S} \times \mathcal{B}$.

De la igualdad (5.5.1), para cada $t \in [a, b]$

$$\begin{aligned} f(t, x_k(t), u_k(t)) = {} & f(t, x_0(t), u_0(t)) + f_x(t, x_0(t), u_0(t))(x_k(t) - x_0(t)) + f_u(t, x_0(t), u_0(t))(u_k(t) - u_0(t)) \\ & + R(t, x_0(t), u_0(t); 0, x_k(t) - x_0(t), u_k(t) - u_0(t)). \end{aligned} \tag{5.5.4}$$

Ahora definamos la función $G(t, y, \epsilon) : [a, b] \times \mathbf{R}^n \times [0, \delta] \to \mathbf{R}^n$ por

$$G(t, y, \epsilon) := \begin{cases} f_x(\tilde{x}_0(t))y + f_u(\tilde{x}_0(t))w(t, \epsilon) + \dfrac{1}{\epsilon} R(\tilde{x}_0(t); 0, \epsilon y, \epsilon w(t, \epsilon)) & \text{si } \epsilon \neq 0, \\ f_x(\tilde{x}_0(t))y + f_u(\tilde{x}_0(t))v(t) & \text{si } \epsilon = 0. \end{cases}$$

Veamos que esta función es continua en los conjuntos $T_i \times \mathbf{R}^n \times [0, \delta]$, donde $T_i = [t_i, t_{i+1}]$ son los subintervalos de $[a, b]$ donde las funciones u_0, v, v_σ ($\sigma = 1, \ldots, m+n$) son todas continuas. Fijemos $(t, y, 0) \in T_i \times \mathbf{R}^n \times [0, \delta]$ y tomemos (s, z, ϵ) en el mismo conjunto con $\epsilon > 0$. Entonces

$$\begin{aligned} |G(s, z, \epsilon) - G(t, y, 0)| & \leq |f_x(\tilde{x}_0(s))z - f_x(\tilde{x}_0(t))y| + |f_u(\tilde{x}_0(s))w(t, \epsilon) - f_u(\tilde{x}_0(t))v(t)| \\ & \quad + \left| \frac{1}{\epsilon} R(\tilde{x}_0(s); 0, \epsilon z, \epsilon w(s, \epsilon)) \right| \\ & \leq |f_x(\tilde{x}_0(s))z - f_x(\tilde{x}_0(t))y| + |f_u(\tilde{x}_0(s))w(t, \epsilon) - f_u(\tilde{x}_0(t))v(t)| \\ & \quad + \frac{M\epsilon |(z, w(s, \epsilon)|^2}{2}, \end{aligned}$$

donde se usó (5.5.2) para llegar a la última desigualdad. Como en el conjunto $T_i \times \mathbf{R}^n \times [0,\delta]$ las funciones

$$g_1(s,z) := f_x(\tilde{x_0}(s))z, \quad g_2(s,\epsilon) := f_u(\tilde{x}_0(s))w(s,\epsilon)$$

son continuas en (t,y) y $(t,0)$ respectivamente y el término $|(z,w(s,\epsilon))|^2$ en el último sumando es acotado cuando (s,z,ϵ) tiende a $(t,y,0)$, podemos hacer la diferencia $|G(s,z,\epsilon)-G(t,y,0)|$ tan pequeña como queramos haciendo a (s,z,ϵ) tender a $(t,y,0)$. La continuidad en los puntos (s,z,ϵ) con $\epsilon>0$ es trivial por cómo se definen los conjuntos T_i. Esto prueba la continuidad en los conjuntos antes mencionados.

Ahora notemos que, por la fórmula para $u(t,\epsilon)$, esta función tiende a $u_0(t)$ uniformemente en $[a,b]$ cuando ϵ tiende a 0, por lo que existe $0<\delta'\leq\delta$ con $\delta'\leq 1$ de forma que, si $0\leq\epsilon\leq\delta'$ entonces $\|u(\cdot,\epsilon)-u_0\|\leq\rho/2$ (ρ la misma constante para la que se satisface la desigualdad (5.4.3)). Además, definamos $V\subset\mathbf{R}^n$ como la $\rho/2$-vecindad abierta alrededor de 0. Veamos que G restringida a $T_i\times V\times[0,\delta']$ es localmente Lipschitz continua con respecto de y, es decir existe $M>0$ de forma que

$$|G(t,y,\epsilon)-G(t,z,\epsilon)|\leq M|y-z| \ \text{ para todo } \ t\in T_i,\ y,z\in V, \epsilon\in[0,\delta'].$$

Si $\epsilon>0$ entonces

$$\begin{aligned}|G(t,y,\epsilon)-G(t,z,\epsilon)| &\leq |f_x(\bar{x}_0(t))||y-z|+\frac{1}{\epsilon}\left|R(\bar{x}_0(t);0,\epsilon y,\epsilon w(t,\epsilon))-R(\bar{x}_0(t);0,\epsilon z,\epsilon w(t,\epsilon))\right|\\ &=|f_x(\bar{x}_0(t))||y-z|\\ &\quad+\frac{1}{\epsilon}\left|R(\bar{x}_0(t);0,\epsilon y,u(t,\epsilon)-u_0(t))-R(\bar{x}_0(t);0,\epsilon z,u(t,\epsilon)-u_0(t))\right|,\end{aligned}$$

mientras que si $\epsilon=0$

$$|G(t,y,0)-G(t,z,0)|=|f_x(\bar{x}_0(t))||y-z|,$$

pero

$$|(0,\epsilon y,u(t,\epsilon)-u_0(t))|\leq\epsilon|y|+|u(t,\epsilon)-u_0(t)|\leq\rho$$

$$|(0,\epsilon z,u(t,\epsilon)-u_0(t))|\leq\epsilon|z|+|u(t,\epsilon)-u_0(t)|\leq\rho$$

para todo (t, y, ϵ), $(t, z, \epsilon) \in T_i \times V \times [0, \delta']$, por lo que

$$(0, \epsilon y, u(t, \epsilon) - u_0(t)), (0, \epsilon z, u(t, \epsilon) - u_0(t)) \in \mathcal{B} \text{ para todo } (t, y, \epsilon), (t, z, \epsilon) \in T_i \times V \times [0, \delta']$$

y, como $f_x(\tilde{x}_0(t))$ es acotada en $[a, b]$ por una constante $M_3 > 0$, tendremos por (5.5.3) que

$$|G(t, y, \epsilon) - G(t, z, \epsilon)| \leq M_3 |y - z| + M_2 |y - z|$$

si $\epsilon > 0$ y

$$|G(t, y, \epsilon) - G(t, z, \epsilon)| \leq M_3 |y - z|$$

si $\epsilon = 0$. En cualquier caso $|G(t, y, \epsilon) - G(t, z, \epsilon)| \leq M|y - z|$ donde $M := M_2 + M_3$, que es precisamente lo que queríamos demostrar.

Por lo tanto podemos invocar [2, p. 110, Teorema 8.3] para asegurar que las soluciones de la ecuación diferencial

$$\dot{z}(t) = G(t, z(t), \epsilon), \quad (t, z(t), \epsilon) \in T_i \times V \times [0, \delta'],$$

dependen continuamente del parámetro ϵ. Por otro lado, como $f \in C^2$, también es localmente Lipschitz continua, por lo que podemos usar de nuevo [2, p. 110, Teorema 8.3] para asegurar que la ecuación diferencial

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

depende continuamente del parámetro ϵ. Usando esto y que $x(a, \epsilon) = X_a = x_0(a)$ para todo $0 \leq \epsilon \leq \delta'$, podemos asegurar que $x(\cdot, \epsilon)$ tiende a x_0 uniformemente en $[a, b]$ cuando ϵ tiende a 0. Esto nos garantiza que, para k suficientemente grande y $t \in [a, b]$, $(t, x_k(t) - x_0(t), t_k) \in T_i \times V \times [0, \delta']$ para algún i. De lo anterior y (5.5.4) obtenemos que

$$\begin{aligned}
\dot{z}_k(t) = \frac{\dot{x}_k(t) - \dot{x}_0(t)}{t_k} &= \frac{f(t, x_k(t), u_k(t)) - f(t, x_0(t), u_0(t))}{t_k} \\
&= f_x(\tilde{x}_0(t)) \left[\frac{x_k(t) - x_0(t)}{t_k} \right] + f_u(\tilde{x}_0(t)) \left[\frac{u_k(t) - u_0(t)}{t_k} \right] \\
&\quad + \frac{1}{t_k} R(\tilde{x}_0(t); 0, x_k(t) - x_0(t), u_k(t) - u_0(t)) \\
&= G\left(t, \frac{x_k(t) - x_0(t)}{t_k}, t_k \right),
\end{aligned}$$

por lo que z_k es solución de la ecuación diferencial $\dot{z}(t) = G(t, z(t), 1/k)$ con $(t, z_k(t), t_k) \in T_i \times V \times [0, \delta']$ para k suficientemente grande. Más aún, como cada z_k es continua en $[a, b]$ y $z_k(a) = 0$ para todo k, z_k es la única solución en $[a, b]$ de la ecuación diferencial antes mencionada con condición inicial $z_k(a) = 0$. Como y es la única solución de la ecuación diferencial $\dot{z}(t) = G(t, z(t), 0)$ en $[a, b]$ con condición inicial $y(a) = 0$, la dependencia continua del parámetro nos asegura que z_k tiende a y uniformemente en $[a, b]$.

Por último veamos que $\dot{z}_k$ tiende a $\dot{y}$ uniformemente en $[a, b]$.

$$\begin{aligned}
|\dot{z}_k(t) - \dot{y}(t)| &= |f_x(\tilde{x}_0(t))z_k(t) + f_u(\tilde{x}_0(t))w_k(t) + \frac{1}{t_k}R(\tilde{x}_0(t); 0, t_k z_k(t), t_k w_k(t)) \\
&\quad - f_x(\tilde{x}_0(t))y(t) - f_u(\tilde{x}_0(t))v(t)| \\
&\leq |f_x(\tilde{x}_0(t))||z_k(t) - y(t)| + |f_u(\tilde{x}_0(t))||w_k(t) - v(t)| \\
&\quad + \frac{1}{t_k}|R(\tilde{x}_0(t); 0, x_k(t) - x_0(t), u_k(t) - u_0(t))| \\
&\leq |f_x(\tilde{x}_0(t))||z_k(t) - y(t)| + |f_u(\tilde{x}_0(t))||w_k(t) - v(t)| \\
&\quad + \frac{M_1}{2t_k}|(x_k(t) - x_0(t), u_k(t) - u_0(t))|^2 \\
&= |f_x(\tilde{x}_0(t))||z_k(t) - y(t)| + |f_u(\tilde{x}_0(t))||w_k(t) - v(t)| \\
&\quad + \frac{M_1}{2}|(z_k(t), w_k(t))||(x_k(t) - x_0(t), u_k(t) - u_0(t))|,
\end{aligned}$$

la segunda desigualdad es por (5.5.2) y, como los términos $|f_x(\tilde{x}_0(t))|$, $|f_u(\tilde{x}_0(t))|$ y $|(z_k(t), w_k(t))|$ son acotados en $[a, b]$ y los términos restantes de la última igualdad tienden a cero uniformemente en $[a, b]$ cuando k tiende a infinito, tendremos la convergencia deseada. $\square$

5.6. Normalidad Implica Regularidad

En esta sección se extienden los resultados del Capítulo 4. Para este fin conservaremos la mayor parte de la notación en ese capítulo. Fijemos $(x_0, u_0) \in S$ y sigamos suponiendo sin pérdida de generalidad que $I_a(x_0, u_0) = A$.

Comencemos enunciando las definiciones básicas. Recordemos que $(x_0, u_0) \in S$ es llamado normal relativo de S si la única solución $\lambda \in \mathbf{R}^{m+n}$ al sistema

S-1) $\lambda_\alpha \geq 0,\ \lambda_\alpha I_\alpha(x_0, u_0) = 0\ (\alpha \in A)$,

S-2) $\sum_{\gamma=1}^{m+n} \lambda_\gamma F_\gamma(y,v) = 0$ para todo $(y,v) \in \mathbf{Y}_a$,

es $\lambda = 0$.

Definamos los conjuntos Γ y Δ como en el Capítulo 4, es decir

$$\Gamma := \{\alpha \in A \mid \exists (y,v) \in R_S(x_0,u_0) \ni F_\alpha(y,v) < 0\} \;\text{ y }\; \Delta := A - \Gamma.$$

Entonces es claro que las restricciones tangenciales $R_S(x_0,u_0)$ pueden reescribirse como

$$R_S(x_0,u_0) = \{(y,v) \in \mathbf{Y} \mid F_\alpha(y,v) \leq 0 \; (\alpha \in \Gamma),\; F_\beta(y,v) = 0 \; (\beta \in \Delta \cup B)\}$$

que son justamente las restricciones tangenciales en (x_0,u_0) del conjunto S^* definido por

$$S^* := \{(x,u) \in D \mid I_\alpha(x,u) \leq 0 \; (\alpha \in \Gamma),\; I_\beta(x,u) = 0 \; (\beta \in \Delta \cup B) \}.$$

El siguiente lema del Capítulo 3 puede llevarse al caso de control sin ninguna dificultad.

Lema 5.6.1. *Sean $\mathcal{K}, \mathcal{P} \subset Z$.*

1. *$T_\mathcal{P}(x_0)$ es un cono cerrado en $\mathbf{Y}$.*

2. *Si $x_0 \in \mathcal{K} \subset \mathcal{P}$, entonces $T_\mathcal{K}(x_0) \subset T_\mathcal{P}(x_0)$.*

Mediante el resultado anterior, podemos basarnos en la demostración del Lema 4.1.3 y obtener lo siguiente.

Lema 5.6.2. *Si (x_0,u_0) es regular de S^*, entonces también es regular de S.*

Ahora definamos al conjunto H por

$$H = \{(y,v) \in \mathbf{Y} \mid F_\alpha(y,v) < 0 \; (\alpha \in \Gamma),\; F_\beta(y,v) = 0 \; (\beta \in \Delta \cup B)\} \cup \{0\}.$$

Lema 5.6.3. *H es un cono convexo cuya cerradura $\overline{H}$ (relativa a $\mathbf{Y}$) es $R_S(x_0,u_0)$.*

Demostración. Ver que H es cono convexo es trivial. Demostremos primero que $\overline{H} \subset R_S(x_0,u_0)$. Si tomamos $(y,v) \in \overline{H}$ debe existir una sucesión $\{(y_k,v_k)\}_{k=1}^\infty \subset H$ de forma que $\lim_{k\to\infty}(y_k,v_k) = (y,v)$. Por la continuidad de los operadores F_α $(\alpha \in \Gamma)$ tenemos que

$$0 \leq \lim_{k\to\infty} F_\alpha(y_k,v_k) = F_\alpha(y.v),$$

por lo que $(y,v) \in R_{S^*}(x_0,u_0) = R_S(x_0,u_0)$.

Por otro lado, si $(y,v) \in R_S(x_0,u_0)$ y existe $0 \neq (z,w) \in H$, entonces $(y,v) + \epsilon(z,w) \in H$ para cualquier $\epsilon > 0$. Definiendo

$$(y_k, v_k) := (y,v) + \frac{1}{k}(z,w)$$

tendremos que

$$\|(y_k,v_k) - (y,v)\| = \frac{1}{k}\|(z,w)\|$$

por lo que $\lim_{k\to\infty}(y_k,v_k) = (y,v)$ y por tanto $(y,v) \in \overline{H}$. Si el único elemento en H es 0, entonces debemos tener que $\Gamma = \emptyset$ y $H = R_{S^*}(x_0,u_0)$, por lo que $\{0\} = \overline{H} = R_{S^*}(x_0,u_0) = R_S(x_0,u_0)$. □

Gracias al Lema 5.6.1 y la demostración del Lema 4.1.5 se obtiene el siguiente resultado.

Lema 5.6.4. (x_0,u_0) *es regular de* S *si y sólo si* $H \subset T_S(x_0,u_0)$. (x_0,u_0) *es regular de* S^* *si y sólo si* $H \subset T_{S^*}(x_0,u_0)$.

Sea K el conjunto dado nuevamente por

$$K := \{(y,v) \in D \mid I_\alpha(y,v) \leq 0 \ (\alpha \in \Delta),\ I_\beta(y,v) = 0 \ (\beta \in B)\},$$

y K_0 el conjunto dado por

$$K_0 := \{(y,v) \in D \mid I_\beta(y,v) = 0 \ (\beta \in \Delta \cup B)\}.$$

Lema 5.6.5. $R_K(x_0,u_0) = R_{K_0}(x_0,u_0)$.

Demostración. Si $\Gamma = \emptyset$ entonces $S = K$ y $S^* = K_0$, por lo que

$$R_K(x_0,u_0) = R_S(x_0,u_0) = R_{S^*}(x_0,u_0) = R_{K_0}(x_0,u_0).$$

Si $\Gamma \neq \emptyset$ y $\Delta = \emptyset$, entonces $K = K_0$ por lo que el resultado es trivial. Si $\Gamma \neq \emptyset$ y $\Delta \neq \emptyset$, tomemos $0 \neq (z,w) \in H$ y $(y,v) \in R_K(x_0,u_0)$ arbitrario. Como

$$F_\gamma((y,v) + h(z,w)) = F_\gamma(y,v) + hF_\gamma(z,w),$$

para todo h, podemos tomar $\bar{h} := \text{máx}_{\gamma\in\Gamma}\{|F_\gamma(y,v)|/|F_\gamma(z,w)|\}$ para obtener que

$$\begin{aligned} F_\gamma(y,v) + \bar{h}F_\gamma(z,w) &\leq F_\gamma(y,v) + \frac{|F_\gamma(y,v)|}{|F_\gamma(z,w)|}F_\gamma(z,w) \\ &= F_\gamma(y,v) - |F_\gamma(y,v)| \leq 0 \end{aligned}$$

para todo $\gamma \in \Gamma$. Además

$$F_\gamma(y,v) + \bar{h}F_\gamma(z,w) = F_\gamma(y,v) \leq 0 \;\text{ si }\; \gamma \in \Delta$$

y

$$F_\gamma(y,v) + \bar{h}F_\gamma(z,w) = 0 \;\text{ si }\; \gamma \in B,$$

por lo que $(y,v) + \bar{h}(z,w) \in R_S(x_0,u_0) = R_{S^*}(x_0,u_0)$, de donde

$$F_\gamma(y,v) + \bar{h}F_\gamma(z,w) = F_\gamma(y,v) = 0 \;\text{ si }\; \gamma \in \Delta,$$

por lo que $(y,v) \in R_{K_0}(x_0,u_0)$. La contención contraria es trivial. □

Para los siguientes resultados es conveniente definir el conjunto de índices B^* por

$$B^* := B \cup \{m+1,\ldots,m+n\} = \{q+1,\ldots,m,m+1,\ldots,m+n\}.$$

Lema 5.6.6. *Si $\Delta \neq \emptyset$ entonces los funcionales $\{F_\gamma \mid \gamma \in \Delta \cup B^*\}$ son linealmente dependientes en $\mathbf{Y}_a$.*

Demostración. Si $\Delta \neq \emptyset$, notando que

$$R_K(x_0,u_0) = \{(y,v) \in \mathbf{Y}_a \mid F_\alpha(y,v) \leq 0 \;(\alpha \in \Delta),\; F_\beta(y,v) = 0 \;(\beta \in B^*)\}$$

y

$$R_{K_0}(x_0,u_0) = \{(y,v) \in \mathbf{Y}_a \mid F_\alpha(y,v) = 0 \;(\alpha \in \Delta \cup B^*)\},$$

podemos aplicar el Lema 1.3.4 a los funcionales F_γ $(\gamma \in \Delta \cup B^*)$ para concluir, por el lema anterior, que existen $\lambda_\gamma \in \mathbf{R}$ $(\gamma \in \Delta \cup B^*)$ de forma que $\lambda_\alpha > 0$ si $\alpha \in \Delta$ y

$$\sum_{\gamma\in\Delta\cup B^*} \lambda_\gamma F_\gamma(y,v) = 0 \;\text{ para todo }\; (y,v) \in \mathbf{Y}_a,$$

que es justamente lo que queríamos probar. □

Notemos que también tenemos las herramientas para, basados en la demostración de la Proposición 4.1.8, concluir lo siguiente.

Proposición 5.6.7. *Si (x_0, u_0) es regular de K entonces también es regular de S. Si (x_0, u_0) es regular de K_0 también es regular de K, S^* y S.*

Demostración. Para demostrar la primer conclusión, por el Lema 5.6.4 basta ver que $H \subset T_S(x_0, u_0)$. Si $\Gamma \neq \emptyset$ tomemos $0 \neq (y, v) \in H$, entonces como (x_0, u_0) es regular de K y $(y, v) \in H \subset R_{K_0}(x_0, u_0) = R_K(x_0, u_0)$, tendremos que $(y, v) \in T_K(x_0, u_0)$ por lo que existe $\{(x_k, u_k)\} \subset K$ y $\{t_k\} \subset \mathbf{R}^+$ como en el inciso (1) de la Definición 5.5.1. Entonces

$$\lim_k \frac{I_\gamma(x_k, u_k)}{t_k} = \lim_k \frac{I_\gamma(x_k, u_k) - I_\gamma(x_0, u_0)}{t_k} = F_\gamma(y, v).$$

Como $F_\gamma(y, v) < 0$ para $\gamma \in \Gamma$, podemos escoger k_0 de forma que $I_\gamma(x_k, u_k) < 0$ si $k > k_0$. Por lo tanto $\{(x_k, u_k)\} \subset S$ si $k > k_0$ por lo que $(y, v) \in T_S(x_0, u_0)$.

Si (x_0, u_0) es regular de K_0, y $(y, v) \in H$ entonces $(y, v) \in R_{K_0}(x_0, u_0) = T_{K_0}(x_0, u_0)$ por lo que podemos proceder como antes para verificar que $H \subset T_{S^*}(x_0, u_0)$, y de nuevo el Lema 5.6.4 nos asegurará la regularidad de (x_0, u_0) en S^* que a su vez, por el Lema 5.6.2, nos garantiza la regularidad de (x_0, u_0) en S. Para verificar que la regularidad de K_0 implica la de K notemos que

$$R_K(x_0, u_0) = R_{K_0}(x_0, u_0) = T_{K_0}(x_0, u_0) \subset T_K(x_0, u_0) \subset R_K(x_0, u_0),$$

por lo que $R_K(x_0) = T_K(x_0)$. □

Teorema 5.6.8. *Si los funcionales $\{F_\gamma \mid \gamma \in \Delta \cup B^*\}$ son linealmente independientes en $\mathbf{Y}_a$, entonces $\Delta = \emptyset$ y (x_0, u_0) es regular de S. Si $\Delta = \emptyset$ y los funcionales $\{F_\beta \mid \beta \in B^*\}$ son linealmente independientes en $\mathbf{Y}_a$, entonces (x_0, u_0) es regular de S.*

Demostración. Por el Lema 5.6.7 debemos tener $\Delta = \emptyset$, ya que de lo contrario, los funcionales $\{F_\gamma \mid \gamma \in \Delta \cup B^*\}$ no pueden ser linealmente independientes en $\mathbf{Y}_a$. Además, como K_0 consta únicamente de restricciones de igualdad, normalidad fuerte de (x_0, u_0) en K_0 coincide con normalidad relativa a K_0 que es precisamente la independencia lineal de los funcionales $\{F_\gamma \mid \gamma \in \Delta \cup B^*\}$ en $\mathbf{Y}_a$. Por lo tanto el Teorema 5.5.7 y la

Proposición 5.5.2 nos garantizan que

$$C_{K_0}(x_0,u_0) \subset T_{K_0}(x_0,u_0) \subset R_{K_0}(x_0,u_0).$$

Por otro lado, [18, p. 274, Teorema 7.1] nos dice que, si (x_0,u_0) es normal fuerte de K_0 entonces $C_{K_0}(x_0,u_0) = R_{K_0}(x_0,u_0)$. Combinando este resultado con las contenciones anteriores obtenemos que $T_{K_0}(x_0,u_0) = R_{K_0}(x_0,u_0)$, es decir, (x_0,u_0) es regular de K_0. Finalmente por la Proposición 5.6.7 tendremos que (x_0,u_0) es normal de S. La segunda conclusión es inmediata de la primera. $\square$

Teorema 5.6.9. *(x_0,u_0) es normal relativo a S si y sólo si los funcionales $\{F_\beta \mid \beta \in B^*\}$ son linealmente independientes en $\mathbf{Y}_a$ y existe un vector $0 \neq (z,w) \in \mathbf{Y}$ con*

$$F_\alpha(z,w) < 0 \ (\alpha \in A) \qquad F_\beta(z,w) = 0 \ (\beta \in B).$$

Demostración. Primero supongamos que los funcionales $\{F_\beta \mid \beta \in B^*\}$ son linealmente independientes en $\mathbf{Y}_a$ y existe un vector $0 \neq (z,w) \in \mathbf{Y}$ con las propiedades indicadas. Si $\lambda_1,\ldots,\lambda_{m+n}$ son multiplicadores como en la definición de normalidad relativa a S entonces

$$0 = \sum_{\gamma=1}^{m+n} \lambda_\gamma F_\gamma(z,w) = \sum_{\gamma=1}^{q} \lambda_\gamma F_\gamma(z,w).$$

Como $F_\gamma(z,w) < 0$ y $\lambda_\gamma \geq 0$ si $\gamma \in \Gamma = A$, debemos tener que $\lambda_\gamma = 0$ si $\gamma \in A$. De esto que

$$\sum_{\gamma=q+1}^{m+n} \lambda_\gamma F_\gamma(y,v) = 0 \quad \text{para todo } (y,v) \in \mathbf{Y}_a$$

y, por la independencia lineal de estos funcionales, $\lambda_{q+1} = \cdots = \lambda_{m+n} = 0$. Por lo tanto (x_0,u_0) es normal relativo a S.

Ahora supongamos que (x_0,u_0) es normal de S. En particular haciendo $\lambda_\alpha = 0$ para $\alpha \in A$, la única solución de

$$0 = \sum_{\gamma=1}^{m+n} \lambda_\gamma F_\gamma(y,v) = \sum_{\gamma=q+1}^{n+m} \lambda_\gamma F_\gamma(y,v) \ \text{ para todo } \ (y,v) \in \mathbf{Y}_a$$

es $\lambda_\gamma = 0$ $(\gamma \in B^*)$, por lo que los funcionales $\{F_\gamma \mid \gamma \in B^*\}$ son linealmente independientes en $\mathbf{Y}_a$. Más aún, $\Delta = \emptyset$ ya que de lo contrario, por el Lema 5.6.5, existiría una solución de

$$\sum_{\gamma\in\Delta\cup B^*} \lambda_\gamma F_\gamma(y,v) = 0 \;\text{ para todo }\; (y,v)\in \mathbf{Y}_a,$$

con $\lambda_\alpha > 0$ si $\alpha\in\Delta$, lo que por hipótesis de normalidad relativa a S no es posible. Entonces, como $\Gamma = A$, para cada $\alpha \in A$ existe $(y_\alpha, v_\alpha) \in R_S(x_0,u_0)$ con

$$F_\alpha(y_\alpha, v_\alpha) < 0,$$

por lo que tomando $(z,w) = \sum_{\alpha\in A}(y_\alpha, v_\alpha)$, se cumple con lo que necesitamos. □

Finalmente tenemos lo que queríamos demostrar desde un principio.

Teorema 5.6.10. *Sea $(x_0,u_0)\in S$. Si (x_0,u_0) es normal de S, también es regular de S.*

Demostración. Notemos que la existencia de $0 \neq (z,w) \in \mathbf{Y}$ con las propiedades enunciadas en el teorema anterior, es justamente que $\Delta = \emptyset$. En vista de esto y del teorema anterior, la segunda conclusión del Teorema 5.6.7 puede ser reescrita como «normalidad implica regularidad». □

5.7. Demostración Alternativa del Teorema 5.2.5

Fijemos $(x_0,u_0,\lambda,p)\in\mathcal{E}$ y supongamos de nuevo que todos los índices son activos. En particular $\Gamma_0(\lambda) \subset I_a(x_0,u_0)$. Para cada $\alpha\in\Gamma_0(\lambda)$ definamos

$$S^\alpha := \{(x,u) \mid I_\alpha(x,u) \le 0,\;\; I_\beta(x,u) = 0 \ (\beta\in\Gamma_+(\lambda)\cup B)\}$$

y notemos que las restricciones tangenciales de cada S^α en (x_0,u_0) están dadas por

$$R_{S^\alpha}(x_0,u_0) = \{(y,v)\in\mathbf{Y} \mid I'_\alpha(x_0,u_0;y,v) \le 0,\;\; I'_\beta(x_0,u_0;y,v) = 0 \ (\beta\in\Gamma_+(\lambda)\cup B)\}.$$

Lema 5.7.1. *Si (x_0,u_0) es normal relativo a $S_1(\lambda)$, entonces (x_0,u_0) es normal fuerte relativo a S^α para todo $\alpha\in\Gamma_0(\lambda)$.*

Demostración. Por el Teorema 5.6.9 aplicado a $S_1(\lambda)$, los funcionales $F_\beta(x_0,u_0;\cdot)$ $(\beta \in \Gamma_+(\lambda)\cup B^*)$ son linealmente independientes en $\mathbf{Y}_a$ y existe un elemento $(z,w)\in R_{S_1}(x_0,u_0)$ de forma que

$$F_\alpha(z,w) < 0 \ (\alpha\in\Gamma_0(\lambda)),\quad F_\beta(z,w) = 0 \ (\beta\in\Gamma_+(\lambda)\cup B^*).$$

Por la independencia lineal de los funcionales F_β existen elementos $(y_\sigma, v_\sigma) \in \mathbf{Y}_a$ $(\sigma \in \Gamma_+(\lambda) \cup B^*)$ de forma que la matriz $M = (F_\beta(y_\sigma, v_\sigma))_{\beta\sigma}$ tiene determinante distinto de cero. Por lo tanto, si para cada $\alpha \in \Gamma_0(\lambda)$ definimos la matriz M^α por

$$M^\alpha = \begin{pmatrix} F_\alpha(z,w) & V_\alpha \\ 0 & M \end{pmatrix}$$

donde V_α es el vector fila $[F_\alpha(y_\sigma, v_\sigma)]$ $(\sigma \in \Gamma_+(\lambda) \cup B^*)$ y 0 es el vector columna que consiste solo de ceros de dimensión $|\Gamma_+(\lambda) \cup B^*|$, tendremos que

$$|M^\alpha| = F_\alpha(z,w)|M| \neq 0,$$

por lo que los vectores $(z,w), (y_\sigma, v_\sigma)$ $(\sigma \in \Gamma_+(\lambda) \cup B^*)$ cumplen con la definición de normalidad fuerte para S^α. □

El siguiente lema se puede verificar en la demostración de [18, Teorema 9.1, p. 285]. Damos la prueba a continuación por completitud.

Lema 5.7.2. *Si (x_0, u_0) es solución local de $CP(S)$ y $(y,v) \in C_{S_1}(x_0, u_0)$, entonces $J''(x_0, u_0; y, v) \geq 0$.*

Demostración. Sea $(y,v) \in C_{S_1}(x_0, u_0)$ y $(x(\epsilon), u(\epsilon)) \in S_1(\lambda)$ $(0 \leq \epsilon < \delta)$ la familia uniparamétrica con las propiedades necesarias. Del hecho de que esta familia pertenece a $S_1(\lambda)$ podemos garantizar que

$$\sum_{\gamma \in A \cup B} \lambda_\gamma I_\gamma(x(\epsilon), u(\epsilon)) = 0.$$

Por lo tanto, usando la igualdad (5.3.3), tendremos que $J(x(\epsilon), u(\epsilon)) = I(x(\epsilon), u(\epsilon))$ $(0 \leq \epsilon < \delta)$. Si definimos la función g por

$$g(\epsilon) = J(x(\epsilon), u(\epsilon)) = I(x(\epsilon), u(\epsilon)) \quad (0 \leq \epsilon < \delta),$$

tendremos que existe $\delta^* > 0$ con $g(\epsilon) \geq g(0)$ $(0 \leq \epsilon < \delta^*)$ y, por la Proposición 5.3.5, $g'(0) = J'(x_0, u_0; y, v) = 0$. Por lo tanto debemos tener que $g''(0) \geq 0$. Por otro lado, en la Proposición 5.3.5 se mostró que, si $(x_0, u_0, \lambda, p) \in \mathcal{E}$, entonces se dan las igualdades

$$F_x(\tilde{x}_0(t)) = -H_x(t, x_0(t), u_0(t), \lambda, p(t)) - \dot{p}(t) = 0, \quad F_u(\tilde{x}_0(t)) = -H_u(t, x_0(t), u_0(t), \lambda, p(t)) = 0,$$

con las que al desarrollar $g''(0)$ podemos concluir que

$$0 \leq g''(0) = J''(x_0, u_0; y, v). \qquad \square$$

Finalmente damos una demostración alternativa del resultado principal.

Demostración del Teorema 5.2.5. Primero tomemos $(y, v) \in R_{S_1}(x_0, u_0)$ de modo que $F_\alpha(y, v) < 0$ para todo $\alpha \in \Gamma_0(\lambda)$. Por el Lema 5.7.1 podemos concluir que

$$R_{S_1}(x_0, u_0) = \bigcap_{\alpha \in \Gamma_0(\lambda)} R_{S^\alpha}(x_0, u_0) = \bigcap_{\alpha \in \Gamma_0(\lambda)} C_{S^\alpha}(x_0, u_0),$$

por lo tanto si fijamos $\alpha^* \in \Gamma_0(\lambda)$ tendremos que existe una familia uniparamétrica $(x(\epsilon), u(\epsilon)) \in S^{\alpha^*}$ $(0 \leq \epsilon < \eta)$ con las propiedades necesarias para garantizar que $(y, v) \in C_{S^{\alpha^*}}(x_0, u_0)$. Ahora definamos

$$g_\gamma(\epsilon) := I_\gamma(x(\epsilon), u(\epsilon)) \ \text{ para todo } \gamma \in \Gamma_0(\lambda) - \{\alpha^*\}.$$

Entonces, como $g_\gamma(0) = 0$ y $g'_\gamma(0) = I'_\gamma(x_0, u_0; y, v) < 0$ para todo $\gamma \in \Gamma_0(\lambda) - \{\alpha^*\}$, existe $\delta_\gamma > 0$ de forma que $I_\gamma(x(\epsilon), u(\epsilon)) = g_\gamma(\epsilon) \leq 0$ si $0 \leq \epsilon < \delta_\gamma$. De esto que, tomando $\delta^* = \text{mín}\{\eta, \delta_\gamma \mid \gamma \in \Gamma_0(\lambda) - \{\alpha^*\}\}$, podemos garantizar que $(x(\epsilon), u(\epsilon)) \in S_1(\lambda)$ $(0 \leq \epsilon < \delta^*)$, por lo que $(y, v) \in C_{S_1}(x_0, u_0)$ y a su vez por el Lema 5.7.2 tendremos que $J''(x_0, u_0; y, v) \geq 0$.

Ahora tomemos $(y, v) \in R_{S_1}(x_0, u_0)$ arbitrario. Como (x_0, u_0) es normal relativo a $S_1(\lambda)$, existe $(z, w) \in R_{S_1}(x_0, u_0)$ de forma que $F_\alpha(z, w) < 0$ para todo $\alpha \in \Gamma_0(\lambda)$. Si definimos $(y_k, v_k) := (y, v) + \frac{1}{k}(z, w)$ tendremos que $F_\alpha(y_k, v_k) < 0$ para todo $\alpha \in \Gamma_0(\lambda)$ y (y_k, v_k) tiende a (y, v), por lo que

$$J''(x_0, u_0; y, v) = \lim_{k \to \infty} J''(x_0, u_0; y_k, v_k) \geq 0$$

debido a la continuidad de $J''(x_0, u_0; \cdot)$ y lo mostrado anteriormente. $\square$

5.8. Ejemplo

Por último daremos un ejemplo donde el Teorema 5.2.5 proporciona más información acerca de un extremo que la teoría clásica. Consideremos el problema de minimizar

$$I(x,u) = \int_0^1 \{x_1^2(t) + x_2^2(t) - u_1^2(t) - u_2^2(t)\}dt$$

$$\text{sujeto a las restricciones} \begin{cases} I_1(x) = \int_0^1 \{x_1(t) + x_2(t)\}dt \leq 0, \\ I_2(x) = \int_0^1 \{x_1(t) + x_2(t) + u_1^2(t) + u_2^2(t)\}dt \leq 0, \\ \dot{x}(t)^* = [u_1(t)(x_1(t)+1),\ u_2(t)(x_2(t)+1)], \\ x(0) = 0,\ x(1) = 0. \end{cases}$$

En este caso $n = m = p = 2$,

$$L(t,x,u) = x_1^2 + x_2^2 - u_1^2 - u_2^2 \qquad L_1(t,x,u) = x_1 + x_2 \qquad L_2(t,x,u) = x_1 + x_2 + u_1^2 + u_2^2$$

y la dinámica está dada por

$$f(t,x,u) = \begin{bmatrix} u_1(x_1+1) \\ u_2(x_2+1) \end{bmatrix}.$$

Claramente el arco $(x_0, u_0) \equiv (0,0)$ es admisible y ambos índices son activos. ¿Será solución local del problema? Primero notemos que al definir $p \equiv 0$ y $\lambda = (\lambda_1, \lambda_2) = (0,0)$ tendremos que $(x_0, u_0, \lambda, p) \in \mathcal{E}$: como

$$H(t,x,u,p,\lambda) = p_1u_1(x_1+1) + p_2u_2(x_2+1) - (x_1^2 + x_2^2 - u_1^2 - u_2^2) - \lambda_1(x_1+x_2) - \lambda_2(x_1 + x_2 + u_1^2 + u_2^2)$$

las derivadas parciales de H están dadas por

$$\begin{aligned} H_{x_1} &= p_1u_1 - 2x_1 - \lambda_1 - \lambda_2 & \qquad H_{u_1} &= p_1(x_1+1) + 2u_1 - 2\lambda_2u_1 \\ H_{x_2} &= p_2u_2 - 2x_2 - \lambda_1 - \lambda_2 & \qquad H_{u_2} &= p_2(x_2+1) + 2u_2 - 2\lambda_2u_2, \end{aligned}$$

de donde es fácil verificar que

$$\dot{p}_i(t) = -H_{x_i}(t, x_0(t), u_0(t), p(t), \lambda), \qquad 0 = H_{u_i}(t, x_0(t), u_0(t), p(t), \lambda)$$

para $i = 1, 2$, por lo que efectivamente $(x_0, u_0, \lambda, p) \in \mathcal{E}$. Además en este caso $S_1(\lambda) = S$. Ahora veamos que (x_0, u_0) es normal relativo a S pero no es normal relativo a S_0, por lo que no podemos recurrir al resultado clásico sobre condiciones necesarias de segundo orden. Supongamos que $\mu_1, \mu_2, \mu_3, \mu_4$ son tales que $\mu_1, \mu_2 \geq 0$ y

$$\mu_1 I_1'(x_0, u_0; y, v) + \mu_2 I_2'(x_0, u_0; y, v) + \mu_3 y_1(1) + \mu_4 y_2(1) = 0 \text{ para todo } (y, v) = (y_1, y_2, v_1, v_2) \in \mathbf{Y}_a.$$

Entonces

$$\begin{aligned}
&\mu_1 I_1'(x_0, u_0; y, v) + \mu_2 I_2'(x_0, u_0; y, v) + \mu_3 y_1(1) + \mu_4 y_2(1) \\
&= \mu_1 \int_0^1 \{y_1(t) + y_2(t)\} dt + \mu_2 \int_0^1 \{y_1(t) + y_2(t) + 2u_{01}(t)v_1(t) + 2u_{02}(t)v_2(t)\} dt + \mu_3 y_1(1) + \mu_4 y_2(1) \\
&= (\mu_1 + \mu_2) \int_0^1 \{y_1(t) + y_2(t)\} dt + \mu_3 y_1(1) + \mu_4 y_2(1)
\end{aligned}$$

ya que $u_{01} \equiv 0 \equiv u_{02}$. Por otro lado, la ecuación de variación está dada por

$$\begin{aligned}
\dot{y}(t) &= \begin{bmatrix} u_{01}(t) & 0 \\ 0 & u_{02}(t) \end{bmatrix} \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} + \begin{bmatrix} x_{01}(t) + 1 & 0 \\ 0 & x_{02}(t) + 1 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} \\
&= \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix}.
\end{aligned}$$

Si $y(t)^* = (\operatorname{sen} \pi t, \operatorname{sen} \pi t)$ y $v(t)^* = (\pi \cos \pi t, \pi \cos \pi t)$ entonces $(y, v) \in \mathbf{Y}_a$ y

$$0 = (\mu_1 + \mu_2) \int_0^1 \{y_1(t) + y_2(t)\} dt + \mu_3 y_1(1) + \mu_4 y_2(1) = 2(\mu_1 + \mu_2) \int_0^1 \operatorname{sen} \pi t dt.$$

Como la integral de la derecha es estrictamente positiva y μ_1, μ_2 son no negativos, debemos tener que $\mu_1 = \mu_2 = 0$. Tomando las parejas de funciones $y(t)^* = (t, t), v(t)^* = (1, 1)$ y $y(t)^* = (t, -t), v(t)^* = (1, -1)$ en $\mathbf{Y}_a$ podemos ver que

$$\mu_3 + \mu_4 = 0, \qquad \mu_3 - \mu_4 = 0,$$

de donde $\mu_3 = \mu_4 = 0$, por lo que (x_0, u_0) es normal relativo de S. Para ver que (x_0, u_0) no es normal relativo de S_0, basta tomar $0 \neq \mu \in \mathbf{R}$ y notar que, al quitar la restricción sobre los signos de los dos primeros multiplicadores, $(\mu, -\mu, 0, 0)$ es solución no trivial del sistema que define normalidad de S_0.

Por otro lado las matrices de segundas derivadas parciales de H evaluadas en $(t, x_0(t), u_0(t), p(t), \lambda)$ son

$$H_{xx} = \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} \qquad H_{xu} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad H_{uu} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},$$

por lo que la segunda variación de J está dada por

$$J''(x_0, u_0, y, v) = \int_0^1 \{2y_1^2(t) + 2y_2^2(t) - 2v_1^2(t) - 2v_2^2(t)\}dt.$$

Tomemos la pareja (y, v) definida por

$$y(t)^* = \begin{cases} (-t, -t) & \text{si } t \in [0, 1/2] \\ (t-1, t-1) & \text{si } t \in [1/2, 1] \end{cases} \qquad v(t)^* = \begin{cases} (-1, -1) & \text{si } t \in [0, 1/2] \\ (1, 1) & \text{si } t \in [1/2, 1] \end{cases}$$

Entonces claramente $(y, v) \in \mathbf{Y}$ e

$$\begin{aligned} I_1'(x_0, u_0; y, v) = I_2'(x_0, u_0; y, v) &= \int_0^{1/2} \{-t - t\}dt + \int_0^{1/2} \{(1-t) + (1-t)\}dt \\ &= -2\frac{t^2}{2}\bigg|_0^{1/2} + 2\frac{(1-t)^2}{2}\bigg|_{1/2}^1 = -\frac{1}{2} < 0 \end{aligned}$$

por lo que $(y, v) \in R_S(x_0, u_0)$. Además

$$\begin{aligned} J''(x_0, u_0; y, v) &= \int_0^{1/2} \{4t^2 - 4\}dt + \int_{1/2}^1 \{4(t-1)^2 - 4\}dt \\ &= 4\frac{t^3}{3}\bigg|_0^{1/2} - 2 + 4\frac{(t-1)^3}{3}\bigg|_{1/2}^1 - 2 \\ &= \frac{1}{6} + \frac{1}{6} - 4 < 0 \end{aligned}$$

por lo que el Teorema 5.2.5 nos garantiza que $(x_0, u_0) \equiv (0, 0)$ no puede ser solución local del problema.

Conclusiones

En este trabajo se analizaron problemas isoperimétricos de optimización mediante la teoría de conos tangentes. En particular se observó que una generalización del cono tangente curvilíneo (a quien denotamos por C_S), usualmente visto en cursos de cálculo de varias variables, nos permite demostrar los resultados clásicos en el área, mientras que una generalización del cono tangente secuencial, con una norma apropiada, nos permite obtener resultados más generales sobre condiciones de segundo orden. De este modo se logró debilitar la hipótesis de normalidad fuerte ampliamente usada en la literatura de condiciones necesarias de segundo orden para el problema isoperimétrico de Lagrange con puntos extremos fijos en el cálculo de variaciones. Esta condición de normalidad fuerte aparece en distintos contextos y problemas como alguna de las equivalencias enunciadas en el Capítulo 2. Incluso en fuentes donde se abordan problemas de optimización sin condiciones de normalidad a priori, se necesita imponer esta condición para obtener un resultado análogo al nuestro (ver por ejemplo [13]).

Una de las principales aplicaciones de las condiciones de segundo orden es reducir nuestro conjunto $\mathcal{E}$ de candidatos a solución del problema, definidos como los elementos que satisfacen las condiciones de primer orden. Como se observó en el ejemplo del Capítulo 4, existen problemas para los que nuestro resultado nos da más información acerca del conjunto $\mathcal{E}$ que el resultado clásico.

También se consiguió trasladar el resultado antes mencionado al contexto de control óptimo gobernado por ecuaciones diferenciales ordinarias. En este caso, los funcionales y el espacio que definen la normalidad débil con la que se demostró el resultado principal, se modificaron adecuadamente para garantizar la controlabilidad del sistema. A esta nueva noción de normalidad se le asoció un conjunto $\bar{\mathcal{E}}$ de extremos, con la desventaja de que no corresponden, al menos en un principio, al conjunto de candidatos a solución del problema. Sin embargo en la Sección 5.3 se probó que esta desventaja solo es aparente, ya que este conjunto determina completamente el conjunto de candidatos a soluciones y viceversa.

Existen varias formas en las que se pueden ampliar los resultados principales de este trabajo, por ejemplo obteniendo análogos en el problema con puntos extremos variables o en los problemas de Mayer y Bolza. También se pueden buscar resultados equivalentes en el problema más complejo donde las restricciones no son funcionales integrales, como el resultado parcial que se obtuvo en la Sección 3.4. Para estos problemas aún no se tiene una conclusión equivalente, tanto en cálculo de variaciones como en control óptimo, aunque para algunos casos particulares se tienen resultados análogos a los presentados en este trabajo (ver por ejemplo [9]).

Bibliografía

[1] Aubin JP, Frankowska H (1990) *Set-Valued Analysis*, Birkhäuser.

[2] Amman H (1990) *Ordinary Differential Equations. An Introduction to Nonlinear Analysis*, de Gruyter Studies in Mathematics **13**, Berlin.

[3] Arutyunov AV, Pereira FL (2006) *Second-Order Necessary Optimality Conditions for Problems Without a Priori Normality Assumptions*, Mathematics of Operations Research, **31**: 1-12.

[4] Arutyunov AV, Vereshchagina YS (2002) *On Necessary Second-Order Conditions in Optimal Control Problems*, Differential Equations, **38**: 1531-1540.

[5] Becerril JA, Rosenblueth JF (2017) *Necessity for isoperimetric inequality constraints*, Discrete and Continuous Dynamical Systems, **37**: 1129-1158

[6] Becerril JA, Rosenblueth JF (2017) *The Importance of Being Normal, Regular and Proper in the Calculus of Variations*, Journal of Optimization Theory and Applications, **172**: 759-773.

[7] Ben-Tal A (1980) *Second-Order and Related Extremality Conditions in Nonlinear Programming*, Journal of Optimization Theory and Applications, **30**: 143-165.

[8] Caicedo JF (2005) *Cálculo Avanzado: Introducción*, Universidad Nacional de Colombia.

[9] Cortez del Río KL, Rosenblueth JF (2016) *A Second Order Constraint Qualification for Certain Classes of Optimal Control Problems*, WSEAS Transactions on Systems and Control, **11**: 419-424.

[10] de Pinho MR, Rosenblueth JF (2007) *Mixed Constraints in Optimal Control: An Implicit Function Theorem Approach*, IMA Journal of Mathematical Control and Information, **24**: 197-218.

[11] Fleming W (1977) *Functions of Several Variables*, Second Edition, Springer-Verlag.

[12] Fleming W, Rishel R (1975) *Deterministic and Stochastic Optimal Control*, Springer-Verlag.

[13] Gilbert EG, Bernstein DS (1983) *Second Order Necessary Conditions in Optimal Control: Accessory-Problem Results Without Normality Conditions*, Journal of Optimization Theory and Applications, **41**: 75-106.

[14] Giorgi G, Guerragio A, Thierfelder J (2004) *Mathematics of Optimization: Smooth and Nonsmooth Case*, Elsevier, Amsterdam.

[15] Levitin E, Milyutin A, Osmolovskii NP (1978) *Conditions of High Order for a Local Minimum for Problems With Constraints*, Russian Math Surveys, **33**: 97-178.

[16] Loewen PD, Zheng H (1978) *Generalized Conjugate Arcs in Optimal Control*, Proceedings of the 33rd IEEE Conference on Desicion and Control, Lake Buena Vista, Florida, **4**: 4004-4008.

[17] Loewen PD, Zheng H (1994) *Generalized Conjugate Points for Optimal Control Problems*, Nonlinear Analysis, Theory, Methods & Applications **22**: 771-791.

[18] Hestenes MR (1966) *Calculus of Variations and Optimal Control Theory*, John Wiley, New York.

[19] Hestenes MR (1975) *Optimization Theory: The Finite Dimensional Case*, John Wiley, New York.

[20] McShane EJ (1973) *The Lagrange Multiplier Rule*, The American Mathematical Monthly, **80**: 922-925.

[21] Milyutin AA, Osmolovskii NP (1998) *Calculus of Variations and Optimal Control*, Translations of Mathematical Monographs, 180, American Mathematical Society, Providence, Rhode Island.

[22] Osmolovskii NP (1975) *Second Order Conditions for a Weak Local Minimum in an Optimal Control Problem (Necessity, Sufficiency)*, Soviet Math Dolk, **16**: 1480-1484.

[23] Rockafellar R, Wets R (2009) *Variational Analysis*, Springer-Verlag.

[24] Rosenblueth JF (2004) *A New Notion of Conjugacy for Isoperimetric Problems*, Applied Mathematics and Optimization, **50**: 209-228.

[25] Rudin W (1976) *Principles of Mathematical Analysis*, Third Edition, McGraw-Hill.

[26] Russak IB (1975) *Second Order Necessary Condition for Problems with State Inequality Constraints*, SIAM Journal on Control, **13**: 372-388.

[27] Russak IB (1975) *Second Order Necessary Condition for General Problems with State Inequality Constraints*, Journal of Optimization Theory and Applications, **17**: 43-92.

[28] Stefani G, Zezza PL (1996) *Optimality Conditions for a Constrained Control Problem*, SIAM Journal on Control & Optimization, **34**: 635-659.

[29] Tikhomirov V (1986) *Fundamental Principles in the Theory of Extremal Problems*, Wiley.

[30] Valentine FA (1996) *The Problem of Lagrange with Differential Inequalities as Added Side Conditions*, Contributions to the Calculus of Variations 1933-37, The University of Chicago Press, 1937, **34**: 635-659.

[31] Warga J (1996) *A Second-Order Lagrangian Condition for Restricted Control Problems*, Journal of Optimization Theory & Applications, **24**: 475-483.

[32] Zeidan V, Zezza P (1990) *Normality for the Problem of Bolza with an Inequality State Constraint*, Rocky Mountain Journal of Mathematics, **20**: 1235-1248.