UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE CIENCIAS

DOS NUEVOS TEOREMAS DE SUFICIENCIA EN CONTROL ÓPTIMO PARA MÍNIMOS DÉBILES Y PARA EL PROBLEMA DE LAGRANGE CON PUNTOS FIJOS Y CON RESTRICCIONES CON DESIGUALDADES E IGUALDADES MIXTAS

T E S I S

QUE PARA OBTENER EL TÍTULO DE:

ACTUARIO

P R E S E N T A :

MAURICIO BALBUENA RODRÍGUEZ

DIRECTOR DE TESIS:
DR. GERARDO SÁNCHEZ LICEA
2018

1. Datos del alumno
Balbuena
Rodríguez
Mauricio
56 37 55 71
Universidad Nacional Autónoma de México
Facultad de Ciencias
Actuaría
300098215

2. Datos del tutor
Dr
Gerardo
Sánchez
Licea

3. Datos del sinodal 1
Dr
Javier Fernando
Rosenblueth
Laguette

4. Datos del sinodal 2
Dr
Pablo
Barrera
Sánchez

5. Datos del sinodal 3
Dr
Luis Octavio
Silva
Pereyra

6. Datos del sinodal 4
M en C
Christian Gabriel
Miranda
Ruíz

7. Datos del trabajo escrito
Dos nuevos teoremas de suficiencia en control óptimo para
mínimos débiles y para el problema de Lagrange con puntos
fijos y con restricciones con desigualdades e igualdades mixtas
95 p
2018

Agradecimientos

A mis padres

Agradezco su confianza, el esfuerzo y apoyo incondicional que me han brindado para lograr cada uno de los objetivos que me he planteado en la vida, en particular concluir mis estudios universitarios.

Gracias, a cada una de las personas que me han apoyado y que de alguna forma contribuyeron a que hoy concluya satisfactoriamente esta etapa de mi vida.

Debo agradecer de manera especial y sincera al Dr. Gerardo Sánchez Licea por su apoyo invaluable en la realización y dirección de esta tesis.

Sinceramente,
Mauricio Balbuena Rodríguez

Dos nuevos teoremas de suficiencia en control óptimo para mínimos débiles y para el problema de Lagrange con puntos fijos y con restricciones con desigualdades e igualdades mixtas

Mauricio Balbuena Rodríguez

Contenido

1 Introducción

La teoría del cálculo de variaciones es una rama de las matemáticas que tuvo su mayor auge en los siglos XVIII y XIX y fue hasta la primera mitad del siglo XX cuando tomó mayor relevancia. De hecho, en los últimos años ha habido un interés por introducir un nuevo enfoque del cálculo de variaciones motivado en gran parte por las investigaciones concernientes en las áreas de la economía, las finanzas, las ciencias de los materiales y otras disciplinas. Por otra parte, la teoría del control óptimo empezó a desarrollarse entre los años 50 y 60 por un equipo de matemáticos rusos dirigidos por Pontryagin. Esta teoría constituye una herramienta complementaria para resolver problemas como los de optimización dinámica, incluyendo a la teoría clásica del cálculo de variaciones y el principio de optimalidad asociado a la ecuación de Bellman. El objetivo primordial de esta tesis, como lo indica el título, es proporcionar dos nuevos teoremas de suficiencia en control óptimo para mínimos débiles y para el problema de Lagrange con puntos fijos y con restricciones con desigualdades e igualdades mixtas. Adicionalmente, estudiamos la teoría clásica del cálculo de variaciones de forma detallada ya que ésta fue una herramienta fundamental para el desarrollo de los resultados principales de control óptimo que obtuvimos en este trabajo.

En la teoría clásica del cálculo de variaciones la condición necesaria más importante de primer orden es la condición de Euler. Las soluciones de la ecuación de Euler se llaman extremos. La condición necesaria más importante de segundo orden es la condición de Jacobi. Desafortunadamente, la condición de Jacobi únicamente es aplicable para extremos no singulares suaves. Sin embargo, vale la pena mencionar que la verificabilidad de las condiciones mencionadas anteriormente es una tarea muy esperanzadora. Por otro lado, la no negatividad de la segunda variación sobre el conjunto de variaciones admisibles, también es una condición necesaria de segundo orden que no requiere de hipótesis de no singularidad ni de suavidad. No obstante, esta condición tiene el rasgo desafortunado de que, en general, su verificabilidad tiene el mismo grado de dificultad que aquel que se presenta al tratar de resolver el problema directamente. Debido a estas componentes desafortunadas, en general, las condiciones necesarias clásicas de segundo orden en el cálculo de variaciones dan una información limitada cuando los extremos son suaves y singulares. Con el propósito de llenar esta laguna en la teoría, en [1, 2] la no negatividad de la segunda variación ha sido caracterizada en términos de una noción de puntos conjugados extendidos. Los resultados principales de estas teorías tienen que ver con el hecho de que, sin imponer hipótesis de suavidad o de no singularidad en la trayectoria óptima propuesta, estos conjuntos de puntos conjugados extendidos son vacíos en el intervalo semi-cerrado de tiempo subyacente si y solo si la segunda variación es no negativa en el conjunto de variaciones admisibles. Estas teorías parecen ser exitosas puesto que cuando uno puede detectar la existencia de una variación admisible negativa, precisamente esta variación admisible negativa puede ser utilizada para encontrar un punto conjugado extendido lo cual implica que el extremo bajo consideración no es un óptimo. En otras palabras, cuando uno se enfrenta a un problema en el que el candidato a ser un óptimo es singular, nunca es más difícil checar la no vacuidad de los conjuntos introducidos en [1, 2] que verificar la negación de la condición necesaria clásica de segundo orden en términos de la segunda variación. Más aún, puesto que las condiciones suficientes clásicas de segundo orden del cálculo de variaciones se obtienen al hacerles un ligero reforzamiento a las condiciones necesarias más importantes, imponiendo en particular hipótesis de no singularidad y de suavidad, la teoría de suficiencia clásica del cálculo de variaciones no da información para extremos singulares no suaves.

Por otro lado, para los problemas de control óptimo como el que estudiaremos en esta tesis, debido a que las trayectorias están restringidas a una dinámica que involucra variables tiempo-estado-control y también por el hecho de que en general existen algunas restricciones adicionales en las que nuevamente intervienen variables del tipo tiempo-estado-control, el objetivo de verificar la positividad de la segunda variación sobre el conjunto de variaciones admisibles no nulas no presenta esta misma clase de dificultades como sucede en el caso del problema más simple del cálculo de variaciones. De hecho, el objetivo principal de esta tesis consiste en ilustrar cómo un enfoque de suficiencia basado principalmente en la positividad de la segunda variación, puede ser aplicado de manera exitosa para obtener dos nuevos teoremas de suficiencia que dan respuesta al problema de control óptimo de Lagrange con puntos fijos finales, dinámicas no lineales, y restricciones explícitas tiempo-estado-control con desigualdades e igualdades.

Las componentes principales del Capítulo 2 de esta tesis están dadas en el Teorema 2.1. Este teorema proporciona condiciones suficientes para mínimos estrictos débiles del problema (P) de control óptimo que estaremos estudiando. Las hipótesis fundamentales de este teorema son la condición necesaria de primer orden en forma normal conocida como la condición necesaria de Pontryagin, la condición necesaria de Legendre-Clebsch, la positividad de la segunda variación sobre un conjunto apropiado de procesos, que les llamaremos variaciones admisibles no nulas, y una condición que está relacionada con la función exceso de Weierstrass de una función que es muy parecida a la función Hamiltoniana del problema (P). Cabe señalar que el Teorema 2.1 no requiere de la hipótesis estándar conocida como la condición reforzada de Legendre-Clebsch. Además, el Corolario 2.2 muestra cómo el Teorema 2.1 no sólo es aplicable para procesos singulares, es decir, procesos que no satisfacen la condición estricta de Legendre sino que de alguna manera disminuye la laguna existente entre las condiciones necesarias y las suficientes para problemes de este tipo, ver por ejemplo [3, 4, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21], en cuyas referencias todos los enfoques de suficiencia requieren de esta hipótesis modificada de Legendre.

En el Capítulo 3, estudiamos el problema más simple del cálculo de variaciones tal como se lleva a cabo en la literatura clásica, ver por ejemplo [6, 7, 17]. En particular, demostramos que si x_0 es un mínimo débil del problema que también lo denotamos por (P), entonces x_0 satisface la forma integral de la ecuación de Euler la cual es equivalente a la anulación de la primera variación sobre el conjunto de variaciones admisibles. Además, x_0 debe de satisfacer una de las condiciones necesarias de Weierstrass. Si x_0 es un mínimo fuerte de (P), entonces x_0 satisface una nueva condición de Weierstrass que en general es más restrictiva que la condición de Weierstrass correspondiente para mínimos débiles. Una nueva condición necesaria de segundo orden la cual es una consecuencia de las condiciones de Weierstrass es la condición necesaria de Legendre. Además, si x_0 es un mínimo débil de (P) que no es necesariamente C^1 ni no singular, entonces la segunda variación de la función de costo debe de ser no negativa sobre el conjunto de variaciones admisibles. Para el caso en que un arco x_0 es suave y no singular, demostramos la existencia de una nueva condición necesaria de segundo orden conocida como la condición necesaria de Jacobi. Resulta que cuando x_0 es C^1 y no singular, las condiciones de Legendre y Jacobi son equivalentes a la no negatividad de la segunda variación sobre el conjunto de variaciones admisibles. Más aún, si x_0 es suave demostramos que las condiciones reforzadas de Legendre y Jacobi son equivalentes a la positividad de la segunda variación sobre el conjunto de variaciones admisibles no nulas. Por otro lado, demostramos que al hacerles un ligero reforzamiento a las condiciones necesarias de Euler, Legendre y Jacobi, entonces estas condiciones reforzadas se convierten en suficientes para la obtención de mínimos estrictos débiles de (P). Si a las condiciones anteriores les agregamos la condición reforzada de Weierstrass, entonces estas nuevas condiciones reforzadas son suficientes para obtener mínimos estrictos fuertes de (P). Finalmente, en las Secciones 21, 22, 23 y 24 proporcionamos algunos nuevos teoremas de suficiencia que no requieren de hipótesis de suavidad ni de no singularidad y que en algunas ocasiones disminuyen el hueco existente entre las condiciones necesarias y suficientes para un óptimo en el cálculo de variaciones. Vale la pena señalar que en contraste con la teoría clásica de suficiencia, estos nuevos teoremas, en efecto, sí nos entregan alguna información cuando los extremos son singulares y no suaves.

2 Dos nuevos teoremas de suficiencia para mínimos débiles con restricciones mixtas

1. El problema y teoremas principales

Supongamos que tenemos dados un intervalo $T := [t_0, t_1]$ en $\mathbf{R}$, dos puntos fijos ξ_0 y ξ_1 en $\mathbf{R}^n$, y funciones L, f y $\varphi = (\varphi_1, \ldots, \varphi_q)$ que mapean $T \times \mathbf{R}^n \times \mathbf{R}^m$ a $\mathbf{R}$, $\mathbf{R}^n$ y $\mathbf{R}^q$ respectivamente. Sea

$$\mathcal{A} := \{(t, x, u) \in T \times \mathbf{R}^n \times \mathbf{R}^m \mid \varphi_\alpha(t, x, u) \leq 0 \, (\alpha \in R), \, \varphi_\beta(t, x, u) = 0 \, (\beta \in Q)\}$$

donde $R = \{1, \ldots, r\}$ y $Q = \{r + 1, \ldots, q\}$ $(0 \leq r \leq q)$. Si $r = 0$ entonces $R = \emptyset$ y hacemos caso omiso de las funciones φ_α. Similarmente, Si $r = q$ entonces $Q = \emptyset$ y hacemos caso omiso de las funciones φ_β.

El problema de control óptimo de puntos fijos que consideraremos, denotado por (P), es el de minimizar la funcional

$$I(x, u) := \int_{t_0}^{t_1} L(t, x(t), u(t)) dt$$

sobre todas las parejas (x, u) con $x: T \to \mathbf{R}^n$ absolutamente continua y $u: T \to \mathbf{R}^m$ esencialmente acotada, que satisfacen las restricciones:

(a) $\dot{x}(t) = f(t, x(t), u(t))$ (c.s. en T).

(b) $x(t_0) = \xi_0$, $x(t_1) = \xi_1$.

(c) $(t, x(t), u(t)) \in \mathcal{A}$ $(t \in T)$.

Denotemos por $\mathcal{X}$ al espacio de las funciones absolutamente continuas que mapean T a $\mathbf{R}^n$, por $\mathcal{U}_s := L^\infty(T; \mathbf{R}^s)$ $(s \in \mathbf{N})$. Los elementos de $\mathcal{X} \times \mathcal{U}_m$ los llamaremos *procesos* y un proceso (x, u) es *admisible* si satisface (a)-(c). Un proceso (x, u) *resuelve* (P) si es admisible e $I(x, u) \leq I(y, v)$ para todos los procesos admisibles (y, v). Para mínimos locales débiles, un proceso admisible (x, u) es llamado un *mínimo débil* de (P) si es un mínimo para I relativo a la norma

$$\|(x, u)\| := \text{ínf}\{C > 0 : |(x(t), u(t))| \leq C \, (\text{c.s. en } T)\},$$

es decir, si para alguna $\epsilon > 0$, $I(x, u) \leq I(y, v)$ para todos los procesos admisibles (y, v) que satisfacen $\|(y, v) - (x, u)\| < \epsilon$. Este es un *mínimo estricto* si $I(x, u) = I(y, v)$ solamente en el caso en que $(x, u) = (y, v)$.

Para cualquiera $(x, u) \in \mathcal{X} \times \mathcal{U}_m$ usaremos la notación $\tilde{x}(t)$ para representar $(t, x(t), u(t))$, similarmente $\tilde{x}_0(t)$ representa $(t, x_0(t), u_0(t))$. Se asumirá a lo largo de este capítulo que las funciones L, f y φ son continuas y de clase C^2 con respecto a x y a u en $T \times \mathbf{R}^n \times \mathbf{R}^m$.

Para continuar con el desarrollo de la teoría de este capítulo, es conveniente introducir las siguientes definiciones.

- Para toda $(t, x, u, p, \mu) \in T \times \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^n \times \mathbf{R}^q$ sea

$$H(t, x, u, p, \mu) := \langle p, f(t, x, u) \rangle - L(t, x, u) - \langle \mu, \varphi(t, x, u) \rangle.$$

- Dadas $p \in \mathcal{X}$ y $\mu \in \mathcal{U}_q$ definamos, para toda $(t, x, u) \in T \times \mathbf{R}^n \times \mathbf{R}^m$,

$$F(t, x, u) := -H(t, x, u, p(t), \mu(t)) - \langle \dot{p}(t), x \rangle.$$

Con respecto a F (quien depende de p y μ), sea

$$J(x, u) := \langle p(t_1), \xi_1 \rangle - \langle p(t_0), \xi_0 \rangle + \int_{t_0}^{t_1} F(\tilde{x}(t)) dt.$$

Consideremos la *primera variación* de J con respecto a $(x, u) \in \mathcal{X} \times \mathcal{U}_m$ sobre $(y, v) \in \mathcal{X} \times L^2(T; \mathbf{R}^m)$ dada por

$$J'((x, u); (y, v)) := \int_{t_0}^{t_1} \{F_x(\tilde{x}(t)) y(t) + F_u(\tilde{x}(t)) v(t)\} dt,$$

y la *segunda variación de* J con respecto a $(x,u) \in \mathcal{X} \times \mathcal{U}_m$ sobre $(y,v) \in \mathcal{X} \times L^2(T;\mathbf{R}^m)$ dada por

$$J''((x,u);(y,v)) := \int_{t_0}^{t_1} 2\Omega(t,y(t),v(t))dt$$

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

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

• Dadas $p \in \mathcal{X}$ y $\mu \in \mathcal{U}_q$, decimos que (x_0,u_0,p,μ) (o a veces que (x_0,u_0)) es *no singular* si el determinante

$$|F_{uu}(\tilde{x}_0(t))| = |-H_{uu}(\tilde{x}_0(t),p(t),\mu(t))| \neq 0 \quad (t \in T).$$

• Dadas $p \in \mathcal{X}$ y $\mu \in \mathcal{U}_q$, decimos que (x_0,u_0,p,μ) (o a veces que (x_0,u_0)) satisface la condición de *Legendre-Clebsch* si

$$F_{uu}(\tilde{x}_0(t)) = -H_{uu}(\tilde{x}_0(t),p(t),\mu(t)) \geq 0 \quad \text{(c.s. en } T).$$

• Dadas $p \in \mathcal{X}$ y $\mu \in \mathcal{U}_q$, decimos que (x_0,u_0,p,μ) satisface la condición *reforzada de Legendre-Clebsch* si

$$F_{uu}(\tilde{x}_0(t)) = -H_{uu}(\tilde{x}_0(t),p(t),\mu(t)) > 0 \quad (t \in T).$$

• Denotemos por E a la función *exceso de Weierstrass* con respecto a F,

$$E(t,x,u,v) := F(t,x,v) - F(t,x,u) - F_u(t,x,u)(v-u).$$

• Para toda $u \in L^1(T;\mathbf{R}^m)$ sea

$$D(u) := \int_{t_0}^{t_1} V(u(t))dt \quad \text{donde} \quad V(b) := (1+|b|^2)^{1/2} - 1.$$

• Para toda $(t,x,u) \in T \times \mathbf{R}^n \times \mathbf{R}^m$, sea

$$\mathcal{I}_a(t,x,u) := \{\alpha \in R \mid \varphi_\alpha(t,x,u) = 0\},$$

el conjunto de índices activos de (t,x,u).

• La notación $*$ denota transpuesta.

Vale la pena mencionar que bajo ciertas hipótesis de normalidad (veáse [5, 7]), las condiciones necesarias de primer orden afirman que, si (x_0,u_0) es un mínimo débil del problema (P), entonces existen $p \in \mathcal{X}$ y $\mu \in \mathcal{U}_q$ con $\mu_\alpha(t) \geq 0$ y $\mu_\alpha(t)\varphi_\alpha(\tilde{x}_0(t)) = 0$ $(\alpha \in R, t \in T)$ tales que

$$\dot{p}(t) = -H_x^*(\tilde{x}_0(t),p(t),\mu(t)) \text{ (c.s. en } T), \quad H_u(\tilde{x}_0(t),p(t),\mu(t)) = 0 \ (t \in T).$$

En este caso (x_0,u_0,p,μ) será llamado un *extremo*.

Ahora estamos en condiciones de establecer el resultado principal de la tesis, un resultado de suficiencia para mínimos estrictos débiles del problema (P). Las condiciones impuestas incluyen, con respecto a un extremo dado, la condición necesaria de Legendre-Clebsch, *pero no su forma reforzada*, la positividad de la segunda variación sobre el conjunto de variaciones admisibles no nulas, y una condición relacionada con la función exceso de Weierstrass.

2.1 Teorema: *Sea (x_0,u_0) un proceso admisible. Asumamos que $\mathcal{I}_a(\tilde{x}_0(\cdot))$ es constante a pedazos en T, y supongamos que existen $p \in \mathcal{X}$, $\mu \in \mathcal{U}_q$ con $\mu_\alpha(t) \geq 0$ y $\mu_\alpha(t)\varphi_\alpha(\tilde{x}_0(t)) = 0$ $(\alpha \in R, t \in T)$ tales que*

$\dot{p}(t) = -H_x^*(\tilde{x}_0(t),p(t),\mu(t))$ (c.s. en T),
$H_u(\tilde{x}_0(t),p(t),\mu(t)) = 0$ $(t \in T)$,

y las siguientes condiciones son satisfechas:

(i) $F_{uu}(\tilde{x}_0(t)) \geq 0$ *(c.s. en* T*).*

(ii) $J''((x_0,u_0);(y,v)) > 0$ *para toda* $(y,v) \neq (0,0)$, $(y,v) \in \mathcal{X} \times L^2(T;\mathbf{R}^m)$ *que satisface*

(a) $\dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t)$ *(c.s. en* T*),* $y(t_0) = y(t_1) = 0$.

(b) $\varphi_{\alpha x}(\tilde{x}_0(t))y(t) + \varphi_{\alpha u}(\tilde{x}_0(t))v(t) \leq 0$ *(c.s. en* T*,* $\alpha \in \mathcal{I}_a(\tilde{x}_0(t))$*),* $\varphi_{\beta x}(\tilde{x}_0(t))y(t) + \varphi_{\beta u}(\tilde{x}_0(t))v(t) = 0$ *(c.s. en* T*,* $\beta \in Q$*).*

(iii) *Para algunas* $h, \epsilon > 0$*, si* (x,u) *es admisible con* $\|(x,u) - (x_0,u_0)\| < \epsilon$*, entonces*

$$\int_{t_0}^{t_1} E(t, x(t), u_0(t), u(t))dt \geq hD(u - u_0).$$

Entonces existen $\rho, \delta > 0$ *tales que si* (x,u) *es admisible con* $\|(x,u) - (x_0,u_0)\| < \rho$*,*

$$I(x,u) \geq I(x_0,u_0) + \delta D(u - u_0).$$

En particular, (x_0,u_0) *es un mínimo estricto débil de* (P)*.*

El siguiente teorema es el segundo resultado principal de la tesis. Como veremos abajo, este resultado es un corolario del Teorema 2.1 y se puede ver como una generalización directa del teorema clásico de suficiencia para mínimos débiles en el cálculo de variaciones, ya que a diferencia del Teorema 2.1 éste asume la condición reforzada de Legendre y la continuidad del control óptimo propuesto.

2.2 Corolario: *Sea* (x_0,u_0) *un proceso admisible con* u_0 *continua. Asumamos que* $\mathcal{I}_a(\tilde{x}_0(\cdot))$ *es constante a pedazos en* T*, y supongamos que existen* $p \in \mathcal{X}$*,* $\mu \in \mathcal{U}_q$ *con* $\mu_\alpha(t) \geq 0$ *y* $\mu_\alpha(t)\varphi_\alpha(\tilde{x}_0(t)) = 0$ $(\alpha \in R, t \in T)$ *tales que*

$\dot{p}(t) = -H_x^*(\tilde{x}_0(t), p(t), \mu(t))$ *(c.s. en* T*),*
$H_u(\tilde{x}_0(t), p(t), \mu(t)) = 0$ $(t \in T)$*,*

y las siguientes condiciones son satisfechas:

(i) $F_{uu}(\tilde{x}_0(t)) > 0$ $(t \in T)$*.*

(ii) $J''((x_0,u_0);(y,v)) > 0$ *para todas las* $(y,v) \neq (0,0)$, $(y,v) \in \mathcal{X} \times L^2(T;\mathbf{R}^m)$ *que satisfacen*

(a) $\dot{y}(t) = f_x(\tilde{x}_0(t))y(t) + f_u(\tilde{x}_0(t))v(t)$ *(c.s. en* T*),* $y(t_0) = y(t_1) = 0$.

(b) $\varphi_{\alpha x}(\tilde{x}_0(t))y(t) + \varphi_{\alpha u}(\tilde{x}_0(t))v(t) \leq 0$ *(c.s. en* T*,* $\alpha \in \mathcal{I}_a(\tilde{x}_0(t))$*),* $\varphi_{\beta x}(\tilde{x}_0(t))y(t) + \varphi_{\beta u}(\tilde{x}_0(t))v(t) = 0$ *(c.s. en* T*,* $\beta \in Q$*).*

Entonces existen $\rho, \delta > 0$ *tales que si* (x,u) *es admisible con* $\|(x,u) - (x_0,u_0)\| < \rho$*,*

$$I(x,u) \geq I(x_0,u_0) + \delta D(u - u_0).$$

En particular, (x_0,u_0) *es un mínimo estricto débil de* (P)*.*

Demostración: Definamos un tubo restringido de radio $\epsilon > 0$ con centro en (x_0,u_0) por

$$\mathcal{T}_1((x_0,u_0);\epsilon) := \{(t,y,v) \in T \times \mathbf{R}^n \times \mathbf{R}^m : |y - x_0(t)| < \epsilon,\ |v - u_0(t)| < \epsilon\}.$$

Por el Corolario 2.2(i), el Lema 3.93 y dado que u_0 es continua, existen $h, \epsilon > 0$ tales que

$$\langle c, F_{uu}(t,x,u)c\rangle \geq h|c|^2 \quad (c \in \mathbf{R}^m,\ (t,x,u) \in \mathcal{T}_1((x_0,u_0);\epsilon)).$$

Por lo tanto, para (t,x,u,v) con (t,x,u) y (t,x,v) en $\mathcal{T}_1((x_0,u_0);\epsilon)$,

$$\begin{aligned} E(t,x,u,v) &= \int_0^1 (1-\lambda)\langle v - u, F_{uu}(t,x,u + \lambda[v-u])(v-u)\rangle d\lambda \\ &\geq \tfrac{1}{2}h|v-u|^2 \geq hV(v-u). \end{aligned}$$

La conclusión del corolario se sigue del Teorema 2.1. ▮

2. Lemas auxiliares

Ahora, estableceremos tres resultados auxiliares que nos ayudarán a hacer la demostración del Teorema 2.1.

En los siguientes tres lemas asumiremos que tenemos dadas $u_0 \in L^1(T;\mathbf{R}^m)$ y una sucesión $\{u_q\}$ en $L^1(T;\mathbf{R}^m)$ tales que

$$\lim_{q\to\infty} D(u_q - u_0) = 0 \quad \text{y} \quad d_q := [2D(u_q - u_0)]^{1/2} > 0 \quad (q \in \mathbf{N}).$$

Para toda $q \in \mathbf{N}$ y $t \in T$ definimos

$$v_q(t) := \frac{u_q(t) - u_0(t)}{d_q}.$$

2.3 Lema: *Para alguna $v_0 \in L^2(T;\mathbf{R}^m)$ y alguna subsucesión de $\{u_q\}$ (sin reetiquetar), $\{v_q\}$ converge débilmente a v_0 en $L^1(T;\mathbf{R}^m)$.*

Demostración: Definamos para toda $q \in \mathbf{N}$ y $t \in T$,

$$W_q(t) := [1 + \tfrac{1}{2}V(u_q(t) - u_0(t))]^{1/2},$$

y observemos que $V(c)(2 + V(c)) = |c|^2$ $(c \in \mathbf{R}^m)$. Entonces para toda $q \in \mathbf{N}$,

$$\int_{t_0}^{t_1} \frac{|v_q(t)|^2}{W_q^2(t)} dt = 1. \tag{2.1}$$

De esta manera, existen $v_0 \in L^2(T;\mathbf{R}^m)$ y una subsucesión de $\{u_q\}$, nuevamente denotada por $\{u_q\}$, tal que $\{v_q/W_q\}$ converge débil a v_0 en $L^2(T;\mathbf{R}^m)$. Puesto que $W_q^2(t) \geq W_q(t) \geq 1$ para toda $t \in T$, se tiene que

$$0 \leq \int_{t_0}^{t_1} [W_q(t) - 1]dt \leq \int_{t_0}^{t_1} [W_q^2(t) - 1]dt \leq \int_{t_0}^{t_1} V(u_q(t) - u_0(t))dt = D(u_q - u_0).$$

Obsérvese también que

$$\int_{t_0}^{t_1} [W_q(t) - 1]^2 dt = \int_{t_0}^{t_1} [W_q^2(t) - 1]dt - 2\int_{t_0}^{t_1} [W_q(t) - 1]dt.$$

En consecuencia, dada $u \in L^\infty(T;\mathbf{R}^m)$, se sigue que $\|W_q u - u\|_2 \to 0$ cuando $q \to \infty$, y entonces

$$\lim_{q\to\infty} \int_{t_0}^{t_1} \langle v_q(t), u(t)\rangle dt = \lim_{q\to\infty} \int_{t_0}^{t_1} \left\langle \frac{v_q(t)}{W_q(t)}, W_q(t)u(t) \right\rangle dt = \int_{t_0}^{t_1} \langle v_0(t), u(t)\rangle dt,$$

esto es, $\{v_q\}$ converge débil a v_0 en $L^1(T;\mathbf{R}^m)$. ∎

2.4 Lema: *Sean $A_q \in L^\infty(T;\mathbf{R}^{n\times n})$ y $B_q \in L^\infty(T;\mathbf{R}^{n\times m})$ matrices de funciones para las cuales existen constantes $m_0, m_1 > 0$ tales que $\|A_q\|_\infty \leq m_0$, $\|B_q\|_\infty \leq m_1$ $(q \in \mathbf{N})$, y para toda $q \in \mathbf{N}$ denotemos por y_q a la solución del problema con valor inicial,*

$$\dot{y}(t) = A_q(t)y(t) + B_q(t)v_q(t) \ \ (c.s. \text{ en } T), \quad y(t_0) = 0.$$

Entonces existen $\sigma_0 \in L^2(T;\mathbf{R}^n)$ y una subsucesión de $\{y_q\}$, nuevamente denotada por $\{y_q\}$, tales que $\{\dot{y}_q\}$ converge débilmente en $L^1(T;\mathbf{R}^n)$ a σ_0, y por lo tanto si $y_0(t) := \int_{t_0}^t \sigma_0(s)ds$ $(t \in T)$, entonces $y_q(t) \to y_0(t)$ uniformemente en T.

Demostración: Denotemos por $L^2(T;\mathbf{R}^n)'$ al espacio dual de $L^2(T;\mathbf{R}^n)$ y sea $l \in L^2(T;\mathbf{R}^n)'$. Recordando la definición de W_q dada al principio de la demostración del Lema 2.3, tenemos la existencia de alguna $u_l \in L^2(T;\mathbf{R}^n)$ tal que

$$l\left(\frac{\dot{y}_q}{W_q}\right) = \int_{t_0}^{t_1} \left\langle u_l(t), \frac{\dot{y}_q(t)}{W_q(t)} \right\rangle dt = \int_{t_0}^{t_1} \left\langle A_q^*(t)u_l(t), \frac{y_q(t)}{W_q(t)} \right\rangle dt + \int_{t_0}^{t_1} \left\langle u_l(t), B_q(t)\frac{v_q(t)}{W_q(t)} \right\rangle dt.$$

Para toda $q \in \mathbf{N}$, sean Φ_q, Φ_q^{-1}, las soluciones de los problemas con valor inicial

$$\dot{\Phi}(t) = A_q(t)\Phi(t) \text{ c.s. en } T, \quad \Phi(t_0) = I_{n\times n},$$

$$\dot{\Phi}^{-1}(t) = -\Phi^{-1}(t)A_q(t) \text{ c.s. en } T, \quad \Phi^{-1}(t_0) = I_{n\times n},$$

respectivamente. Puesto que $\|A_q\|_\infty \le m_0$ $(q \in \mathbf{N})$, aplicando el lema de Gronwall, se verifica la existencia de alguna $c_0 >$ tal que

$$\text{máx}\{\|\Phi_q\|_\infty, \|\Phi_q^{-1}\|_\infty\} \le c_0 \quad (q \in \mathbf{N}).$$

Ahora, puesto que $\|B_q\|_\infty \le m_1$ $(q \in \mathbf{N})$ y el hecho de que $\{v_q\}$ converge débil en $L^1(T;\mathbf{R}^m)$ a v_0, existe $c_1 > 0$ tal que para toda $q \in \mathbf{N}$ y $t \in T$,

$$\begin{aligned}|y_q(t)| &\le |\Phi_q(t)| \int_{t_0}^{t_1} |\Phi_q^{-1}(s)B_q(s)v_q(s)|ds \le \|\Phi_q\|_\infty \cdot \|\Phi_q^{-1}\|_\infty \cdot \|B_q\|_\infty \int_{t_0}^{t_1} |v_q(s)|ds \\ &\le c_0^2 \cdot m_1 \int_{t_0}^{t_1} |v_q(s)|ds \le c_0^2 \cdot m_1 \cdot c_1.\end{aligned}$$

De esta manera, existe $c_2 > 0$ tal que $\|y_q\|_\infty \le c_2$ $(q \in \mathbf{N})$. Entonces para toda $q \in \mathbf{N}$,

$$\begin{aligned}\left| \int_{t_0}^{t_1} \left\langle A_q^*(t)u_l(t), \frac{y_q(t)}{W_q(t)} \right\rangle dt \right| &\le \int_{t_0}^{t_1} |A_q^*(t)u_l(t)| \cdot |y_q(t)|dt \\ &\le (t_1 - t_0)^{1/2} \cdot \|A_q^*\|_\infty \cdot \|u_l\|_2 \cdot \|y_q\|_\infty \\ &\le (t_1 - t_0)^{1/2} \cdot m_0 \cdot \|u_l\|_2 \cdot c_2.\end{aligned}$$

Por (2.1), para toda $q \in \mathbf{N}$,

$$\left| \int_{t_0}^{t_1} \left\langle u_l(t), B_q(t)\frac{v_q(t)}{W_q(t)} \right\rangle dt \right| \le \|u_l\|_2 \cdot m_1 \cdot \left\| \frac{v_q}{W_q} \right\|_2 = \|u_l\|_2 \cdot m_1.$$

Por lo tanto, $\{l(\dot{y}_q/W_q)\}_q$ está acotada en $\mathbf{R}$ para toda $l \in L^2(T;\mathbf{R}^n)'$ y en consecuencia, $\{\dot{y}_q/W_q\}$ está acotada en $L^2(T;\mathbf{R}^n)$. Entonces, existe una función $\sigma_0 \in L^2(T;\mathbf{R}^n)$ tal que alguna subsucesión (sin reetiquetar) converge débil en $L^2(T;\mathbf{R}^n)$ a σ_0. Consecuentemente, $\{\dot{y}_q\}$ converge débil en $L^1(T;\mathbf{R}^n)$ a σ_0, y así $\{\dot{y}_q\}$ es equi-integrable en T, esto es, para toda $\epsilon > 0$, existe $\delta_\epsilon > 0$ tal que

$$m(\omega) < \delta_\epsilon \Longrightarrow \left| \int_\omega \dot{y}_q(t)dt \right| < \epsilon \quad (q \in \mathbf{N}).$$

Por lo tanto, la sucesión $\{y_q\}$ es equi-continua en T, y en consecuencia,

$$y_q(t) = \int_{t_0}^{t} \dot{y}_q(s)ds \to y_0(t) := \int_{t_0}^{t} \sigma_0(s)ds$$

uniformemente en T. ▮

2.5 Lema: *Supongamos que $u_q(t) \to u_0(t)$ uniformemente en T. Sean $R_q, R_0 \in L^\infty(T;\mathbf{R}^{m\times m})$, asumamos que $R_q(t) \to R_0(t)$ uniformemente en T, $R_0(t) \ge 0$ (c.s. en T), y sea v_0 la función considerada en el Lema 2.3. Entonces,*

$$\liminf_{q\to\infty} \int_{t_0}^{t_1} \langle R_q(t)v_q(t), v_q(t)\rangle dt \ge \int_{t_0}^{t_1} \langle R_0(t)v_0(t), v_0(t)\rangle dt.$$

Demostración: Nuevamente, recordando la definición de W_q, obsérvese que como $u_q(t) \to u_0(t)$ uniformemente en T, entonces $W_q(t) \to 1$ uniformemente en T. De esta manera, para toda $h \in L^2(T;\mathbf{R}^m)$,

$$\lim_{q\to\infty} \int_{t_0}^{t_1} \langle v_q(t), h(t)\rangle dt = \lim_{q\to\infty} \int_{t_0}^{t_1} \left\langle \frac{v_q(t)}{W_q(t)}, W_q(t)h(t) \right\rangle dt = \int_{t_0}^{t_1} \langle v_0(t), h(t)\rangle dt,$$

esto es, $\{v_q\}$ converge débil en $L^2(T;\mathbf{R}^m)$ a v_0. Puesto que $R_0(t) \geq 0$ (c.s. en T), la función

$$v \mapsto \int_{t_0}^{t_1} \langle v(t), R_0(t)v(t)\rangle dt$$

es convexa en $L^2(T;\mathbf{R}^m)$ y puesto que ésta es fuertemente continua en $L^2(T;\mathbf{R}^m)$, entonces dicha función es débilmente semicontinua inferior en $L^2(T;\mathbf{R}^m)$. De esta manera,

$$\liminf_{q\to\infty} \int_{t_0}^{t_1} \langle v_q(t), R_0(t)v_q(t)\rangle dt \geq \int_{t_0}^{t_1} \langle v_0(t), R_0(t)v_0(t)\rangle dt.$$

Como $R_q(t) \to R_0(t)$ uniformemente en T, se sigue que

$$\liminf_{q\to\infty} \int_{t_0}^{t_1} \langle v_q(t), R_q(t)v_q(t)\rangle dt \geq \int_{t_0}^{t_1} \langle v_0(t), R_0(t)v_0(t)\rangle dt. \blacksquare$$

3. Demostración del Teorema 2.1

La demostración se realizará por contraposición, es decir, asumiremos que para toda $\rho, \delta > 0$, existe un proceso admisible (x,u) tal que

$$\|(x,u) - (x_0,u_0)\| < \rho \quad \text{e} \quad I(x,u) < I(x_0,u_0) + \delta D(u-u_0). \tag{2.2}$$

También asumiremos que $\mathcal{I}_a(\tilde{x}_0(\cdot))$ es constante a pedazos en T, que la condición clásica de primer orden se cumple, las condiciones (i), (iii) del Teorema 2.1 son satisfechas, y obtendremos la negación de la condición (ii) del Teorema 2.1. Primero que todo, notemos que como $\mu_\alpha(t) \geq 0$ ($\alpha \in R, t \in T$), si (x,u) es admisible, entonces $I(x,u) \geq J(x,u)$. También dado que $\mu_\alpha(t)\varphi_\alpha(\tilde{x}_0(t)) = 0$ ($\alpha \in R, t \in T$), entonces $I(x_0,u_0) = J(x_0,u_0)$. Así, (2.2) implica que para toda $\rho, \delta > 0$, existe (x,u) admisible con $\|(x,u) - (x_0,u_0)\| < \rho$ y

$$J(x,u) < J(x_0,u_0) + \delta D(u-u_0). \tag{2.3}$$

Sea $z_0 := (x_0,u_0)$. Notemos que, para todo proceso admisible $z = (x,u)$,

$$J(z) = J(z_0) + J'(z_0; z - z_0) + \mathcal{K}(z) + \mathcal{E}(z) \tag{2.4}$$

donde

$$\mathcal{E}(x,u) := \int_{t_0}^{t_1} E(t, x(t), u_0(t), u(t))dt,$$

$$\mathcal{K}(x,u) := \int_{t_0}^{t_1} \{M(t,x(t)) + \langle u(t) - u_0(t), N(t,x(t))\rangle\}dt,$$

y las funciones M y N están dadas por

$$M(t,y) := F(t,y,u_0(t)) - F(\tilde{x}_0(t)) - F_x(\tilde{x}_0(t))(y - x_0(t)),$$

$$N(t,y) := F_u^*(t,y,u_0(t)) - F_u^*(\tilde{x}_0(t)).$$

Tenemos que,

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

donde

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

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

Ahora, por (2.3), para toda $q \in \mathbf{N}$ existe $z_q := (x_q, u_q)$ admisible tal que

$$\|z_q - z_0\| < \text{mín}\{\epsilon, 1/q\}, \quad J(z_q) - J(z_0) < \frac{1}{q}D(u_q - u_0). \tag{2.5}$$

Consecuentemente, por la segunda desigualdad en (2.5),

$$d_q := [2D(u_q - u_0)]^{1/2} > 0 \quad (q \in \mathbf{N}),$$

y por la primera desigualdad en (2.5),

$$\lim_{q\to\infty} D(u_q - u_0) = 0.$$

Para toda $t \in T$ y $q \in \mathbf{N}$, definamos

$$v_q(t) := \frac{u_q(t) - u_0(t)}{d_q} \quad \text{y} \quad y_q(t) := \frac{x_q(t) - x_0(t)}{d_q}.$$

Por el Lema 2.3, existen $v_0 \in L^2(T; \mathbf{R}^m)$ y alguna subsucesión de $\{z_q\}$ (sin reetiquetar) tales que $\{v_q\}$ converge débilmente en $L^1(T; \mathbf{R}^m)$ a v_0. Para toda $q \in \mathbf{N}$, tenemos que

$$\dot{y}_q(t) = A_q(t)y_q(t) + B_q(t)v_q(t) \text{ (c.s. en } T), \quad y_q(t_0) = 0,$$

donde

$$A_q(t) = \int_0^1 f_x(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)])d\lambda,$$

$$B_q(t) = \int_0^1 f_u(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)])d\lambda.$$

Por la continuidad de f_x y f_u, existen $m_0, m_1 > 0$ tales que $\|A_q\|_\infty \le m_0$, $\|B_q\|_\infty \le m_1$ ($q \in \mathbf{N}$). Por el Lema 2.4, existen $\sigma_0 \in L^2(T; \mathbf{R}^n)$ y una subsucesión de $\{z_q\}$ (sin reetiquetar) tales que, si para toda $t \in T$,

$$y_0(t) := \int_{t_0}^t \sigma_0(s)ds,$$

entonces $y_q(t) \to y_0(t)$ uniformemente en T.

Afirmamos que $J''(z_0; (y_0, v_0)) \le 0$, $(y_0, v_0) \ne (0,0)$, $y_0(t_0) = y_0(t_1) = 0$ y $\dot{y}_0(t) = f_x(\tilde{x}_0(t))y_0(t) + f_u(\tilde{x}_0(t))v_0(t)$ (c.s. en T).

El hecho de que $y_0(t_0) = y_0(t_1) = 0$ se sigue de la definición de y_q, la admisibilidad de z_q y dado que $y_q(t) \to y_0(t)$ uniformemente en T.

Para toda $q \in \mathbf{N}$, tenemos que

$$\frac{\mathcal{K}(z_q)}{d_q^2} = \int_{t_0}^{t_1} \left\{ \frac{M(t, x_q(t))}{d_q^2} + \left\langle v_q(t), \frac{N(t, x_q(t))}{d_q} \right\rangle \right\} dt.$$

Por el Lema 2.4,

$$\lim_{q\to\infty} \frac{M(t, x_q(t))}{d_q^2} = \tfrac{1}{2}\langle y_0(t), F_{xx}(\tilde{x}_0(t))y_0(t)\rangle,$$

$$\lim_{q\to\infty} \frac{N(t, x_q(t))}{d_q} = F_{ux}(\tilde{x}_0(t))y_0(t),$$

ambos uniformemente en T y, dado que $\{v_q\}$ converge débilmente a v_0 en $L^1(T; \mathbf{R}^m)$,

$$\tfrac{1}{2}J''(z_0; (y_0, v_0)) = \lim_{q\to\infty} \frac{\mathcal{K}(z_q)}{d_q^2} + \tfrac{1}{2}\int_{t_0}^{t_1} \langle v_0(t), F_{uu}(\tilde{x}_0(t))v_0(t)\rangle dt. \tag{2.6}$$

Tenemos que,

$$\liminf_{q\to\infty} \frac{\mathcal{E}(z_q)}{d_q^2} \geq \tfrac{1}{2}\int_{t_0}^{t_1} \langle v_0(t), F_{uu}(\tilde{x}_0(t))v_0(t)\rangle dt. \tag{2.7}$$

En efecto, para toda $t \in T$ y $q \in \mathbf{N}$,

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

donde

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

Claramente,

$$\lim_{q\to\infty} R_q(t) = R_0(t) := F_{uu}(\tilde{x}_0(t)) \quad \text{uniformemente en } T.$$

Por el Teorema 2.1(i), $R_0(t) \geq 0$ (c.s. en T), y por lo tanto, (2.7) se sigue por el Lema 2.5. Por otra parte, dado que

$$\dot{p}(t) = -H_x^*(\tilde{x}_0(t), p(t), \mu(t)) \quad \text{(c.s. en } T),$$

$$H_u(\tilde{x}_0(t), p(t), \mu(t)) = 0 \quad (t \in T),$$

se sigue que $J'(z_0; (y, v)) = 0$ para toda $(y, v) \in \mathcal{X} \times L^2(T; \mathbf{R}^m)$. Con esto en mente, (2.4), (2.5), (2.6) y (2.7),

$$\tfrac{1}{2}J''(z_0; (y_0, v_0)) \leq \lim_{q\to\infty} \frac{\mathcal{K}(z_q)}{d_q^2} + \liminf_{q\to\infty} \frac{\mathcal{E}(z_q)}{d_q^2} = \liminf_{q\to\infty} \frac{J(z_q) - J(z_0)}{d_q^2} \leq 0.$$

Adicionalmente, si $(y_0, v_0) = (0, 0)$, entonces

$$\lim_{q\to\infty} \frac{\mathcal{K}(z_q)}{d_q^2} = 0$$

y en consecuencia, por el Teorema 2.1(iii),

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

lo cual contradice la positividad de h.

También, dado que

$$A_q(t) \to A_0(t) := f_x(\tilde{x}_0(t)) \quad \text{y} \quad B_q(t) \to B_0(t) := f_u(\tilde{x}_0(t))$$

ambas uniformemente en T, $y_q(t) \to y_0(t)$ uniformemente en T y $\{v_q\}$ converge débilmente en $L^1(T; \mathbf{R}^m)$ a v_0, se sigue que $\{\dot{y}_q\}$ converge débilmente en $L^1(T; \mathbf{R}^n)$ a $A_0y_0 + B_0v_0$. Por el Lema 2.4, $\{\dot{y}_q\}$ converge débilmente en $L^1(T; \mathbf{R}^n)$ a $\sigma_0 = \dot{y}_0$. Por lo tanto,

$$\dot{y}_0(t) = A_0(t)y_0(t) + B_0(t)v_0(t) \quad \text{(c.s. en } T)$$

lo cual demuestra nuestra afirmación.

Ahora, probaremos que para toda $\alpha \in \mathcal{I}_a(\tilde{x}_0(t))$,

$$\varphi_{\alpha x}(\tilde{x}_0(t))y_0(t) + \varphi_{\alpha u}(\tilde{x}_0(t))v_0(t) \leq 0 \quad \text{(c.s. en } T). \tag{2.8}$$

En efecto, para toda $\alpha \in R$, $q \in \mathbf{N}$, $t \in T$ y $\lambda \in [0,1]$, definamos

$$G_q^\alpha(t;\lambda) := \varphi_\alpha(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)]),$$
$$W_q^\alpha(t) := [-\varphi_\alpha(\tilde{x}_q(t))]^{1/2},$$
$$Z_0^\alpha(t) := -\varphi_{\alpha x}(\tilde{x}_0(t))y_0(t) - \varphi_{\alpha u}(\tilde{x}_0(t))v_0(t),$$

donde, como de costumbre, $\tilde{x}_q(t) := (t, x_q(t), u_q(t))$. Dada $t \in [t_0, t_1)$ un punto de continuidad de $\mathcal{I}_a(\tilde{x}_0(\cdot))$ y $\alpha \in \mathcal{I}_a(\tilde{x}_0(t))$, como $\mathcal{I}_a(\tilde{x}_0(\cdot))$ es constante a pedazos en T, existe un intervalo $[t, \bar{t}] \subset T$ con $t < \bar{t}$ tal que $\varphi_\alpha(\tilde{x}_0(s)) = 0$ para toda $s \in [t, \bar{t}]$. Tenemos que,

$$\begin{aligned}
0 &\le \lim_{q\to\infty} \int_t^{\bar{t}} \frac{(W_q^\alpha(s))^2}{d_q} ds \\
&= \lim_{q\to\infty} \frac{1}{d_q} \int_t^{\bar{t}} \{-\varphi_\alpha(\tilde{x}_q(s)) + \varphi_\alpha(\tilde{x}_0(s))\} ds \\
&= -\lim_{q\to\infty} \frac{1}{d_q} \int_t^{\bar{t}} \{G_q^\alpha(s;1) - G_q^\alpha(s;0)\} ds \\
&= -\lim_{q\to\infty} \frac{1}{d_q} \int_t^{\bar{t}} \int_0^1 \frac{\partial}{\partial\lambda} G_q^\alpha(s;\lambda) d\lambda ds \\
&= -\lim_{q\to\infty} \frac{1}{d_q} \int_t^{\bar{t}} \int_0^1 \varphi_{\alpha x}(t, x_0(s) + \lambda[x_q(s) - x_0(s)], u_0(s) + \lambda[u_q(s) - u_0(s)])(x_q(s) - x_0(s)) d\lambda ds \\
&\quad - \lim_{q\to\infty} \frac{1}{d_q} \int_t^{\bar{t}} \int_0^1 \varphi_{\alpha u}(t, x_0(s) + \lambda[x_q(s) - x_0(s)], u_0(s) + \lambda[u_q(s) - u_0(s)])(u_q(s) - u_0(s)) d\lambda ds \\
&= -\lim_{q\to\infty} \int_t^{\bar{t}} \int_0^1 \varphi_{\alpha x}(t, x_0(s) + \lambda[x_q(s) - x_0(s)], u_0(s) + \lambda[u_q(s) - u_0(s)]) y_q(s) d\lambda ds \\
&\quad - \lim_{q\to\infty} \int_t^{\bar{t}} \int_0^1 \varphi_{\alpha u}(t, x_0(s) + \lambda[x_q(s) - x_0(s)], u_0(s) + \lambda[u_q(s) - u_0(s)]) v_q(s) d\lambda ds \\
&= \int_t^{\bar{t}} \{-\varphi_{\alpha x}(\tilde{x}_0(s))y_0(s) - \varphi_{\alpha u}(\tilde{x}_0(s))v_0(s)\} ds \\
&= \int_t^{\bar{t}} Z_0^\alpha(s) ds.
\end{aligned}$$

Si $Z_0^\alpha(s) < 0$ en un conjunto medible ω tal que $\omega \subset [t, \bar{t}]$ y $m(\omega) > 0$, entonces

$$0 > \int_\omega Z_0^\alpha(s) ds = \lim_{q\to\infty} \int_\omega \frac{(W_q^\alpha(s))^2}{d_q} ds \ge 0$$

lo cual es una contradicción. Por lo tanto, $Z_0^\alpha(s) \ge 0$ c.s. en $[t, \bar{t}]$ con $t \in [t_0, t_1)$ un punto arbitrario de continuidad de $\mathcal{I}_a(\tilde{x}_0(\cdot))$ y por consiguiente $Z_0^\alpha(t) \ge 0$ (c.s. en T) lo que muestra que (2.8) se cumple.

Finalmente, probaremos que para toda $\beta \in Q$,

$$\varphi_{\beta x}(\tilde{x}_0(t))y_0(t) + \varphi_{\beta u}(\tilde{x}_0(t))v_0(t) = 0 \text{ (c.s. en } T). \tag{2.9}$$

En efecto, para toda $\beta \in Q$, $q \in \mathbf{N}$, $t \in T$ y $\lambda \in [0,1]$, definamos

$$H_q^\beta(t;\lambda) := \varphi_\beta(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)]).$$

Para toda $\beta \in Q$, $q \in \mathbf{N}$ y $t \in T$, tenemos que

$$\begin{aligned}
0 = H_q^\beta(t;1) - H_q^\beta(t;0) &= \int_0^1 \frac{\partial}{\partial\lambda} H_q^\beta(t;\lambda) d\lambda \\
&= \int_0^1 \varphi_{\beta x}(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)])(x_q(t) - x_0(t)) d\lambda \\
&\quad + \int_0^1 \varphi_{\beta u}(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)])(u_q(t) - u_0(t)) d\lambda.
\end{aligned}$$

En consecuencia para toda $\beta \in Q$, $q \in \mathbf{N}$ y $t \in T$,

$$\begin{aligned} 0 = \int_0^1 &\{\varphi_{\beta x}(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)])y_q(t) \\ &+ \varphi_{\beta u}(t, x_0(t) + \lambda[x_q(t) - x_0(t)], u_0(t) + \lambda[u_q(t) - u_0(t)])v_q(t)\}d\lambda. \end{aligned} \tag{2.10}$$

Por (2.10), para toda $t \in T$ y $\beta \in Q$,

$$0 = \int_{t_0}^{t} \{\varphi_{\beta x}(\tilde{x}_0(s))y_0(s) + \varphi_{\beta u}(\tilde{x}_0(s))v_0(s)\}ds$$

y entonces (2.9) se cumple. ▮

4. Ejemplo

Ahora, exhibamos un ejemplo simple de un problema de control óptimo con puntos fijos finales y con restricciones con desigualdades para el cual una aplicación del Corolario 2.2 muestra que el extremo considerado de hecho es un mínimo estricto débil del problema (P) bajo consideración.

2.6 Ejemplo: *Sea $\gamma > 0$ y sea (P_γ) el problema de minimizar*

$$I_\gamma(x, u) = \int_0^\pi \{(\gamma/2)u_2^2(t) + \tfrac{1}{2}u_2(t) - \tfrac{1}{2}x^2(t)\}dt$$

sujeto a $\dot{x}(t) = u_1(t)$ $(t \in [0, \pi])$, $x(0) = x(\pi) = 0$,

$$u_1^2(t) - u_2(t) - u_1(t) \leq 0 \quad y \quad u_1(t) - u_2(t) \leq 0 \quad (t \in [0, \pi]).$$

En este caso $T = [0, \pi]$, $n = 1$, $m = 2$, $q = 2$, $r = 2$, $\xi_0 = \xi_1 = 0$,

$$L(t, x, u) = (\gamma/2)u_2^2 + \tfrac{1}{2}u_2 - \tfrac{1}{2}x^2, \quad f(t, x, u) = u_1,$$

$$\varphi_1(u) = u_1^2 - u_2 - u_1, \quad \varphi_2(u) = u_1 - u_2.$$

En consecuencia

$$H(t, x, u, p, \mu) = pu_1 - (\gamma/2)u_2^2 - \tfrac{1}{2}u_2 + \tfrac{1}{2}x^2 - \mu_1[u_1^2 - u_2 - u_1] - \mu_2[u_1 - u_2],$$

y por lo tanto

$$H_x(t, x, u, p, \mu) = x,$$

$$H_u(t, x, u, p, \mu) = (p - 2\mu_1 u_1 + \mu_1 - \mu_2, -\gamma u_2 - \tfrac{1}{2} + \mu_1 + \mu_2).$$

Sea $(x_0, u_0) \equiv (0, 0, 0)$ la cual es admisible. Notemos que $\mathcal{I}_a(u_0(t)) = \{1, 2\}$ es constante en T. Si definimos $(p, \mu) \equiv (-\frac{1}{2}, \frac{1}{2}, 0)$, entonces, $\mu_1(t), \mu_2(t) \geq 0$, $\mu_1(t)\varphi_1(u_0(t)) = 0$ y $\mu_2(t)\varphi_2(u_0(t)) = 0$ $(t \in T)$. También, (x_0, u_0, p, μ) satisface la condición de primer orden del Teorema 2.2. Tenemos que, para toda $(t, x, u) \in T \times \mathbf{R} \times \mathbf{R}^2$

$$F(t, x, u) = \tfrac{1}{2}u_1^2 + (\gamma/2)u_2^2 - \tfrac{1}{2}x^2.$$

Por lo tanto,

$$F_{uu}(\tilde{x}_0(t)) = \begin{pmatrix} 1 & 0 \\ 0 & \gamma \end{pmatrix} \quad (t \in T)$$

y por lo tanto el Teorema 2.2(i) se verifica. También, $F_{xx}(\tilde{x}_0(t)) = -1$ y $F_{xu}(\tilde{x}_0(t)) = (0, 0)$ $(t \in T)$, y en consecuencia

$$J''((x_0, u_0); (y, v)) = \int_0^\pi \{v_1^2(t) + \gamma v_2^2(t) - y^2(t)\}dt > 0$$

para toda $(y, v) \neq (0, 0)$, $(y, v) \in \mathcal{X} \times L^2(T; \mathbf{R}^2)$ que satisface $\dot{y}(t) = v_1(t)$ (c.s. en T), $y(0) = y(\pi) = 0$, $-v_1(t) - v_2(t) \leq 0$, $v_1(t) - v_2(t) \leq 0$ (c.s. en T). Consecuentemente, el Teorema 2.2(ii) es satisfecho. Por el Teorema 2.2, (x_0, u_0) es un mínimo estricto débil de (P_γ).

5. Aplicación

En la siguiente aplicación apelamos al Teorema 2.1 para resolver el problema (P) correspondiente. Cabe señalar que el *extremo singular* bajo consideración se convierte en un mínimo débil estricto del problema (P). También, es importante mencionar que en esta aplicación, las restricciones mixtas tiempo-estado-control son cruciales para obtener la conclusión mencionada arriba.

La aplicación que estudiaremos está relacionada con un modelo de un sector de la economía que toma en cuenta el crecimiento poblacional. En este modelo económico asumiremos que el único factor que puede hacer que disminuya el capital por trabajador es la incorporación de nuevos trabajadores a la economía. Denotemos por $x(t)$ el capital por trabajador al tiempo t, y por $P(t)$ la tasa de producción por trabajador al tiempo t. Por lo tanto $P(t)x(t)$ es la tasa de incremento del capital por trabajador a causa de la producción. Si la población tiene una tasa de crecimiento $G(t)$ en el tiempo t, hay una tasa de decrecimiento $-G(t)x(t)$ del capital por trabajador a causa del crecimiento poblacional. Una fracción $u(t)$ de la producción generada se mantiene en la economía y la fracción restante $1 - u(t)$ es consumida. De esta manera, obtenemos la ecuación $\dot{x}(t) = u(t)P(t)x(t) - G(t)x(t)$ que proporciona la tasa de cambio del capital por trabajador dada por la diferencia entre la tasa de incremento del capital no consumido menos la tasa de incremento de capital a causa del crecimiento de la población.

Consideremos el problema de elegir un plan de ahorro $u(t)$ para incrementar o disminuir el capital por trabajador de un punto inicial fijo ξ_0 a un punto terminal fijo ξ_1, en un intervalo de tiempo $T := [t_0, t_1]$, mientras que se maximiza la tasa global de consumo

$$\int_{t_0}^{t_1} [1 - u(t)]P(t)x(t)dt.$$

Consideraremos un modelo de crecimiento poblacional suponiendo que el intervalo de tiempo es $T = [0, 1]$, $\xi_0 = (1/3)\exp(2/3)$ y $\xi_1 = 1/3$. La función que nos dará la tasa de crecimiento G está dada por $G(t) := t^{1/2}$ y la función de la producción está dada por $P(t) := t^{1/2}/2$. Además, introduciremos dos restricciones con igualdades tiempo-estado-control $\varphi_\beta(t, x(t), u(t)) = 0$ ($\beta = 1, 2, t \in T$). En este caso, las funciones $\varphi_\beta: T \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ ($\beta = 1, 2$) están dadas por

$$\varphi_1(t, x, u) := xu\exp(-u) - (1/2)x^2u \quad \text{y} \quad \varphi_2(t, x, u) := t^2x^2u^4 - 2t^2xu^4 + t^2u^4 - 4txu^2.$$

Demostraremos que al escoger un plan de ahorro $u(t)$ en el cual todo el capital producido $P(t)x(t)$ es consumido al tiempo t, la tasa global de consumo $\int_0^1 [1 - u(t)]P(t)x(t)dt$ se maximiza localmente.

Como uno puede verificar sin ninguna dificultad, en este caso, el problema de control óptimo (P) consiste en minimizar

$$I(x, u) := \int_0^1 (t^{1/2}/2)x(t)[u(t) - 1]dt$$

sobre todas las parejas (x, u) con $x: T \to \mathbf{R}$ absolutamente continua y $u: T \to \mathbf{R}$ esencialmente acotada sujeta a las restricciones:

(a) $\dot{x}(t) = (t^{1/2}/2)x(t)u(t) - t^{1/2}x(t)$ (c.s. en $[0, 1]$).

(b) $x(0) = (1/3)\exp(2/3)$ y $x(1) = 1/3$.

(c) $(t, x(t), u(t)) \in \mathcal{A}$ $(t \in T)$.

Aquí,

$$\mathcal{A} := \{(t, x, u) \in T \times \mathbf{R} \times \mathbf{R} \mid \varphi_\beta(t, x, u) = 0 \ (\beta \in Q)\}$$

con $Q = \{1, 2\}$.

Para este caso, $n = m = 1$, $q = 2$, $r = 0$, $T = [0, 1]$, $\xi_0 = (1/3)\exp(2/3)$, $\xi_1 = 1/3$,

$$L(t, x, u) = (t^{1/2}/2)x[u - 1], \quad f(t, x, u) = (t^{1/2}/2)xu - t^{1/2}x,$$

$$\varphi_1(t,x,u) = xu\exp(-u) - (1/2)x^2u,$$
$$\varphi_2(t,x,u) = t^2x^2u^4 - 2t^2xu^4 + t^2u^4 - 4txu^2.$$

Sea $(x_0,u_0) \equiv ((1/3)\exp\{(2/3)[1-t^{3/2}]\},0)$ la cual es admisible. Tenemos que

$$\begin{aligned}
H(t,x,u,p,\mu) &= p(t^{1/2}/2)xu - pt^{1/2}x + (t^{1/2}/2)x[1-u] \\
&\quad - \mu_1[xu\exp(-u) - (1/2)x^2u] \\
&\quad - \mu_2[t^2x^2u^4 - 2t^2xu^4 + t^2u^4 - 4txu^2], \\
H_x(t,x,u,p,\mu) &= p(t^{1/2}/2)u - pt^{1/2} + (t^{1/2}/2)[1-u] \\
&\quad - \mu_1[u\exp(-u) - xu] - \mu_2[2t^2xu^4 - 2t^2u^4 - 4tu^2], \\
H_u(t,x,u,p,\mu) &= p(t^{1/2}/2)x - (t^{1/2}/2)x \\
&\quad - \mu_1[x\exp(-u) - xu\exp(-u) - (1/2)x^2] \\
&\quad - \mu_2[4t^2x^2u^3 - 8t^2xu^3 + 4t^2u^3 - 8txu].
\end{aligned}$$

Sea $p \equiv 1/2$,

$$\mu_1(t) := -\frac{t^{1/2}}{4[1-(1/6)\exp\{(2/3)[1-t^{3/2}]\}]} \qquad \text{y} \quad \mu_2(t) := 0 \quad (t \in T).$$

Se puede verificar fácilmente que (x_0,u_0,p,μ) satisface la condición de primer orden del Teorema 2.1. Además, la función F está dada por

$$F(t,x,u) = \frac{t^{1/2}xu}{4} - \frac{t^{1/2}[xu\exp(-u) - (1/2)x^2u]}{4[1-(1/6)\exp\{(2/3)[1-t^{3/2}]\}]}$$

y por lo tanto,

$$F_{uu}(\tilde{x}_0(t)) = \frac{t^{1/2}\exp\{(2/3)[1-t^{3/2}]\}}{6[1-(1/6)\exp\{(2/3)[1-t^{3/2}]\}]} \quad (t \in T)$$

lo que implica que (x_0,u_0) es *singular* y se cumple la condición (i) del Teorema 2.1.

También, observemos que, para toda $t \in T$,

$$f_x(\tilde{x}_0(t)) = -t^{1/2} \quad \text{y} \quad f_u(\tilde{x}_0(t)) = \frac{t^{1/2}\exp\{(2/3)[1-t^{3/2}]\}}{6}$$

y por lo tanto (y,v) satisface (a) y (b) del Teorema 2.1(ii) si y solo si

$$\dot{y}(t) = -t^{1/2}y(t) + \frac{t^{1/2}\exp\{(2/3)[1-t^{3/2}]\}}{6}v(t) \text{ (c.s. en } T),$$

$y(0) = y(1) = 0$, y

$$(1/3)\exp\{(2/3)[1-t^{3/2}]\}[1-(1/6)\exp\{(2/3)[1-t^{3/2}]\}]v(t) = 0 \text{ (c.s. en } T).$$

En consecuencia, no existe ninguna $(y,v) \in \mathcal{X} \times L^2(T;\mathbf{R})$ no nula que satisfaga (a) y (b) del Teorema 2.1(ii) y por lo tanto el Teorema 2.1(ii) se cumple.

Ahora supongamos que la condición (iii) del Teorema 2.1 no se cumple. Entonces para toda $q \in \mathbf{N}$, existe un proceso admisible (x_q,u_q) tal que

$$\|(x_q,u_q) - (x_0,u_0)\| < \frac{1}{q},$$

$$\int_0^1 E(t,x_q(t),u_0(t),u_q(t))dt < \frac{1}{q}D(u_q-u_0).$$

Notemos que la última desigualdad implica que $u_q \neq 0$ ($q \in \mathbf{N}$). Por (c), se puede verificar fácilmente que si $\|(x_q, u_q) - (x_0, u_0)\| < 1/q$ (q suficientemente grande), entonces

$$tx_q(t)u_q^2(t) = [V(t^{1/2}u_q(t))]^2 \quad (t \in T).$$

De esta manera, podemos asumir que para toda $q \in \mathbf{N}$ y $t \in T$,

$$\frac{tx_q(t)u_q^2(t)}{2 + V(t^{1/2}u_q(t))} = \frac{[V(t^{1/2}u_q(t))]^2}{2 + V(t^{1/2}u_q(t))}. \tag{2.11}$$

Para toda $q \in \mathbf{N}$, sea

$$\Gamma_q := \{t \in (0,1] \mid u_q(t) \neq 0\},$$

y observemos que para toda $q \in \mathbf{N}$, Γ_q tiene medida positiva. Ahora, dado que $V(c)(2+V(c)) = c^2$ ($c \in \mathbf{R}$), por (2.11), para toda $q \in \mathbf{N}$,

$$x_q(t) = \frac{V(t^{1/2}u_q(t))}{2 + V(t^{1/2}u_q(t))} \quad (t \in \Gamma_q). \tag{2.12}$$

Dado que

$$2\exp(-2) < \frac{1}{3} = x_0(1) \leq x_0(t) = (1/3)\exp\{(2/3)[1 - t^{3/2}]\} \quad (t \in T),$$

vemos que con $\delta > 0$ suficientemente pequeña,

$$\mathcal{T}_0(x_0;\delta) := \{(t,x) \in T \times \mathbf{R} : |x - x_0(t)| < \delta\}$$

está contenido en

$$\mathcal{Q} := \{(t,x) \in \mathbf{R}^2 \mid t \geq 0,\ x \geq 2\exp(-2)\}.$$

Como para toda q suficientemente grande y $t \in T$, $(t, x_q(t)) \in \mathcal{T}_0(x_0;\delta) \subset \mathcal{Q}$, entonces tenemos que para toda q suficientemente grande y $t \in T$,

$$x_q(t) \geq 2\exp(-2). \tag{2.13}$$

Sin embargo, ya que $u_q(t) \to u_0(t)$ uniformemente en T, entonces para toda q suficientemente grande

$$\frac{V(t^{1/2}u_q(t))}{2 + V(t^{1/2}u_q(t))} < 2\exp(-2) \quad (t \in T). \tag{2.14}$$

Por (2.12) y (2.14), tenemos que para toda q suficientemente grande,

$$x_q(t) < 2\exp(-2) \quad (t \in \Gamma_q)$$

lo cual contradice (2.13). En consecuencia, existen $h, \epsilon > 0$ tales que la condición (iii) del Teorema 2.1 es satisfecha. Por lo tanto, (x_0, u_0) es un mínimo estricto débil de (P).

3 Cálculo de variaciones

1. Funciones continuas a pedazos y funciones C^1 a pedazos

En esta sección definiremos los conceptos fundamentales de las funciones continuas a pedazos, las funciones C^1 a pedazos, y el de las funciones absolutamente continuas. También, enunciaremos algunas de sus propiedades más importantes.

3.1 Definición: *Sea $T := [t_0, t_1]$ un intervalo compacto en $\mathbf{R}$. Una función $\psi: T \to \mathbf{R}^n$ tiene una discontinuidad de primera clase en un punto $t \in (t_0, t_1)$, si*

$$\psi(t-0) := \lim_{s \to t-} \psi(s) \quad y \quad \psi(t+0) := \lim_{s \to t+} \psi(s)$$

existen y

$$\psi(t-0) \neq \psi(t+0).$$

Una función ψ es continua a pedazos en T si ψ tiene a lo más un número finito de discontinuidades de primera clase en T y estas discontinuidades se encuentran en (t_0, t_1). Nótese que por definición

$$\psi(t_0 + 0) = \psi(t_0) \quad y \quad \psi(t_1 - 0) = \psi(t_1),$$

es decir, una función continua a pedazos en T es continua en los puntos inicial y final del intervalo T.

3.2 Observación: *Sea $\psi: T \to \mathbf{R}^n$ continua a pedazos en T. Claramente ψ es acotada y puesto que ψ tiene a lo más un número finito de discontinuidades en T, entonces ψ es Riemann integrable en cualquier intervalo de la forma $[t_0, t]$ con $t \in T$.*

3.3 Definición: *Una función $\varphi: T \to \mathbf{R}^n$ es C^1 a pedazos en T si existe una función $\psi: T \to \mathbf{R}^n$ continua a pedazos en T y una constante $c \in \mathbf{R}^n$ tales que para toda $t \in T$,*

$$\varphi(t) = \int_{t_0}^{t} \psi(\tau) d\tau + c.$$

3.4 Definición: *Una función $x: T \to \mathbf{R}^n$ es absolutamente continua en T, si dada $\epsilon > 0$ existe $\delta_\epsilon > 0$ tal que*

$$\sum_{i=1}^{m} |x(t_i) - x(\bar{t}_i)| < \epsilon$$

para cualquier colección finita $\{(t_i, \bar{t}_i)\}$ de intervalos disjuntos que satisfacen

$$\sum_{i=1}^{m} |t_i - \bar{t}_i| < \delta_\epsilon.$$

3.5 Observación: *Obsérvese que una función $x: T \to \mathbf{R}^n$ absolutamente continua en T es uniformemente continua en T.*

La caracterización más importante de las funciones absolutamente continuas está dada por el siguiente resultado.

3.6 Teorema: *Una función $x: T \to \mathbf{R}^n$ es una integral indefinida si y solo si ésta es absolutamente continua en T.*

3.7 Teorema: *Si una función $x: T \to \mathbf{R}^n$ es absolutamente continua en T, entonces x tiene derivada excepto en un conjunto de medida cero en T y la derivada denotada por $\dot{x}$ es un elemento de $L^1(T; \mathbf{R}^n)$.*

3.8 Teorema: *Si $g \in L^1(T; \mathbf{R}^n)$ y para alguna $c \in \mathbf{R}^n$ definimos $x: T \to \mathbf{R}^n$ por $x(t) := \int_{t_0}^{t} g(s) ds + c$ $(t \in T)$, entonces $\dot{x}(t) = g(t)$ excepto en un conjunto de medida cero de T.*

3.9 Teorema: *Una función* $x: T \to \mathbf{R}^n$ *es absolutamente continua en* T *si y solo si* $x(t) = \int_{t_0}^{t} \dot{x}(s)ds + x(t_0)$ $(t \in T)$*. En otras palabras, una función* x *es absolutamente continua si y solo si* x *es la integral indefinida de su derivada.*

3.10 Observación: *Obsérvese que la clase de funciones absolutamente continuas es la clase más grande de funciones para las cuales son válidos el primer y segundo teoremas fundamentales del cálculo al utilizar integrales de Lebesgue.*

3.11 Observación: *Obsérvese que si* $\varphi: T \to \mathbf{R}^n$ *es una función* C^1 *a pedazos en* T*, puesto que* φ *es una integral indefinida, entonces* φ *es absolutamente continua en* T *y en consencuencia* φ *es continua en* T*. Nótese que si* $\varphi(t) = \int_{t_0}^{t} \psi(\tau)d\tau + c$ $(t \in T, c \in \mathbf{R}^n)$ *y* $t = \bar{t}$ *es un punto de discontinuidad de* ψ*, entonces la derivada de* φ *en* $t = \bar{t}$ *no necesariamente existe. De hecho, la Observación 3.13 asegura que la derivada de* ψ *en* $\bar{t}$ *no existe. Si* ψ *tiene una discontinuidad en el punto* $t = \bar{t}$*, diremos que la función* φ *tiene una esquina o un pico en* $t = \bar{t}$*.*

3.12 Proposición: *Si* $\varphi: T \to \mathbf{R}^n$ *es una función* C^1 *a pedazos en* T*, entonces para toda* $t \in (t_0, t_1)$*,* $\varphi'(t-0)$ *y* $\varphi'(t+0)$ *existen, donde*

$$\varphi'(t-0) := \lim_{h \to 0-} \frac{\varphi(t+h) - \varphi(t)}{h} \quad y \quad \varphi'(t+0) := \lim_{h \to 0+} \frac{\varphi(t+h) - \varphi(t)}{h}.$$

Además, si $\varphi(t) = \int_{t_0}^{t} \psi(\tau)d\tau + c$ $(t \in T, c \in \mathbf{R}^n)$*, entonces* $\varphi'(t-0) = \psi(t-0)$ *y* $\varphi'(t+0) = \psi(t+0)$ *para toda* $t \in (t_0, t_1)$*. Adicionalmente, nótese que* $\varphi'(t_0) = \psi(t_0)$ *y* $\varphi'(t_1) = \psi(t_1)$*.*

Demostración: Sea $\psi: T \to \mathbf{R}^n$ una función continua a pedazos en T tal que para toda $t \in T$ y para alguna constante $c \in \mathbf{R}^n$

$$\varphi(t) = \int_{t_0}^{t} \psi(\tau)d\tau + c.$$

Sea $\bar{t}$ fija en el intervalo (t_0, t_1). Puesto que ψ es continua a pedazos en T, para toda $h < 0$ suficientemente pequeña se tiene que ψ es continua en el intervalo $[\bar{t}+h, \bar{t})$. De esta manera, para toda $h < 0$ suficientemente pequeña, existe $\tau_h \in (\bar{t}+h, \bar{t})$ tal que

$$\int_{\bar{t}+h}^{\bar{t}} \psi(\tau)d\tau = -\psi(\tau_h)h.$$

Con esto en mente, se tiene que

$$\lim_{h \to 0-} \frac{\varphi(\bar{t}+h) - \varphi(\bar{t})}{h} = -\lim_{h \to 0-} \frac{1}{h} \int_{\bar{t}+h}^{\bar{t}} \psi(\tau)d\tau = \lim_{h \to 0-} \psi(\tau_h) = \psi(\bar{t}-0),$$

y por lo tanto $\varphi'(\bar{t}-0) = \psi(\bar{t}-0)$. Similarmente, se demuestra que $\varphi'(\bar{t}+0) = \psi(\bar{t}+0)$. ∎

3.13 Observación: *Si una función* $\phi: T \to \mathbf{R}^n$ *es diferenciable en todos los puntos de* T *y* $\phi': T \to \mathbf{R}^n$ *la derivada de* ϕ *es discontinua en un punto* $\bar{t} \in (t_0, t_1)$*, entonces la discontinuidad de* ϕ' *en* $t = \bar{t}$ *es de segunda clase, es decir,*

$$\lim_{s \to \bar{t}-} \phi'(s) \text{ no existe,}$$

o

$$\lim_{s \to \bar{t}+} \phi'(s) \text{ no existe.}$$

De esta manera, si $\psi: T \to \mathbf{R}^n$ *es una función continua a pedazos en* T *con al menos un punto de discontinuidad en* (t_0, t_1) *y* $\phi: T \to \mathbf{R}^n$ *es diferenciable en* T*, entonces la relación*

$$\phi' = \psi$$

es imposible.

2. El problema clásico de cálculo de variaciones con puntos fijos finales

Supongamos que tenemos dados un intervalo compacto $T := [t_0, t_1]$ en $\mathbf{R}$, dos puntos fijos ξ_0 y ξ_1 en $\mathbf{R}^n$, y una *función continua* $L(t, x, \dot{x})\colon T \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$. De hecho, a lo largo de este trabajo asumiremos que L es de clase C^2 en $T \times \mathbf{R}^n \times \mathbf{R}^n$, sin embargo, el lector que esté interesado en los detalles de la suavidad de la función L detectará que en muchas ocasiones no será necesario que la función L sea C^2 sino únicamente C^1 en $T \times \mathbf{R}^n \times \mathbf{R}^n$ o algo muy similar a estas propiedades de suavidad de la función L. Aquí $L(t, x, \dot{x})$ es una función de 1+2n variables $(t, x^1, \ldots, x^n, \dot{x}^1, \ldots, \dot{x}^n)$. Se debe de observar que $\dot{x} = (\dot{x}^1, \ldots, \dot{x}^n)^* \in \mathbf{R}^n$ es considerada como una variable libre y no como una variable que se obtiene al sacar una derivada. El convenio de utilizar $\dot{x}$ en ambos sentidos, es decir, tanto como una derivada o como una variable libre, es estándar en el cálculo de variaciones. El contexto determina la interpretación que debe de usarse. De hecho, la notación $\dot{x}(t)$ denota la derivada de la función x evaluada en el punto t. Por otro lado, si $T = [0, 1]$, $n = 1$ y $L\colon T \times \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ está definida por

$$L(t, x, \dot{x}) := t\dot{x}^2 - t^2x^2,$$

entonces, para toda $(t, x, \dot{x}) \in T \times \mathbf{R} \times \mathbf{R}$, se tiene que

$$L_t(t, x, \dot{x}) = \dot{x}^2 - 2tx^2, \quad L_x(t, x, \dot{x}) = -2t^2x, \quad L_{\dot{x}}(t, x, \dot{x}) = 2t\dot{x},$$

$$L_{tt}(t, x, \dot{x}) = -2x^2, \quad L_{tx}(t, x, \dot{x}) = -4tx, \quad L_{t\dot{x}}(t, x, \dot{x}) = 2\dot{x},$$

$$L_{xt}(t, x, \dot{x}) = -4tx, \quad L_{xx}(t, x, \dot{x}) = -2t^2, \quad L_{x\dot{x}}(t, x, \dot{x}) = 0,$$

$$L_{\dot{x}t}(t, x, \dot{x}) = 2\dot{x}, \quad L_{\dot{x}x}(t, x, \dot{x}) = 0, \quad L_{\dot{x}\dot{x}}(t, x, \dot{x}) = 2t.$$

Adicionalmente, si $T = [0, 1]$, $n = 2$ y $L\colon T \times \mathbf{R}^2 \times \mathbf{R}^2 \to \mathbf{R}$ está definida por

$$L(t, x, \dot{x}) := tx^1 + (x^2)^2 + x^2\dot{x}^1 + \operatorname{sen}(\dot{x}^2)^2,$$

entonces

$$L_t(t, x, \dot{x}) = x^1, \quad L_x(t, x, \dot{x}) = (t, 2x^2 + \dot{x}^1), \quad L_{\dot{x}}(t, x, \dot{x}) = (x^2, 2\dot{x}^2\cos(\dot{x}^2)^2),$$

$$L_{tt}(t, x, \dot{x}) = 0, \quad L_{tx}(t, x, \dot{x}) = (1, 0), \quad L_{t\dot{x}}(t, x, \dot{x}) = (0, 0),$$

$$L_{xt}(t, x, \dot{x}) = (1, 0), \quad L_{xx}(t, x, \dot{x}) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix}, \quad L_{x\dot{x}}(t, x, \dot{x}) = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix},$$

$$L_{\dot{x}t}(t, x, \dot{x}) = (0, 0), \quad L_{\dot{x}x}(t, x, \dot{x}) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad L_{\dot{x}\dot{x}}(t, x, \dot{x}) = \begin{pmatrix} 0 & 0 \\ 0 & 2\cos(\dot{x}^2)^2 - 4(\dot{x}^2)^2\operatorname{sen}(\dot{x}^2)^2 \end{pmatrix}.$$

El *problema clásico de cálculo de variaciones con puntos fijos finales* el cual denotaremos por (P), consiste en minimizar la funcional

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt$$

sujeta a

(a) $x\colon T \to \mathbf{R}^n$ es C^1 a pedazos.

(b) $x(t_0) = \xi_0$ y $x(t_1) = \xi_1$.

3.14 Observación: *Nótese que como x es C^1 a pedazos en T, entonces existe la posibilidad de que x tenga un número finito de esquinas o picos en algunos puntos denotados por $t^1, \ldots, t^m$ que se encuentran en el intervalo abierto (t_0, t_1) y así la derivada de x en esos puntos no existe. De hecho, sabemos que para toda $i = 1, \ldots, m$, $\dot{x}(t^i - 0)$ y $\dot{x}(t^i + 0)$ existen. En la teoría clásica del cálculo de variaciones es usual hacer el convenio de que en un punto esquina $t = \bar{t}$ de x, la notación $\dot{x}(\bar{t})$ denota tanto a $\dot{x}(\bar{t} - 0)$ como a $\dot{x}(\bar{t} + 0)$. Por lo tanto, en un sentido estricto la aplicación $t \mapsto \dot{x}(t)$, en general, no es una función. Sin embargo, en la teoría clásica del cálculo de variaciones es estándar hacer el convenio de que en un punto esquina $t = \bar{t}$,*

$\dot{x}(\bar{t})$ representa las dos derivadas $\dot{x}(\bar{t}-0)$ y $\dot{x}(\bar{t}+0)$ y que la aplicación $t \mapsto \dot{x}(t)$ es llamada una función. Por ejemplo, con este convenio en mente podemos observar que la función

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

es continua a pedazos en T y por lo tanto ésta es Riemann integrable en T. De esta manera, la funcional I está bien definida para todas las trayectorias o arcos x que satisfacen la restricción (a).

3.15 Definición: *Un arco o trayectoria $x: T \to \mathbf{R}^n$ que satisface las restricciones (a) y (b) se le llama arco o trayectoria admisible.*

3.16 Definición: *Si $x: T \to \mathbf{R}^n$ es C^1 a pedazos, definimos*

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

$$\|x\|_1 := \sup_{t \in T} (|x(t)| + |\dot{x}(t)|).$$

3.17 Observación: *Si definimos*

$$X := \{x: T \to \mathbf{R}^n \mid x \text{ es } C^1 \text{ a pedazos}\},$$

obsérvese que las funciones $\|\cdot\|_0: X \to \mathbf{R}$, $\|\cdot\|_1: X \to \mathbf{R}$ son normas en X, es decir, $(X, \|\cdot\|_0)$ y $(X, \|\cdot\|_1)$ son espacios vectoriales normados.

3.18 Definición: *Un arco admisible x es una solución global de (P) si $I(x) \leq I(y)$ para toda trayectoria admisible y. Diremos que x es un mínimo fuerte de (P) si existe $\epsilon > 0$ tal que $I(x) \leq I(y)$ para toda trayectoria admisible y que satisface $\|y - x\|_0 < \epsilon$. Diremos que x es un mínimo débil de (P) si existe $\epsilon > 0$ tal que $I(x) \leq I(y)$ para toda trayectoria admisible y que satisface $\|y - x\|_1 < \epsilon$. Diremos que el arco admisible x es una solución global estricta de (P) si x es una solución global de (P) y además $I(x) = I(y)$ sólo en el caso en que $x = y$. Las definiciones para mínimos estrictos fuertes y débiles se establecen análogamente.*

3.19 Observación: *Obsérvese que si x es una solución global de (P), entonces x es un mínimo fuerte de (P) y, a su vez, si x es un mínimo fuerte de (P), entonces x es un mínimo débil de (P).*

3. Condiciones necesarias de Euler, Weierstrass y Legendre

En esta sección enunciaremos tres teoremas fundamentales que dan condiciones necesarias de primer orden para óptimos locales. Estos resultados son los Teoremas 3.20, 3.22 y 3.23. Antes de hacer sus demostraciones, estableceremos algunos corolarios que se derivan de éstos. El Teorema 3.20 es uno de los teoremas más importantes en la teoría del cálculo de variaciones. La ecuación (3.1) se conoce como *la forma integral de la ecuación de Euler*. Como veremos más adelante, en general, esta ecuación integral es equivalente a una ecuación diferencial de segundo grado. Puesto que el área de las ecuaciones diferenciales ha sido ampliamente estudiada, cuando la forma integral de la ecuación de Euler es equivalente a una ecuación diferencial, ésta se convierte en una herramienta muy cómoda para buscar candidatos que pueden resolver el problema (P).

3.20 Teorema (Euler): *Supongamos que x_0 es un mínimo débil de (P). Entonces existe una constante $c \in \mathbf{R}^n$ tal que*

$$L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^{t} L_x(s, x_0(s), \dot{x}_0(s)) ds + c^* \quad (t \in T). \tag{3.1}$$

3.21 Definición: *Definamos $E: T \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ por*

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

A la función E se le llama la función exceso de Weierstrass de L.

El teorema siguiente proporciona otra condición necesaria de primer orden para un mínimo débil.

3.22 Teorema (Weierstrass): *Sea x_0 un mínimo débil de (P). Entonces existe $\sigma > 0$ tal que*

$$E(t, x_0(t), \dot{x}_0(t), u) \geq 0 \quad \textit{para toda } (t, u) \in T \times \mathbf{R}^n \textit{ con } |u - \dot{x}_0(t)| < \sigma.$$

Otro resultado que involucra nuevamente a la función exceso de Weierstrass y que ahora proporciona una condición necesaria de primer orden para un mínimo fuerte se encuentra en el siguiente:

3.23 Teorema (Weierstrass): *Sea x_0 un mínimo fuerte de (P). Entonces*

$$E(t, x_0(t), \dot{x}_0(t), u) \geq 0 \quad \textit{para toda } (t, u) \in T \times \mathbf{R}^n.$$

3.24 Observación: *Obsérvese que en un punto esquina de x_0, en el espacio-$tx\dot{x}$ hay dos elementos,*

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

También, la ecuación (3.1) es la forma integral de la ecuación de Euler

$$\frac{d}{dt} L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = L_x(t, x_0(t), \dot{x}_0(t)) \quad (t \in T). \tag{3.2}$$

En un punto esquina de x_0 la derivada d/dt se interpreta como una derivada izquierda o derecha. La ecuación (3.2) se cumple aun cuando x_0 no tenga una segunda derivada. Si x es C^2 en T y L es C^2 en $T \times \mathbf{R}^n \times \mathbf{R}^n$, la ecuación de Euler se convierte en

$$L^*_{\dot{x}t} + L_{\dot{x}x}\dot{x}(t) + L_{\dot{x}\dot{x}}\ddot{x}(t) = L^*_x \quad (t \in T)$$

donde los argumentos en las derivadas de L son $(t, x(t), \dot{x}(t))$. Por lo tanto, si x es C^2 en T y L es C^2 en $T \times \mathbf{R}^n \times \mathbf{R}^n$, entonces la ecuación de Euler es una ecuación diferencial de segundo grado y de esta manera dicha ecuación es una herramienta fundamental para obtener candidatos x a resolver el problema (P).

3.25 Corolario: *Si x_0 satisface la forma integral de la ecuación de Euler y la condición de Weierstrass para un mínimo débil, entonces la función*

$$t \mapsto L(t, x_0(t), \dot{x}_0(t)) - L_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\dot{x}_0(t)$$

es continua en T.

Demostración: Definamos $p: T \to \mathbf{R}^n$ por

$$p(t) := L^*_{\dot{x}}(t, x_0(t), \dot{x}_0(t)).$$

Puesto que x_0 satisface la forma integral de la ecuación de Euler, entonces para alguna constante $c \in \mathbf{R}^n$,

$$p(t) = \int_{t_0}^{t} L^*_x(s, x_0(s), \dot{x}_0(s))ds + c \quad (t \in T).$$

De esta manera, p es una función C^1 a pedazos en T y en consecuencia p es una función continua en T. Por lo tanto, tenemos la condición de esquina

$$p(t-0) = p(t+0) \tag{3.3}$$

para toda $t \in T$. La condición (3.3) es comúnmente llamada la *condición de esquina de Weierstrass-Erdmann.* Para finalizar con la demostración del corolario, definamos $F: T \times \mathbf{R}^n \to \mathbf{R}$ por

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

Nótese que es suficiente demostrar que para toda $t \in T$,

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

Como x_0 satisface la condición de Weierstrass para un mínimo débil, entonces existe $\sigma > 0$ tal que

$$E(t, x_0(t), \dot{x}_0(t), u) \geq 0 \quad \text{para toda } (t,u) \in T \times \mathbf{R}^n \text{ con } |u - \dot{x}_0(t-0)| < \sigma \text{ o } |u - \dot{x}_0(t+0)| < \sigma.$$

En particular,

$$E(t, x_0(t), \dot{x}_0(t-0), u) \geq 0 \quad \text{para toda } (t,u) \in T \times \mathbf{R}^n \text{ con } |u - \dot{x}_0(t+0)| < \sigma,$$

y

$$E(t, x_0(t), \dot{x}_0(t+0), u) \geq 0 \quad \text{para toda } (t,u) \in T \times \mathbf{R}^n \text{ con } |u - \dot{x}_0(t-0)| < \sigma.$$

Entonces,

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

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

Accidentalmente, hemos demostrado el siguiente resultado.

3.26 Corolario: *En un punto esquina de un arco admisible x_0 que satisface la forma integral de la ecuación de Euler y la condición de Weierstrass para un mínimo débil, se tiene que*

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

El siguiente resultado da una nueva condición necesaria de segundo orden. La condición que se obtiene al reemplazar el mayor o igual que por el mayor que se le conoce como la condición *reforzada de Legendre.* Como veremos más adelante, esta condición reforzada jugará un papel fundamental en algunas condiciones necesarias de segundo orden, así como en la teoría clásica de condiciones suficientes para un óptimo.

3.27 Corolario (Legendre): *Si x_0 satisface la condición de Weierstrass para un mínimo débil, entonces la matriz Hessiana $L_{\dot{x}\dot{x}}$ es semi-definida positiva a lo largo de x_0, esto es, para toda $h \in \mathbf{R}^n$ y para toda $t \in T$,*

$$\langle h, L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t))h\rangle \geq 0. \tag{3.4}$$

Demostración: Sea $t = \bar{t}$ un punto que no corresponda a una esquina de x_0 y sea $\sigma > 0$ el número que aparece en la correspondiente condición de Weierstrass. Definamos $\bar{x} := x_0(\bar{t})$, $\bar{u} := \dot{x}_0(\bar{t})$ y $G: B(\bar{u}, \sigma) \to \mathbf{R}$ por

$$G(u) := E(\bar{t}, \bar{x}, \bar{u}, u).$$

Obsérvese que $G(\bar{u}) = 0$ y por la condición de Weierstrass $G(u) \geq 0$ para toda $u \in B(\bar{u}, \sigma)$. Por lo tanto,

$$0 \leq G''(\bar{u}) = L_{\dot{x}\dot{x}}(\bar{t}, \bar{x}, \bar{u}).$$

En consecuencia, la desigualdad (3.4) se cumple en todos los puntos en los que x_0 no tiene esquinas. Por continuidad, (3.4) se cumple también en los puntos esquina de x_0. $\blacksquare$

La condición (3.4) se conoce como la *condición de Legendre.*

3.28 Definición: *A una solución x_0 de la forma integral de la ecuación de Euler se le llamará extremal. A un extremal sin esquinas se le llamará extremo. Un extremal está constituido de un número finito de subarcos los cuales son extremos. A un arco admisible x_0 se le dirá no singular si el determinante*

$$|L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t))| \neq 0$$

para toda $t \in T$.

3.29 Observación: *Puede suceder que la forma integral de la ecuación de Euler no sea equivalente a la ecuación de Euler. De hecho, si un arco x satisface la forma integral de la ecuación de Euler, entonces x satisface la ecuación de Euler. El recíproco, puede ser falso si la función $t \mapsto L_{\dot{x}}(t, x(t), \dot{x}(t))$ no es continua en T. Por ejemplo, si $T = [0, 2\pi]$, $n = 1$, $L(t, x, \dot{x}) = \dot{x}^2 - x^2$, $\xi_0 = \xi_1 = 0$, entonces como uno puede verificar fácilmente la ecuación de Euler se convierte en*

$$\frac{d}{dt}\dot{x}(t) = -x(t) \quad (t \in T).$$

El arco admisible

$$x(t) := \begin{cases} \operatorname{sen} t & si\ t \in [0, \pi] \\ 0 & si\ t \in [\pi, 2\pi] \end{cases}$$

claramente satisface la ecuación de Euler, pero x no satisface la forma integral de la ecuación de Euler la cual está dada por

$$\dot{x}(t) = \int_0^t -x(s)ds + c \quad (t \in T),$$

puesto que $\dot{x}$ es discontinua en T, mientras que para toda $c \in \mathbf{R}$, la función

$$\int_0^t -x(s)ds + c \quad (t \in T),$$

es continua en T. Sin embargo, obsérvese que si la función $t \mapsto L_{\dot{x}}(t, x(t), \dot{x}(t))$ es continua en T, entonces si x satisface la ecuación de Euler sucede que x también satisface la forma integral de la ecuación de Euler. En efecto, denotemos por κ a la función $t \mapsto L_{\dot{x}}(t, x(t), \dot{x}(t))$, supongamos que κ es continua en T y que x satisface la ecuación de Euler. Por simplicidad en la notación, supongamos que κ tiene una sola esquina en $t = \bar{t}$ que pertenece a (t_0, t_1). Puesto que κ es continua en $t = \bar{t}$, y κ es C^1 en los intervalos $[t_0, \bar{t}]$ y $[\bar{t}, t_1]$, obsérvese que κ se puede escribir como

$$\kappa(t) = \begin{cases} \int_{t_0}^t \dot{\kappa}(s)ds + \kappa(t_0) & si\ t \in [t_0, \bar{t}] \\ \int_{\bar{t}}^t \dot{\kappa}(s)ds + \kappa(\bar{t}) & si\ t \in [\bar{t}, t_1]. \end{cases}$$

Entonces

$$\kappa(\bar{t}) = \int_{t_0}^{\bar{t}} \dot{\kappa}(s)ds + \kappa(t_0),$$

por lo que para toda $t \in [\bar{t}, t_1]$,

$$\kappa(t) = \int_{\bar{t}}^t \dot{\kappa}(s)ds + \kappa(\bar{t}) = \int_{\bar{t}}^t \dot{\kappa}(s)ds + \int_{t_0}^{\bar{t}} \dot{\kappa}(s)ds + \kappa(t_0) = \int_{t_0}^t \dot{\kappa}(s)ds + \kappa(t_0).$$

Así, $\kappa(t) = \int_{t_0}^t \dot{\kappa}(s)ds + \kappa(t_0)$ $(t \in T)$, lo cual implica que κ es C^1 a pedazos en T y por lo tanto κ es absolutamente continua en T. Consecuentemente,

$$L_{\dot{x}}(t, x(t), \dot{x}(t)) = \int_{t_0}^t L_x(s, x(s), \dot{x}(s))ds + L_{\dot{x}}(t_0, x(t_0), \dot{x}(t_0)) \quad (t \in T),$$

esto es, x satisface la forma integral de la ecuación de Euler.

4. Teorema de Diferenciabilidad de Hilbert

El Teorema 3.32 nos da algunas condiciones que con frecuencia nos ayudan a asegurar ciertas propiedades de suavidad de un arco no singular que satisface la ecuación de Euler. Este resultado, en muchos ocasiones, se convierte en una herramienta muy útil para verificar que la ecuación de Euler es una ecuación diferencial de segundo grado.

3.30 Teorema (Teorema de la función implícita): *Sea S un subconjunto abierto de $\mathbf{R}^{m+n}$. Sea $f(t,x)\colon S \to \mathbf{R}^n$ continua en S. Supongamos que $f_x(t,x)\colon S \to \mathbf{R}^{n\times n}$ es continua en S y asumamos que para algún punto $(t_0, x_0) \in S$, se tiene que*

$$f(t_0, x_0) = 0 \quad y \quad |f_x(t_0, x_0)| \neq 0.$$

Entonces, existen $\delta, \epsilon > 0$ y $x\colon B(t_0, \epsilon) \to \mathbf{R}^n$ continua tales que

$$x(t_0) = x_0 \quad y \quad f(t, x(t)) = 0 \quad (t \in B(t_0, \epsilon))$$

y tales que las relaciones

$$f(t,x) = 0,\ |x - x(t)| < \delta\ (t \in B(t_0, \epsilon)) \Longrightarrow x = x(t).$$

Si la función f es C^m en S, la función x es C^m en $B(t_0, \epsilon)$.

3.31 Definición: *De ahora en adelante, si decimos que L es de clase C^m en $T \times \mathbf{R}^n \times \mathbf{R}^n$ lo que estamos diciendo es que L es de clase C^m en una vecindad de $T \times \mathbf{R}^n \times \mathbf{R}^n$, es decir, L y todas sus derivadas de orden menor o igual que m existen y son continuas en un conjunto de la forma $(t_0 - \epsilon, t_1 + \epsilon) \times \mathbf{R}^n \times \mathbf{R}^n$ para alguna $\epsilon > 0$.*

3.32 Teorema (Teorema de Diferenciabilidad de Hilbert): *Un extremo no singular x_0 es C^m $(m \geq 2)$ en T si L es C^m en $T \times \mathbf{R}^n \times \mathbf{R}^n$. Entre esquinas un extremal no singular x_0 es C^m en T si L es C^m en $T \times \mathbf{R}^n \times \mathbf{R}^n$.*

Demostración: Supongamos que L es C^m en $T \times \mathbf{R}^n \times \mathbf{R}^n$ y sea $\epsilon > 0$ como en la Definición 3.31. Sea $t = \bar{t}$ un punto en T que no corresponde a una esquina de x_0. Sea $\gamma > 0$ tal que $(\bar{t} - \gamma, \bar{t} + \gamma) \subset (t_0 - \epsilon, t_1 + \epsilon)$ y tal que este intervalo no contenga puntos en los que x_0 tenga esquinas. Nótese que si $\bar{t}$ es igual a t_0 o a t_1, puesto que x_0 es C^1 en $[t_0, t_0 + \gamma)$ y en $(t_1 - \gamma, t_1]$, sin ninguna dificultad a x_0 se le puede extender a los intervalos $(t_0 - \gamma, t_0 + \gamma)$ y $(t_1 - \gamma, t_1 + \gamma)$ de tal forma que x_0 sea de clase C^1 en dichos intervalos. Sea $F\colon (t_0 - \gamma, t_1 + \gamma) \times \mathbf{R}^n \to \mathbf{R}^n$ definida por

$$F(t,u) := L^*_{\dot{x}}(t, x_0(t), u) - \int_{t_0}^{t} L^*_x(s, x_0(s), \dot{x}_0(s))ds - c,$$

donde $c \in \mathbf{R}^n$ es la constante dada por la forma integral de la ecuación de Euler. Se tiene que

$$F_u(t,u) = L_{\dot{x}\dot{x}}(t, x_0(t), u).$$

Por lo tanto, las funciones F y F_u son continuas en el abierto $(\bar{t} - \gamma, \bar{t} + \gamma) \times \mathbf{R}^n$. También, si definimos $\bar{x} := x_0(\bar{t})$ y $\bar{u} := \dot{x}_0(\bar{t})$, tenemos que

$$F(\bar{t}, \bar{u}) = 0, \quad |F_u(\bar{t}, \bar{u})| = |L_{\dot{x}\dot{x}}(\bar{t}, \bar{x}, \bar{u})| \neq 0.$$

Demostremos que

$$x_0 \in C^1 \text{ en } (\bar{t} - \gamma, \bar{t} + \gamma) \Longrightarrow x_0 \in C^2 \text{ en } (\bar{t} - \gamma, \bar{t} + \gamma).$$

En efecto, si $x_0 \in C^1$ en $(\bar{t} - \gamma, \bar{t} + \gamma)$, entonces las funciones

$$F_t(t,u) = L^*_{\dot{x}t}(t, x_0(t), u) + L_{\dot{x}x}(t, x_0(t), u)\dot{x}_0(t) - L^*_x(t, x_0(t), \dot{x}_0(t)),$$

$$F_u(t,u) = L_{\dot{x}\dot{x}}(t, x_0(t), u),$$

son continuas en $(\bar{t}-\gamma, \bar{t}+\gamma) \times \mathbf{R}^n$, esto es, $F \in C^1$ en $(\bar{t}-\gamma, \bar{t}+\gamma) \times \mathbf{R}^n$. Por el Teorema 3.30, podemos disminuir γ de tal forma que $\dot{x}_0 \colon (\bar{t}-\gamma, \bar{t}+\gamma) \to \mathbf{R}^n$ sea la única función continua que en $t = \bar{t}$ tenga el valor $\bar{u}$ y que satisfaga

$$F(t, \dot{x}_0(t)) = 0 \quad (t \in (\bar{t}-\gamma, \bar{t}+\gamma)).$$

Nuevamente por el Teorema 3.30, $\dot{x}_0 \in C^1$ en $(\bar{t}-\gamma, \bar{t}+\gamma)$. Y así, $x_0 \in C^2$ en $(\bar{t}-\gamma, \bar{t}+\gamma)$ lo cual demuestra nuestra afirmación. Similarmente, se puede ver sin dificultad que $x_0 \in C^k$ $(1 \leq k < m; \bar{t}-\gamma < t < \bar{t}+\gamma)$ implica que $x_0 \in C^{k+1}$ en $\bar{t}-\gamma < t < \bar{t}+\gamma$. De esta manera, $L \in C^m$ $(m \geq 2)$ en $T \times \mathbf{R}^n \times \mathbf{R}^n$ implica que $x_0 \in C^m$ en T. ▮

5. Positividad regular

La *positividad regular* es una propiedad de la integral I que nos permite asegurar en muchas ocasiones que las posibles soluciones del problema (P) son no singulares y que no tienen esquinas o picos. Con frecuencia, uno puede apelar al teorema de diferenciabilidad de Hilbert para concluir que las soluciones deben de ser de clase C^m con $m \geq 2$ y así tener la ventaja de obtener todas las posibles soluciones del problema (P) al encontrar la solución general C^2 de la ecuación de Euler puesto que en este caso dicha ecuación es una ecuación diferencial de segundo grado.

3.33 Definición: *La integral I se llamará positiva regular si*

$$\langle h, L_{\dot{x}\dot{x}}(t,x,\dot{x})h \rangle > 0$$

para toda $(t,x,\dot{x}) \in T \times \mathbf{R}^n \times \mathbf{R}^n$ y para toda $h \neq 0$.

3.34 Teorema: *Si I es positiva regular, entonces*

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

para toda $(t,x,\dot{x},u) \in T \times \mathbf{R}^n \times \mathbf{R}^n \times \mathbf{R}^n$ con $u \neq \dot{x}$. Además, todo extremal es un extremo no singular.

Demostración: Obsérvese que

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

lo cual demuestra la primera afirmación del teorema. Para demostrar la segunda afirmación, nótese que si x_0 es un extremal, entonces por la primera conclusión del teorema, x_0 satisface la forma integral de la ecuación de Euler y la condición de Weierstrass para un mínimo débil. Por lo tanto, para toda $t \in T$,

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

De esta manera, para toda $t \in T$, $\dot{x}_0(t-0) = \dot{x}_0(t+0)$. ▮

6. Ejemplos

En el Ejemplo 3.35 mostramos cómo se pueden aplicar la positividad regular y el teorema de diferenciabilidad de Hilbert a un problema (P). También, ilustramos que las condicones necesarias de Euler y Weierstrass no son suficientes para obtener una solución de (P). Adicionalmente, demostramos que el problema (P) bajo consideración no tiene solución.

3.35 Ejemplo: *Consideremos el problema (P) de minimizar*

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

sujeta a

(a) $x\colon [0,b] \to \mathbf{R}$ *es* C^1 *a pedazos.*

(b) $x(0) = x(b) = 0$.

En este caso $T = [0,b]$, $n = 1$, $\xi_0 = \xi_1 = 0$, y $L(t,x,\dot{x}) = \dot{x}^2 - x^2$. Puesto que $L_x(t,x,\dot{x}) = -2x$ y $L_{\dot{x}}(t,x,\dot{x}) = 2\dot{x}$, entonces la forma integral de la ecuación de Euler se convierte en

$$\dot{x}(t) = \int_0^t -x(s)ds + c \quad (t \in T)$$

para alguna constante $c \in \mathbf{R}$. Puesto que $L_{\dot{x}\dot{x}}(t,x,\dot{x}) = 2$ para toda $(t,x,\dot{x}) \in T \times \mathbf{R} \times \mathbf{R}$, entonces la integral I es positiva regular y así todo extremal es un extremo no singular. Por otro lado, obsérvese que L es C^∞ en $T \times \mathbf{R} \times \mathbf{R}$ y entonces por el teorema de diferenciabilidad de Hilbert, todo extremo debe de ser de clase C^∞ en T. En consecuencia, un extremo satisface la forma integral de la ecuación de Euler si y solo si éste satisface la ecuación de Euler la cual toma la forma

$$\ddot{x}(t) + x(t) = 0 \quad (t \in T).$$

La solución general C^2 de esta ecuación diferencial, es de la forma

$$x(t) = a\cos t + b \operatorname{sen} t \quad (t \in T; a, b \in \mathbf{R}).$$

La ecuación de Weierstrass toma la forma

$$E(t,x,\dot{x},u) = (u - \dot{x})^2.$$

Sin embargo, si $b > \pi$, el arco $x_0 \equiv 0$ el cual es admisible no es un mínimo global de (P). En efecto, sea $c \in (\pi, b]$, $c < 2\pi$, y sea $x\colon T \to \mathbf{R}$ definida por

$$x(t) := \begin{cases} \operatorname{sen} t, & t \in [0, c/2], \\ \operatorname{sen}(c - t), & t \in [c/2, c], \\ 0, & t \in [c, b]. \end{cases} \tag{3.5}$$

Obsérvese que x es admisible y además,

$$\begin{aligned} I(x) &= \int_0^{c/2} \{\cos^2 t - \operatorname{sen}^2 t\} dt + \int_{c/2}^c \{\cos^2(c-t) - \operatorname{sen}^2(c-t)\} dt \\ &= 2\int_0^{c/2} \{\cos^2 t - \operatorname{sen}^2 t\} dt = 2\int_0^{c/2} \cos 2t dt = \operatorname{sen} c < 0 = I(x_0). \end{aligned}$$

Por lo tanto, deberíamos de notar que las condiciones necesarias de Euler y Weierstrass no son suficientes para garantizar un mínimo para el problema (P).

Por otro lado como demostraremos más adelante, si x es un mínimo débil de (P), entonces

$$0 = I'(x,y) := \int_0^b \{L_x(t,x(t),\dot{x}(t))y(t) + L_{\dot{x}}(t,x(t),\dot{x}(t))\dot{y}(t)\}dt = 2\int_0^b \{\dot{x}(t)\dot{y}(t) - x(t)y(t)\}dt$$

para toda $y \in Y := \{y\colon T \to \mathbf{R} \mid y \text{ es } C^1 \text{ a pedazos y } y(0) = y(b) = 0\}$. Por lo tanto, si x es un mínimo global de (P), entonces

$$0 = I'(x,x) = 2\int_0^b \{\dot{x}^2(t) - x^2(t)\}dt = 2I(x).$$

Puesto que $I(x_0) = 0$, entonces si x es un mínimo global de (P), también se tiene que x_0 es un mínimo global de (P), es decir

$$0 = I(x_0) \le I(y) \text{ para toda } y \text{ admisible}$$

lo cual no es el caso puesto que la trayectoria x dada en (3.5) es admisible e $I(x) < I(x_0)$. De esta manera, si $b > \pi$ el problema (P) no tiene mínimos globales. Como se puede ver sin dificultad si se sustituye a x por γx con $\gamma > 0$ suficientemente pequeña, uno puede concluir que el problema (P) tampoco tiene mínimos fuertes ni mínimos débiles, es decir, si $b > \pi$ el problema (P) bajo consideración no tiene solución.

En el Ejemplo 3.36 vemos desde una perspectiva diferente cómo un problema variacional (P), puede no tener solución.

3.36 Ejemplo: *Consideremos el problema (P) de minimizar*

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

sujeta a

(a) $x\colon [0,1] \to \mathbf{R}$ *es* C^1 *a pedazos.*

(b) $x(0) = 0$ *y* $x(1) = 1$.

En este caso $T = [0,1]$, $n = 1$, $\xi_0 = 0$, $\xi_1 = 1$, y $L(t,x,\dot{x}) = t^2\dot{x}^2$. Claramente, $I(x) \geq 0$ para toda x admisible. Para $\epsilon \in (0,1]$ los arcos

$$x_\epsilon(t) := \begin{cases} t/\epsilon, & t \in [0,\epsilon], \\ 1, & t \in [\epsilon, 1], \end{cases}$$

son admisibles y satisfacen que

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

En consecuencia,

$$\text{ínf}\{I(x) : x \text{ es admisible}\} = 0.$$

Por otro lado, como uno puede ver fácilmente, la forma integral de la ecuación de Euler se convierte en

$$t^2 \dot{x}(t) = c \quad (t \in T).$$

Haciendo $t = 0$ se tiene que $c = 0$. Consecuentemente, los únicos extremos definidos en el intervalo $0 \leq t \leq 1$ son aquellos para los cuales

$$x \equiv \text{constante}.$$

Sin embargo, ningún extremo de este tipo satisface que $x(0) = 0$ y $x(1) = 1$. De esta manera, el problema variacional descrito anteriormente no tiene solución. Nótese que en este problema se tiene que

$$L_{\dot{x}\dot{x}}(t,x,\dot{x}) = 2t^2$$

y por lo tanto, si $x\colon T \to \mathbf{R}$ es C^1 a pedazos,

$$L_{\dot{x}\dot{x}}(0, x(0), \dot{x}(0)) = 0.$$

Consecuentemente, uno tiene una singularidad en $t = 0$.

7. Demostración del Teorema de Weierstrass: mínimo débil

Ahora demostremos el teorema de Weierstrass para un mínimo débil de (P).

3.37 Teorema (Weierstrass): *Sea x_0 un mínimo débil de (P). Entonces existe $\sigma > 0$ tal que*

$$E(t, x_0(t), \dot{x}_0(t), u) \geq 0 \quad \textit{para toda } (t,u) \in T \times \mathbf{R}^n \textit{ con } |u - \dot{x}_0(t)| < \sigma.$$

Demostración: Consideremos un punto $\bar{t} \in (t_0, t_1)$ el cual no corresponde a un punto esquina de x_0. Definimos $\bar{x} := x_0(\bar{t})$, $\bar{u} := \dot{x}_0(\bar{t})$, y seleccionemos $u \in \mathbf{R}^n$ y $\sigma > 0$ tales que $0 < |u - \bar{u}| < \sigma$. Escogemos $\delta_0 > 0$ tal que $\bar{t} + \delta_0 < t_1$. Para $0 \leq \delta \leq \delta_0$, $0 < \epsilon < 1/2$, definimos

$$x(t, \epsilon, \delta) := \begin{cases} x_0(t), & t \in [t_0, \bar{t}] \cup [\bar{t} + \delta, t_1] \\ x_0(t) + (t - \bar{t})(u - \bar{u}), & t \in [\bar{t}, \bar{t} + \epsilon\delta] \\ x_0(t) + [\epsilon/(1 - \epsilon)](u - \bar{u})(\bar{t} + \delta - t), & t \in [\bar{t} + \epsilon\delta, \bar{t} + \delta]. \end{cases}$$

Observemos que $x(\cdot, \epsilon, \delta)$ es admisible para $(0 \leq \delta \leq \delta_0, 0 < \epsilon < 1/2)$. Puesto que x_0 es un mínimo débil de (P), existe $\gamma > 0$ tal que si x es admisible,

$$[\|x - x_0\|_0 < \gamma \quad \text{y} \quad \|\dot{x} - \dot{x}_0\|_0 < \gamma] \Longrightarrow I(x_0) \leq I(x).$$

Uno puede verificar fácilmente que para $0 \leq \delta \leq \delta_0$, $0 < \epsilon < 1/2$,

$$\|x(\cdot, \epsilon, \delta) - x_0\|_0 < \gamma, \quad \|\dot{x}(\cdot, \epsilon, \delta) - \dot{x}_0\|_0 < \gamma,$$

siempre que σ sea suficientemente pequeña. En consecuencia, si σ es suficientemente pequeña,

$$0 \leq I(x(\cdot, \epsilon, \delta)) - I(x_0) = F(\epsilon, \delta) + G(\epsilon, \delta), \tag{3.6}$$

donde

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

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

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

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

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

Por (3.6),

$$L(\bar{t}, \bar{x}, u) - L(\bar{t}, \bar{x}, \bar{u}) + \frac{L(\bar{t}, \bar{x}, \bar{u} - \eta(u - \bar{u})) - L(\bar{t}, \bar{x}, \bar{u})}{\eta} \geq 0.$$

Tomando el límite cuando $\eta \to 0$, obtenemos

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

Por lo tanto, la condición de Weierstrass se satisface a lo largo de x_0, excepto posiblemente en los puntos finales y esquina de x_0. Por continuidad, ésta también se satisface en estos puntos. ▮

8. Demostración del Teorema de Weierstrass: mínimo fuerte

La demostración del Teorema 3.38 es muy similar a la del Teorema 3.37, sin embargo, es fundamental detectar las diferencias.

3.38 Teorema (Weierstrass): *Sea x_0 un mínimo fuerte de (P). Entonces*

$$E(t, x_0(t), \dot{x}_0(t), u) \geq 0 \quad \textit{para toda } (t, u) \in T \times \mathbf{R}^n.$$

Demostración: Consideremos un punto $\bar{t} \in (t_0, t_1)$ el cual no corresponde a un punto esquina de x_0. Definimos $\bar{x} := x_0(\bar{t})$, $\bar{u} := \dot{x}_0(\bar{t})$, y seleccionemos $u \in \mathbf{R}^n$. Escogemos $\delta_0 > 0$ tal que $\bar{t} + \delta_0 < t_1$. Para $0 \leq \delta \leq \delta_0$, $0 < \epsilon < 1$, definimos

$$x(t, \epsilon, \delta) := \begin{cases} x_0(t), & t \in [t_0, \bar{t}] \cup [\bar{t} + \delta, t_1] \\ x_0(t) + (t - \bar{t})(u - \bar{u}), & t \in [\bar{t}, \bar{t} + \epsilon\delta] \\ x_0(t) + [\epsilon/(1-\epsilon)](u - \bar{u})(\bar{t} + \delta - t), & t \in [\bar{t} + \epsilon\delta, \bar{t} + \delta]. \end{cases}$$

Observemos que $x(\cdot, \epsilon, \delta)$ es admisible para $(0 < \epsilon < 1, 0 \leq \delta \leq \delta_0)$. Puesto que x_0 es un mínimo fuerte de (P), existe $\gamma > 0$ tal que si x es admisibe,

$$\|x - x_0\|_0 < \gamma \Longrightarrow I(x_0) \leq I(x).$$

Uno puede verificar fácilmente que para $0 \leq \delta \leq \delta_0$, $0 < \epsilon < 1$,

$$\|x(\cdot, \epsilon, \delta) - x_0\|_0 < \gamma,$$

siempre que ϵ sea suficientemente pequeña. En consecuencia, si ϵ es suficientemente pequeña,

$$0 \leq I(x(\cdot, \epsilon, \delta)) - I(x_0) = F(\epsilon, \delta) + G(\epsilon, \delta), \tag{3.7}$$

donde

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

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

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

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

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

Por (3.7),

$$L(\bar{t}, \bar{x}, u) - L(\bar{t}, \bar{x}, \bar{u}) + \frac{L(\bar{t}, \bar{x}, \bar{u} - \eta(u - \bar{u})) - L(\bar{t}, \bar{x}, \bar{u})}{\eta} \geq 0.$$

Tomando el límite cuando $\eta \to 0$, obtenemos

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

Por lo tanto, la condición de Weierstrass se satisface a lo largo de x_0, excepto posiblemente en los puntos finales y esquina de x_0. Por continuidad, ésta también se satisface en estos puntos. ▮

9. La primera y segunda variaciones de I

En esta sección demostraremos el Teorema 3.20. Además, proporcionaremos dos nuevas condiciones necesarias para un mínimo débil de (P). Estas condiciones se conocen como la anulación de la primera variación y la no negatividad de la segunda variación sobre el conjunto de variaciones admisibles.

Supongamos que L es de clase C^2 en $T \times \mathbf{R}^n \times \mathbf{R}^n$, sea x un arco admisible del problema (P) y sea $y: T \to \mathbf{R}^n$ una función C^1 a pedazos. La función $F: \mathbf{R} \to \mathbf{R}$ definida por

$$F(\epsilon) := I(x + \epsilon y) = \int_{t_0}^{t_1} L(t, x(t) + \epsilon y(t), \dot{x}(t) + \epsilon \dot{y}(t)) dt$$

es de clase C^2 en $\mathbf{R}$. En efecto, por la regla de Leibniz, para toda $\epsilon \in \mathbf{R}$,

$$\begin{aligned} F'(\epsilon) &= \frac{d}{d\epsilon}\int_{t_0}^{t_1} L(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))dt \\ &= \int_{t_0}^{t_1}\{L_x(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))y(t) + L_{\dot{x}}(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))\dot{y}(t)\}dt. \end{aligned}$$

Nuevamente por la regla de Leibniz, para toda $\epsilon \in \mathbf{R}$,

$$\begin{aligned} F''(\epsilon) &= \frac{d}{d\epsilon}\int_{t_0}^{t_1}\{L_x(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))y(t) + L_{\dot{x}}(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))\dot{y}(t)\} \\ &= \int_{t_0}^{t_1}\{\langle y(t), L_{xx}(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))y(t)\rangle \\ &\qquad + 2\langle y(t), L_{x\dot{x}}(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))\dot{y}(t)\rangle \\ &\qquad + \langle \dot{y}(t), L_{\dot{x}\dot{x}}(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))\dot{y}(t)\rangle\}dt. \end{aligned}$$

Veamos que F'' es continua en $\mathbf{R}$. Para demostrarlo, sea $\epsilon \in \mathbf{R}$ fija y sea $\{\epsilon_n\}$ una sucesión de números reales tales que $\lim_{n\to\infty} \epsilon_n = \epsilon$. Por las hipótesis de suavidad de L y el teorema de la convergencia acotada,

$$\begin{aligned} \lim_{n\to\infty} F''(\epsilon_n) &= \lim_{n\to\infty}\int_{t_0}^{t_1}\{\langle y(t), L_{xx}(t, x(t)+\epsilon_n y(t), \dot{x}(t)+\epsilon_n\dot{y}(t))y(t)\rangle \\ &\qquad + 2\langle y(t), L_{x\dot{x}}(t, x(t)+\epsilon_n y(t), \dot{x}(t)+\epsilon_n\dot{y}(t))\dot{y}(t)\rangle \\ &\qquad + \langle \dot{y}(t), L_{\dot{x}\dot{x}}(t, x(t)+\epsilon_n y(t), \dot{x}(t)+\epsilon_n\dot{y}(t))\dot{y}(t)\rangle\}dt \\ &= \int_{t_0}^{t_1}\lim_{n\to\infty}\{\langle y(t), L_{xx}(t, x(t)+\epsilon_n y(t), \dot{x}(t)+\epsilon_n\dot{y}(t))y(t)\rangle \\ &\qquad + 2\langle y(t), L_{x\dot{x}}(t, x(t)+\epsilon_n y(t), \dot{x}(t)+\epsilon_n\dot{y}(t))\dot{y}(t)\rangle \\ &\qquad + \langle \dot{y}(t), L_{\dot{x}\dot{x}}(t, x(t)+\epsilon_n y(t), \dot{x}(t)+\epsilon_n\dot{y}(t))\dot{y}(t)\rangle\}dt \\ &= \int_{t_0}^{t_1}\{\langle y(t), L_{xx}(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))y(t)\rangle \\ &\qquad + 2\langle y(t), L_{x\dot{x}}(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))\dot{y}(t)\rangle \\ &\qquad + \langle \dot{y}(t), L_{\dot{x}\dot{x}}(t, x(t)+\epsilon y(t), \dot{x}(t)+\epsilon\dot{y}(t))\dot{y}(t)\rangle\}dt \\ &= F''(\epsilon) \end{aligned}$$

lo cual demuestra la afirmación. De esta manera, F es una función de clase C^2 en $\mathbf{R}$.

3.39 Definición: *Definimos a la primera y segunda variaciones de I a lo largo de x y en la dirección y por $F'(0)$ y $F''(0)$ respectivamente. Utilizaremos las notaciones $I'(x,y)$ e $I''(x,y)$ para denotar a la primera y segunda variaciones de I respectivamente. De esta manera, obsérvese que*

$$\begin{aligned} I'(x,y) &= \int_{t_0}^{t_1}\{L_x(t, x(t), \dot{x}(t))y(t) + L_{\dot{x}}(t, x(t), \dot{x}(t))\dot{y}(t)\}dt, \\ I''(x,y) &= \int_{t_0}^{t_1}\{\langle y(t), L_{xx}(t, x(t), \dot{x}(t))y(t)\rangle \\ &\qquad + 2\langle y(t), L_{x\dot{x}}(t, x(t), \dot{x}(t))\dot{y}(t)\rangle \\ &\qquad + \langle \dot{y}(t), L_{\dot{x}\dot{x}}(t, x(t), \dot{x}(t))\dot{y}(t)\rangle\}dt. \end{aligned}$$

Ahora supongamos que x_0 es un mínimo débil de (P). Definamos

$$Y := \{y \in X \mid y(t_0) = y(t_1) = 0\} = \{y\colon T \to \mathbf{R}^n \mid y \text{ es } C^1 \text{ a pedazos y } y(t_0) = y(t_1) = 0\}.$$

Al conjunto Y se le conoce como el conjunto de *variaciones admisibles*. Dadas $y \in Y$ y $\gamma > 0$ fijas, nótese que los arcos $x_0 + \epsilon y$ son admisibles y además

$$\|x_0 + \epsilon y - x_0\|_1 = |\epsilon| \|y\|_1 < \gamma$$

siempre que ϵ sea suficientemente pequeña. Por lo tanto, para $\delta > 0$ suficientemente pequeña,

$$I(x_0) \leq I(x_0 + \epsilon y) \quad (-\delta < \epsilon < \delta).$$

Como lo hicimos anteriormente, si definimos $F\colon (-\delta, \delta) \to \mathbf{R}$ por

$$F(\epsilon) := \int_{t_0}^{t_1} L(t, x_0(t) + \epsilon y(t), \dot{x}_0(t) + \epsilon \dot{y}(t)) dt,$$

se tiene que

$$F(0) \leq F(\epsilon) \quad (-\delta < \epsilon < \delta)$$

con F de clase C^2 en $(-\delta, \delta)$. En consecuencia,

$$F'(0) = I'(x_0, y) = 0 \quad \text{y} \quad F''(0) = I''(x_0, y) \geq 0.$$

Lo anterior demuestra la primera conclusión del siguiente teorema. Suele suceder que la verificación de la anulación de la primera variación sobre el conjunto de variaciones admisibles tiene el mismo grado de dificultad que el de resolver el problema directamente. Puesto que la forma integral de la ecuación de Euler, en general, es una ecuación diferencial de segundo grado, una componente crucial del Teorema 3.40, es la caracterización entre la anulación de la primera variación sobre el conjunto de variaciones admisibles y la forma integral de la ecuación de Euler.

3.40 Teorema: *Si un arco admisible x_0 es un mínimo débil de (P), entonces*

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

para toda $y \in Y$. La relación

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

se cumple en Y si y solo si para alguna $c \in \mathbf{R}^n$

$$L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^{t} L_x(s, x_0(s), \dot{x}_0(s)) ds + c^* \quad (t \in T),$$

esto es, si y solo si x_0 es un extremal de la integral I.

Antes de hacer la demostración del Teorema 3.40, el siguiente lema nos será de utilidad.

3.41 Lema: *Dado un arco admisible x, existe un único arco $z \in Y$ tal que*

$$I'(x, y) = (y, z) \tag{3.8}$$

para todo arco $y \in Y$, donde

$$(y, z) := \int_{t_0}^{t_1} \langle \dot{y}(t), \dot{z}(t) \rangle dt.$$

El arco z está definido por

$$z(t) := \int_{t_0}^{t} \left\{ L_{\dot{x}}^*(s, x(s), \dot{x}(s)) - \int_{t_0}^{s} L_x^*(\tau, x(\tau), \dot{x}(\tau)) d\tau - c \right\} ds \quad (t \in T)$$

donde

$$c := \frac{1}{t_1 - t_0}\int_{t_0}^{t_1}\left\{L_{\dot{x}}^*(t, x(t), \dot{x}(t)) - \int_{t_0}^{t} L_x^*(s, x(s), \dot{x}(s))ds\right\}dt.$$

Demostración: Primero que todo nótese que z es C^1 a pedazos en T, $z(t_0) = z(t_1) = 0$, y por lo tanto $z \in Y$. Para demostrar (3.8), sea $y \in Y$. Entonces $y(t_0) = y(t_1) = 0$, y

$$\begin{aligned}
0 &= \int_{t_0}^{t_1} \frac{d}{dt}\left\langle y(t), \int_{t_0}^{t} L_x^*(s, x(s), \dot{x}(s))ds + c\right\rangle dt \\
&= \int_{t_0}^{t_1}\left[\left\langle \int_{t_0}^{t} L_x^*(s, x(s), \dot{x}(s))ds + c, \dot{y}(t)\right\rangle + L_x(t, x(t), \dot{x}(t))y(t)\right]dt \\
&= \int_{t_0}^{t_1}\{L_x(t, x(t), \dot{x}(t))y(t) + (L_{\dot{x}}(t, x(t), \dot{x}(t)) - \dot{z}^*(t))\dot{y}(t)\}dt \\
&= \int_{t_0}^{t_1}\{L_x(t, x(t), \dot{x}(t))y(t) + L_{\dot{x}}(t, x(t), \dot{x}(t))\dot{y}(t)\}dt - \int_{t_0}^{t_1}\langle \dot{y}(t), \dot{z}(t)\rangle dt \\
&= I'(x, y) - (y, z).
\end{aligned}$$

En consecuencia, (3.8) se cumple. Si existieran dos funciones z_1 y z_2 en Y tales que (3.8) se cumpla, entonces

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

para toda $y \in Y$. Tomando $y = z_1 - z_2$, vemos que

$$\int_{t_0}^{t_1} |\dot{y}(t)|^2 dt = 0$$

y en consecuencia, $\dot{y}(t) = 0$ en $t_0 \le t \le t_1$. Puesto que $y(t_0) = 0$, se sigue que $y = 0$ y que $z_1 = z_2$. Esto demuestra el lema. ▮

Para demostrar la última afirmación del Teorema 3.40, elijamos $z \in Y$ relacionada a x_0 como se describe en el Lema 3.41. Entonces,

$$I'(x_0, z) = (z, z) = \int_{t_0}^{t_1} |\dot{z}(t)|^2 dt.$$

La relación

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

por lo tanto se cumple en Y si y solo si $\dot{z}(t) = 0$ en $t_0 \le t \le t_1$, esto es, si y solo si la relación

$$L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^{t} L_x(s, x_0(s), \dot{x}_0(s))ds + c^* \quad (t \in T)$$

es satisfecha. Esto completa la demostración del teorema. ▮

10. Una condición necesaria adicional

Como seguramente hemos observado uno de los enfoques fundamentales que se utilizan para atacar el problema (P) es asumir que uno tiene algún tipo de solución de (P) y bajo un estudio detallado detectar qué condiciones debe de satisfacer dicha solución. Por supuesto, es muy importante que las condiciones que debe de satisfacer la solución no tengan el mismo grado de dificultad que el de resolver el problema directamente. Con frecuencia, la siguiente condición necesaria para un mínimo débil del problema (P), se puede verificar sin mucha dificultad.

3.42 Teorema: *Si x_0 es un mínimo débil de (P) y L es C^1 en $T \times \mathbf{R}^n \times \mathbf{R}^n$, entonces para alguna $c \in \mathbf{R}$,*

$$L(t, x_0(t), \dot{x}_0(t)) - L_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\dot{x}_0(t) = \int_{t_0}^{t} L_t(s, x_0(s), \dot{x}_0(s))ds + c \quad (t \in T).$$

Demostración: Para demostrar este teorema escribiremos a x_0 en forma paramétrica

$$x_0\colon \qquad x_0^0(t) := t, \quad x_0^i(t) \qquad (t \in T, i = 1, \ldots, n).$$

Primero que todo, consideremos la clase de arcos paramétricos

$$x\colon \qquad x^0(t), \quad x^i(t) \qquad (t \in T, i = 1, \ldots, n) \tag{3.9}$$

que satisfacen $\dot{x}^0(t) > 0$ $(t \in T)$ y que unen los puntos finales de x_0, esto es,

$$x^k(t_0) = x_0^k(t_0), \quad x^k(t_1) = x_0^k(t_1) \qquad (k = 0, 1, \ldots, n). \tag{3.10}$$

Ahora, definamos F por

$$F(x, \dot{x}) = F(x^0, x^1, \ldots, x^n, \dot{x}^0, \dot{x}^1, \ldots, \dot{x}^n) := L(x^0, x^1, \ldots, x^n, \dot{x}^1/\dot{x}^0, \ldots, \dot{x}^n/\dot{x}^0)\dot{x}^0 \quad (\dot{x}^0 > 0).$$

Obsérvese que para cualquier arco paramétrico (3.9) que satisface (3.10), se tiene que

$$\begin{aligned}\int_{t_0}^{t_1} F(x(s), \dot{x}(s))ds &= \int_{t_0}^{t_1} F(x((x^0)^{-1}(t)), \dot{x}((x^0)^{-1}(t)))\frac{dt}{\dot{x}^0((x^0)^{-1}(t))}\\ &= \int_{t_0}^{t_1} L(t, x^1 \circ (x^0)^{-1}(t), \ldots, x^n \circ (x^0)^{-1}(t), (x^1 \circ (x^0)^{-1})'(t), \ldots, (x^n \circ (x^0)^{-1})'(t))dt\\ &= I(x \circ (x^0)^{-1}).\end{aligned}$$

Puesto que x_0 es un mínimo débil de (P), existe $\epsilon > 0$ tal que $I(x) \geq I(x_0)$ siempre que $x = (x^1, \ldots, x^n)$ sea admisible y $\|x - x_0\|_1 < \epsilon$. Se tiene que existe $\delta > 0$ suficientemente pequeña tal que si

$$\|(x^0, x^1, \ldots, x^n) - (x_0^0, x_0^1, \ldots, x_0^n)\|_1 < \delta,$$

entonces

$$\epsilon > \|x \circ (x^0)^{-1} - x_0\|_1 = \|(x^1, \ldots, x^n) \circ (x^0)^{-1} - (x_0^1, \ldots, x_0^n)\|_1.$$

Puesto que $x \circ (x^0)^{-1} = (x^1, \ldots, x^n) \circ (x^0)^{-1}$ es admisible y $\|x \circ (x^0)^{-1} - x_0\|_1 < \epsilon$, se tiene que

$$I(x \circ (x^0)^{-1}) \geq I(x_0).$$

En consecuencia, si $x = (x^0, x^1, \ldots, x^n)\colon T \to \mathbf{R}^{1+n}$ es un arco paramétrico C^1 a pedazos que une los puntos finales de x_0 y que satisface que $\|(x^0, x^1, \ldots, x^n) - (x_0^0, x_0^1, \ldots, x_0^n)\|_1 < \delta$, sucede que

$$\begin{aligned}\int_{t_0}^{t_1} F(x(t), \dot{x}(t))dt &\geq I(x_0)\\ &= \int_{t_0}^{t_1} L(t, x_0(t), \dot{x}_0(t))dt\\ &= \int_{t_0}^{t_1} L(x_0^0(t), x_0^1(t), \ldots, x_0^n(t), \dot{x}_0^1(t)/\dot{x}_0^0(t), \ldots, \dot{x}_0^n(t)/\dot{x}_0^0(t))\dot{x}_0^0(t)dt\\ &= \int_{t_0}^{t_1} F(x_0^0(t), x_0^1(t), \ldots, x_0^n(t), \dot{x}_0^0(t), \dot{x}_0^1(t), \ldots, \dot{x}_0^n(t))dt\\ &= \int_{t_0}^{t_1} F(x_0(t), \dot{x}_0(t))dt,\end{aligned}$$

esto es, x_0 es un mínimo débil del problema de minimizar la integral $\int_{t_0}^{t_1} F(x(t), \dot{x}(t))dt$ sobre el conjunto de arcos paramétricos $x = (x^0, x^1, \ldots, x^n) \colon T \to \mathbf{R}^{1+n}$ que unen los puntos finales de x_0 y que son C^1 a pedazos. De esta manera, por la forma integral de la ecuación de Euler, tenemos que

$$F_{\dot{x}^0}(x_0(t), \dot{x}_0(t)) = \int_{t_0}^{t} F_{x^0}(x_0(s), \dot{x}_0(s))ds + c \quad (t \in T).$$

Notemos que

$$F_{\dot{x}^0}(x, \dot{x}) = L(x^0, x^1, \ldots, x^n, \dot{x}^1/\dot{x}^0, \ldots, \dot{x}^n/\dot{x}^0) - L_{\dot{x}}(x^0, x^1, \ldots, x^n, \dot{x}^1/\dot{x}^0, \ldots, \dot{x}^n/\dot{x}^0)\frac{\dot{x}}{\dot{x}^0},$$

$$F_{x^0}(x, \dot{x}) = L_t(x^0, x^1, \ldots, x^n, \dot{x}^1/\dot{x}^0, \ldots, \dot{x}^n/\dot{x}^0)\dot{x}^0.$$

Así, para toda $t \in T$,

$$\begin{aligned} &F_{\dot{x}^0}(x_0(t), \dot{x}_0(t)) = L(t, x_0(t), \dot{x}_0(t)) - L_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\dot{x}_0(t), \\ &F_{x^0}(x_0(t), \dot{x}_0(t)) = L_t(t, x_0(t), \dot{x}_0(t)). \end{aligned}$$

Entonces,

$$L(t, x_0(t), \dot{x}_0(t)) - L_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\dot{x}_0(t) = \int_{t_0}^{t} L_t(s, x_0(s), \dot{x}_0(s))ds + c \quad (t \in T). \blacksquare$$

3.43 Observación: *Obsérvese que en un arco de clase C^2, para toda $t \in T$, se tiene que*

$$\begin{aligned} &\frac{d}{dt}[L(t, x(t), \dot{x}(t)) - L_{\dot{x}}(t, x(t), \dot{x}(t))\dot{x}(t)] - L_t(t, x(t), \dot{x}(t)) \\ &+ \left(\frac{d}{dt}L_{\dot{x}}(t, x(t), \dot{x}(t)) - L_x(t, x(t), \dot{x}(t))\right)\dot{x}(t) = 0. \end{aligned}$$

3.44 Observación: *Se sigue que en arcos de clase C^2, la ecuación*

$$\frac{d}{dt}[L(t, x(t), \dot{x}(t)) - L_{\dot{x}}(t, x(t), \dot{x}(t))\dot{x}(t)] = L_t(t, x(t), \dot{x}(t)) \quad (t \in T)$$

es una consecuencia de la ecuación de Euler.

3.45 Observación: *Nótese que si la función*

$$t \mapsto L(t, x_0(t), \dot{x}_0(t)) - L_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\dot{x}_0(t)$$

es continua en T, entonces

$$\frac{d}{dt}[L(t, x_0(t), \dot{x}_0(t)) - L_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\dot{x}_0(t)] = L_t(t, x_0(t), \dot{x}_0(t)) \quad (t \in T)$$

si y solo si para alguna constante $c \in \mathbf{R}$

$$L(t, x_0(t), \dot{x}_0(t)) - L_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\dot{x}_0(t) = \int_{t_0}^{t} L_t(s, x_0(s), \dot{x}_0(s))ds + c \quad (t \in T).$$

En particular, por el Corolario 3.25, si x_0 satisface la forma integral de la ecuación de Euler y la condición necesaria de Weierstrass para un mínimo débil, entonces con respecto a x_0 las dos condiciones mencionadas anteriormente son equivalentes.

11. Problemas isoperimétricos

En el problema (P) suele suceder que no se necesita buscar una solución en un conjunto tan amplio como aquel que está delimitado por las condiciones (a) y (b). Ciertos tipos de restricciones adicionales con frecuencia nos ayudan a delimitar aún más el conjunto sobre el cual uno quiere minimizar a la funcional I. Las restricciones más comunes que se encuentran en las aplicaciones del cálculo de variaciones son aquellas que involucran integrales como la funcional I. A este tipo de restricciones se les conoce como *restricciones isoperimétricas.* Vale la pena señalar que el problema de encontrar un arco de longitud fija que al hacerlo girar como una superficie de revolución alrededor del eje de las abscisas y que encierra la mayor área posible cae dentro de un problema de este tipo.

Consideremos el problema (P) de minimizar la funcional

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt$$

sujeta a

(a) $x: T \to \mathbf{R}^n$ es C^1 a pedazos.

(b) $x(t_0) = \xi_0$ y $x(t_1) = \xi_1$.

(c) $I_i(x) := \int_{t_0}^{t_1} L_i(t, x(t), \dot{x}(t))dt = 0$ $(i = 1, \ldots, p)$.

En esta sección asumiremos que las funciones L, L_i $(i = 1, \ldots, p)$ son de clase C^2 en $T \times \mathbf{R}^n \times \mathbf{R}^n$, y como es natural, un arco x es *admisible* si satisface las restricciones (a)-(c).

3.46 Definición: *Un arco admisible x_0 es normal si éste no es un extremal de una integral de la forma*

$$J(x) := \sum_{i=1}^{p} \lambda_i I_i(x)$$

donde $\lambda_1, \ldots, \lambda_p$ son constantes, no todas cero. Recordemos que x_0 es un extremal de J si y solo si

$$J'(x_0, y) = \sum_{i=1}^{p} \lambda_i I_i'(x_0, y) = 0$$

para todas las $y \in Y = \{y \in X \mid y(t_0) = y(t_1) = 0\}$. En consecuencia, x_0 es normal si y solo si las variaciones $I_i'(x_0, \cdot)$ $(i = 1, \ldots, p)$ de I_i a lo largo de x_0 son linealmente independientes en Y.

El siguiente lema será de utilidad para la demostración del Lema 3.49. Su afirmación así como su demostración se pueden encontrar en [7, p. 12].

3.47 Lema: *Un conjunto de p funcionales lineales $L_1, \ldots, L_p$ es linealmente independiente en un espacio vectorial E si y solo si existen p vectores $x_1, \ldots, x_p$ en E, tales que el determinante de la matriz*

$$\begin{pmatrix} L_1(x_1) & \ldots & L_1(x_p) \\ \vdots & \ddots & \vdots \\ L_p(x_1) & \ldots & L_p(x_p) \end{pmatrix}$$

es diferente de cero.

El Lema 3.48 es una consecuencia inmediata de la Definición 3.46 y el Lema 3.47.

3.48 Lema: *Sea $I_i'(x_0, \cdot)$ la primera variación de I_i a lo largo de x_0. El arco x_0 es normal si y solo si las funcionales lineales $I_i'(x_0, \cdot)$ $(i = 1, \ldots, p)$ son linealmente independientes en Y. El arco x_0 es normal si y solo si existen p arcos $y_1, \ldots, y_p$ en Y tales que*

$$\det \begin{pmatrix} I_1'(x_0, y_1) & \ldots & I_1'(x_0, y_p) \\ \vdots & \ddots & \vdots \\ I_p'(x_0, y_1) & \ldots & I_p'(x_0, y_p) \end{pmatrix} \neq 0. \tag{3.11}$$

Si x_0 es un arco normal admisible, entonces hay arcos vecinos admisibles, tal como se puede ver en el siguiente lema.

3.49 Lema: *Sea x_0 un arco normal admisible, y sea $y \in Y$ que satisfaga las condiciones*

$$I_i'(x_0, y) = 0 \quad (i = 1, \ldots, p). \tag{3.12}$$

Existe una familia uniparamétrica

$$x(\epsilon): \qquad x(t, \epsilon) \quad (t_0 \leq t \leq t_1, |\epsilon| < \epsilon_0)$$

de arcos admisibles que contienen a x_0 para $\epsilon = 0$. Las funciones $x(t, \epsilon)$ tienen primera y segunda derivadas continuas $x_\epsilon(t, \epsilon)$, $x_{\epsilon\epsilon}(t, \epsilon)$ las cuales definen arcos en Y. Finalmente,

$$y(t) = x_\epsilon(t, 0) \quad (t_0 \leq t \leq t_1),$$

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

Demostración: Seleccionemos $y_1, \ldots, y_p$ arcos en Y tales que (3.11) se satisfaga. Consideremos las ecuaciones

$$I_i\left(x_0 + \sum_{j=1}^{p} c_j y_j + \epsilon y\right) = 0 \quad (i = 1, \ldots, p). \tag{3.13}$$

Definamos $f_i: \mathbf{R} \times \mathbf{R}^p \to \mathbf{R}$ $(i = 1, \ldots, p)$ por

$$f_i(\epsilon, c) := I_i\left(x_0 + \sum_{j=1}^{p} c_j y_j + \epsilon y\right).$$

Obsérvese que todas las derivadas parciales de primer y segundo orden de f_i existen y son continuas en $\mathbf{R} \times \mathbf{R}^p$, es decir, para toda $i = 1, \ldots, p$, f_i es de clase C^2 en $\mathbf{R} \times \mathbf{R}^p$. Además, nótese que para toda $i = 1, \ldots, p$,

$$f_i(0, 0) = 0,$$

y en virtud de (3.13), si $f = (f_1, \ldots, f_p)$, entonces

$$|f_c(0, 0)| \neq 0.$$

Por el Teorema 3.30, existen $\epsilon_0 > 0$ y $c: B(0, \epsilon_0) \to \mathbf{R}^p$ de clase C^2 tal que $c(0) = 0$. Sustituyendo c por $c(\epsilon)$ en (3.13), encontramos que al derivar con respecto a ϵ y evaluando en $\epsilon = 0$,

$$0 = \sum_{j=1}^{p} I_i'(x_0, y_j) c_j'(0) + I_i'(x_0, y) = \sum_{j=1}^{p} I_i'(x_0, y_j) c_j'(0) \quad (i = 1, \ldots, p).$$

Por (3.11), tenemos que $c'(0) = 0$. Como se puede verificar fácilmente, el arco

$$x(\epsilon) := x_0 + \sum_{j=1}^{p} c_j(\epsilon) y_j + \epsilon y$$

tiene las propiedades descritas en el lema. ∎

El siguiente lema será de utilidad en la demostración del Teorema 3.51. Su enunciado así como su demostración se pueden encontrar en [7, p. 12-13].

3.50 Lema: *Sean $L, L_1, \ldots, L_p$ funcionales lineales en E. Si $L(x) = 0$ para toda x en E que satisface las relaciones*

$$L_i(x) = 0 \quad (i = 1, \ldots, p),$$

existe un conjunto de multiplicadores $\lambda_1, \ldots, \lambda_p$, *tales que*

$$L(x) = \sum_{i=1}^{p} \lambda_i L_i(x)$$

en E. *Si* $L_1, \ldots, L_p$ *son linealmente independientes, estos multiplicadores son únicos.*

3.51 Teorema: *Supongamos que* x_0 *es normal y que* x_0 *es un mínimo débil de (P). Entonces, existe un único conjunto de multiplicadores* $\lambda_1, \ldots, \lambda_p$ *tales que la primera variación* J' *de la integral*

$$J := I + \lambda_1 I_1 + \cdots + \lambda_p I_p \tag{3.14}$$

satisface la condición

$$J'(x_0, y) = 0 \tag{3.15}$$

para toda $y \in Y$, *y la segunda variación* J'' *satisface la condición*

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

para todos los arcos $y \in Y$ *que satisfacen las condiciones*

$$I_i'(x_0, y) = 0 \quad (i = 1, \ldots, p). \tag{3.16}$$

Demostración: Sea $y \in Y$ que satisfaga (3.16), y sea $x(\epsilon)$ relacionada a x_0 y a y como en el Lema 3.49. Entonces,

$$W(\epsilon) := I(x(\epsilon))$$

tiene un mínimo local en $\epsilon = 0$. En consecuencia,

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

Por el Lema 3.50, existen multiplicadores únicos $\lambda_1, \ldots, \lambda_p$ tales que

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

en Y. De esta manera, si definimos J por (3.14), se tiene que

$$J'(x_0, y) = I'(x_0, y) + \sum_{i=1}^{p} \lambda_i I_i'(x_0, y) = 0$$

en Y. Puesto que $I_i(x(\epsilon)) = 0$ $(i = 1, \ldots, p; \epsilon \in \mathbf{R})$, podemos escribir

$$W(\epsilon) = J(x(\epsilon)) = I(x(\epsilon)) + \sum_{i=1}^{p} \lambda_i I_i(x(\epsilon)).$$

Puesto que $\epsilon = 0$ minimiza $W(\epsilon)$ localmente, tenemos que

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

y

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

puesto que $x_{\epsilon\epsilon} \in Y$. ▮

Sabemos que la anulación de la primera variación sobre el conjunto de variaciones admisibles es equivalente a la forma integral de la ecuación de Euler. Con esto en mente, podemos ver la veracidad de la primera conclusión del siguiente teorema.

3.52 Teorema: *Si x_0 es un mínimo débil de (P), y x_0 es normal, entonces existe un único conjunto de multiplicadores $\lambda_1, \ldots, \lambda_p$ tales que la relación*

$$F_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^{t} F_x(s, x_0(s), \dot{x}_0(s))ds + c^* \quad (t \in T) \tag{3.17}$$

se cumple para alguna constante $c \in \mathbf{R}^n$, donde

$$F := L + \sum_{i=1}^{p} \lambda_i L_i.$$

Además,

$$E(t, x_0(t), \dot{x}_0(t), u) \geq 0 \quad \textit{para toda } (t, u) \in T \times \mathbf{R}^n \tag{3.18}$$

donde E es la función exceso de Weierstrass

$$E(t, x, \dot{x}, u) := F(t, x, u) - F(t, x, \dot{x}) - F_{\dot{x}}(t, x, \dot{x})(u - \dot{x}).$$

Demostración: Como vimos en las demostraciones de las condiciones necesarias de Weierstrass para el problema (P) sin restricciones isoperimétricas, por continuidad, es suficiente considerar esta condición en un punto arbitrario $\bar{t}$ el cual no corresponde a un punto esquina o a un punto inicial o final de x_0. Sea $\bar{x} := x_0(\bar{t})$, $\bar{u} := \dot{x}_0(\bar{t})$, y seleccionemos $u \in \mathbf{R}^n$. Sea $y(\epsilon)$ $(0 \leq \epsilon \leq \epsilon_0)$ el arco en Y definido por las relaciones

$$y(t, \epsilon) := \begin{cases} 0 & (t_0 \leq t \leq \bar{t}) \\ (u - \bar{u})(t - \bar{t}) & (\bar{t} \leq t \leq \bar{t} + \epsilon) \\ \frac{\epsilon(u - \bar{u})(t_1 - t)}{t_1 - \bar{t} - \epsilon} & (\bar{t} + \epsilon \leq t \leq t_1). \end{cases}$$

Sean $y_1, \ldots, y_p$ arcos en Y tales que (3.11) se cumple. Como en la demostración del Lema 3.49, las ecuaciones

$$I_i\left(x_0 + \sum_{j=1}^{p} c_j y_j + y(\epsilon)\right) = 0 \quad (i = 1, \ldots, p)$$

tienen una solución $c(\epsilon)$ de clase C^1 en un intervalo $0 \leq \epsilon \leq \epsilon_1$ tal que $c(0) = 0$, y tal que el arco

$$x(\epsilon) := x_0 + \sum_{j=1}^{p} c_j(\epsilon) y_j + y(\epsilon)$$

es admisible. Puesto que x_0 es un mínimo débil de (P), existe $\gamma > 0$ tal que si x es admisible,

$$[\|x - x_0\|_0 < \gamma \quad \text{y} \quad \|\dot{x} - \dot{x}_0\|_0 < \gamma] \Longrightarrow I(x_0) \leq I(x),$$

de hecho, por la admisibilidad de x y x_0, también se tiene que

$$[\|x - x_0\|_0 < \gamma \quad \text{y} \quad \|\dot{x} - \dot{x}_0\|_0 < \gamma] \Longrightarrow J(x_0) \leq J(x).$$

Uno puede verificar fácilmente que

$$\|x(\epsilon) - x_0\|_0 < \gamma, \quad \|\dot{x}(\epsilon) - \dot{x}_0\|_0 < \gamma,$$

siempre que $\epsilon_1 > 0$ sea suficientemente pequeña. En consecuencia, la función

$$W(\epsilon) := J(x(\epsilon))$$

tiene un mínimo local en $\epsilon = 0$. Puesto que $\epsilon \geq 0$, se tiene que $W'(0) \geq 0$. Para calcular $W'(0)$, obsérvese primero que, por (3.15),

$$W'(0) = \sum_{j=1}^{p} J'(x_0, y_j)c_j'(0) + G'(0) = G'(0),$$

donde

$$\begin{aligned} G(\epsilon) &:= J(x_0 + y(\epsilon)) \\ &= \text{constante} + \int_{\bar{t}}^{\bar{t}+\epsilon} F(t, x_0(t) + (u - \bar{u})(t - \bar{t}), \dot{x}_0(t) + u - \bar{u})dt \\ &\qquad + \int_{\bar{t}+\epsilon}^{t_1} F(t, x_0(t) + \phi(\epsilon)z(t), \dot{x}_0(t) + \phi(\epsilon)\dot{z}(t))dt, \end{aligned}$$

donde

$$\phi(\epsilon) := \frac{\epsilon(t_1 - \bar{t})}{t_1 - \bar{t} - \epsilon} \quad \text{y} \quad z(t) := \frac{(u - \bar{u})(t_1 - t)}{t_1 - \bar{t}}.$$

Observando que $\phi(0) = 0$, $\phi'(0) = 1$, y por (3.17),

$$\begin{aligned} G'(0) &= F(\bar{t}, \bar{x}, u) - F(\bar{t}, \bar{x}, \bar{u}) + \int_{\bar{t}}^{t_1} \{F_x(t, x_0(t), \dot{x}_0(t))z(t) + F_{\dot{x}}(t, x_0(t), \dot{x}_0(t))\dot{z}(t)\}dt \\ &= F(\bar{t}, \bar{x}, u) - F(\bar{t}, \bar{x}, \bar{u}) + \int_{\bar{t}}^{t_1} \frac{d}{dt}[F_{\dot{x}}(t, x_0(t), \dot{x}_0(t))z(t)]dt. \end{aligned}$$

Puesto que $z(\bar{t}) = u - \bar{u}$ y $z(t_1) = 0$, se sigue que

$$G'(0) = F(\bar{t}, \bar{x}, u) - F(\bar{t}, \bar{x}, \bar{u}) - F_{\dot{x}}(\bar{t}, \bar{x}, \bar{u})(u - \bar{u}) = E(\bar{t}, \bar{x}, \bar{u}, u).$$

Como $W'(0) = G'(0) \geq 0$ la desigualdad (3.18) es satisfecha. ∎

3.53 Observación: *Analizando el Teorema 3.22, obsérvese que la desigualdad (3.18) no es necesariamente satisfecha si x_0 es un mínimo débil del problema (P) que no tiene restricciones isoperimétricas. Por otro lado, si x_0 es normal y x_0 es un mínimo débil del problema isoperimétrico (P), entonces, en efecto, x_0 sí tiene que satisfacer la desigualdad (3.18) aunque éste sea un mínimo débil. Por lo tanto, para el problema (P) de esta sección si x_0 es normal únicamente existe una condición necesaria de Weierstrass, en contraste, con el problema sin restricciones para el cual las condiciones necesarias de Weierstrass, en general, son diferentes cuando uno estudia mínimos débiles o fuertes.*

12. Extremos y variables canónicas

En esta sección demostraremos que si x_0 es un extremo no singular y L es de clase C^2 en $T \times \mathbf{R}^n \times \mathbf{R}^n$, entonces la ecuación de Euler es equivalente a una ecuación diferencial de primer grado en el espacio $2n$-dimensional. Este resultado combinado con la teoría de las ecuaciones diferenciales nos permitirá desarrollar la teoría de Jacobi, la cual es una de las componentes más importantes de las condiciones necesarias de segundo orden y de las condiciones suficientes clásicas del problema (P) sin restricciones.

3.54 Teorema (Teorema de la función implícita): *Sean S un subconjunto abierto de $\mathbf{R}^{m+n}$ y K_0 un subconjunto compacto de $\mathbf{R}^m$. Sea $f(t,x)\colon S \to \mathbf{R}^n$ continua en S y supongamos que $f_x(t,x)\colon S \to \mathbf{R}^{n\times n}$ es continua en S. Asumamos que existe una función continua $x_0\colon K_0 \to \mathbf{R}^n$ tal que*

$$t \in K_0 \Longrightarrow (t, x_0(t)) \in S, \quad f(t, x_0(t)) = 0, \quad |f_x(t, x_0(t))| \neq 0.$$

Entonces, existen una vecindad V de K_0, una función continua $x\colon V \to \mathbf{R}^n$ y una constante $\epsilon > 0$ tales que

$$x(t) = x_0(t) \ (t \in K_0) \quad y \quad f(t, x(t)) = 0 \ (t \in V),$$

y tales que las relaciones

$$f(t,x)=0,\ |x-x(t)|<\epsilon\ (t\in V)\Longrightarrow x=x(t).$$

Si la función f es C^m en S, la función x es C^m en V.

De ahora en adelante todo arco x_0 (o en ocasiones x) será un *extremo no singular*, esto es, x_0 será un arco sin esquinas que satisface

$$\frac{d}{dt}L_{\dot x}(t,x_0(t),\dot x_0(t))=L_x(t,x_0(t),\dot x_0(t))\quad (t\in T),$$

$$|L_{\dot x\dot x}(t,x_0(t),\dot x_0(t))|\neq 0\quad (t\in T).$$

3.55 Observación: *Nótese que por el Teorema 3.32, x_0 es C^2 en T, y entonces en este caso la ecuación de Euler se convierte en la ecuación diferencial de segundo grado*

$$L^*_{\dot x t}+L_{\dot x x}\dot x_0(t)+L_{\dot x\dot x}\ddot x_0(t)=L^*_x\quad (t\in T)$$

la cual es una herramienta útil en la obtención de candidatos para resolver el problema (P). Aquí, los argumentos en las derivadas de L son $(t,x_0(t),\dot x_0(t))$.

Definamos $p_0\colon T\to\mathbf{R}^n$ por

$$p_0(t):=L^*_{\dot x}(t,x_0(t),\dot x_0(t)).$$

Claramente, p_0 es continua en T. Definamos $F\colon T\times\mathbf{R}^n\times\mathbf{R}^n\times\mathbf{R}^n\to\mathbf{R}^n$ por

$$F(t,x,p,\dot x):=L^*_{\dot x}(t,x,\dot x)-p.$$

También, definamos

$$K_0:=\{(t,x_0(t),p_0(t))\mid t\in T\}.$$

Claramente, K_0 es compacto en $\mathbf{R}\times\mathbf{R}^n\times\mathbf{R}^n$. Sea $g\colon K_0\to\mathbf{R}^n$ definida por

$$g(t,x_0(t),p_0(t)):=\dot x_0(t).$$

No es difícil ver que g es continua en K_0 y es evidente que

$$(t,x_0(t),p_0(t))\in K_0\Longrightarrow (t,x_0(t),p_0(t),g(t,x_0(t),p_0(t)))=(t,x_0(t),p_0(t),\dot x_0(t))\in S$$

donde S lo definimos como el abierto

$$S:=T\times\mathbf{R}^n\times\mathbf{R}^n\times\mathbf{R}^n$$

el cual es el dominio de F. Además, obsérvese que $F_{\dot x}(t,x,p,\dot x)=L_{\dot x\dot x}(t,x,\dot x)$ para toda $(t,x,p,\dot x)\in S$ y por lo tanto $F\colon S\to\mathbf{R}^n$ y $F_{\dot x}\colon S\to\mathbf{R}^{n\times n}$ son continuas en S. Adicionalmente,

$$F(t,x_0(t),p_0(t),\dot x_0(t))=0\quad\text{y}\quad |F_{\dot x}(t,x_0(t),p_0(t),\dot x_0(t))|\neq 0\quad (t\in T).$$

Así, por el Teorema 3.54, existe una función continua P definida en una vecindad $\mathcal{P}$ de K_0 y una constante $\epsilon>0$ tales que

$$P(t,x_0(t),p_0(t))=\dot x_0(t)\ (t\in T)\quad\text{y}\quad F(t,x,p,P(t,x,p))=0\ ((t,x,p)\in\mathcal{P})$$

si y solo si

$$P(t,x_0(t),p_0(t))=\dot x_0(t)\ (t\in T)\quad\text{y}\quad L^*_{\dot x}(t,x,P(t,x,p))=p\ ((t,x,p)\in\mathcal{P}),\tag{3.19}$$

y tales que las relaciones

$$F(t,x,p,P)=0,\ |P-P(t,x,p)|<\epsilon\ ((t,x,p)\in\mathcal{P})\Longrightarrow P=P(t,x,p).$$

Además, si la función L es de clase C^m en $T \times \mathbf{R}^n \times \mathbf{R}^n$, la función P es de clase C^{m-1} en $\mathcal{P}$.

La función $H\colon \mathcal{P} \to \mathbf{R}$ definida por

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

es llamada el *Hamiltoniano* correspondiente a la integral I. Obsérvese que para toda $(t,x,p) \in \mathcal{P}$,

$$H_t(t,x,p) = -L_t(t,x,P(t,x,p)), \quad H_x(t,x,p) = -L_x(t,x,P(t,x,p)), \quad H_p(t,x,p) = P^*(t,x,p). \tag{3.20}$$

En consecuecia, H es de clase C^m en $\mathcal{P}$ siempre que L sea de clase C^m en $T \times \mathbf{R}^n \times \mathbf{R}^n$.

Obsérvese además que la ecuación de Euler

$$\frac{d}{dt}L_{\dot{x}}(t,x(t),\dot{x}(t)) = L_x(t,x(t),\dot{x}(t)) \quad (t \in T), \tag{3.21}$$

es equivalente a las ecuaciones

$$\dot{x}(t) = H_p^*(t,x(t),p(t)), \quad \dot{p}(t) = -H_x^*(t,x(t),p(t)) \quad (t \in T). \tag{3.22}$$

En efecto, supongamos que x es un arco no singular sin esquinas que satisface (3.21). Primero que todo, recuérdese que $p(\cdot)$ se define por $L_{\dot{x}}^*(\cdot,x(\cdot),\dot{x}(\cdot))$. Además, por la primera igualdad de (3.19) y la tercera igualdad de (3.20), nótese que

$$\dot{x}(t) = P(t,x(t),p(t)) = H_p^*(t,x(t),p(t)) \quad (t \in T),$$

y por lo tanto la primera ecuación de (3.22) es satisfecha. Además, por la segunda igualdad de (3.20), para toda $t \in T$,

$$\begin{aligned}\dot{p}(t) &= \frac{d}{dt}L_{\dot{x}}^*(t,x(t),\dot{x}(t)) \\ &= L_x^*(t,x(t),\dot{x}(t)) \\ &= L_x^*(t,x(t),P(t,x(t),p(t))) \\ &= -H_x^*(t,x(t),p(t))\end{aligned}$$

y así, la segunda ecuación de (3.22) también se verifica.

Recíprocamente, sea (x,p) que satisfaga (3.22) con x un arco sin esquinas no singular. Por la tercera igualdad de (3.20) y la primera igualdad de (3.22), se tiene que

$$\dot{x}(t) = P(t,x(t),p(t)) \quad (t \in T). \tag{3.23}$$

Ahora, por la segunda igualdad de (3.20) y (3.23),

$$\dot{p}(t) = -H_x^*(t,x(t),p(t)) = L_x^*(t,x(t),P(t,x(t),p(t))) = L_x^*(t,x(t),\dot{x}(t)) \quad (t \in T). \tag{3.24}$$

Entonces, por la segunda igualdad de (3.19), (3.23) y (3.24), para toda $t \in T$,

$$\begin{aligned}\frac{d}{dt}L_{\dot{x}}(t,x(t),\dot{x}(t)) &= \frac{d}{dt}L_{\dot{x}}(t,x(t),P(t,x(t),p(t))) \\ &= \frac{d}{dt}p^*(t) \\ &= \dot{p}^*(t) \\ &= L_x(t,x(t),\dot{x}(t)),\end{aligned}$$

y por lo tanto la ecuación de Euler es satisfecha. Al par de ecuaciones (3.22) se le llama la forma *canónica o Hamiltoniana* de la ecuación de Euler.

Finalmente, obsérvese que la ecuación

$$\frac{d}{dt}[L(t, x(t), \dot{x}(t)) - L_{\dot{x}}(t, x(t), \dot{x}(t))\dot{x}(t)] = L_t(t, x(t), \dot{x}(t)) \quad (t \in T)$$

la cual se cumple a lo largo de extremos no singulares, toma la forma simple

$$\frac{d}{dt}H(t, x(t), p(t)) = H_t(t, x(t), p(t)) \quad (t \in T).$$

13. La condición de Jacobi

En esta sección daremos una introducción de la teoría de Jacobi, la cual consiste en caracterizar a la no negatividad de la segunda variación sobre el conjunto de variaciones admisibles con las condiciones de Legendre y Jacobi. La verificabilidad de estas dos últimas condiciones es una tarea muy esperanzadora, en contraste, con la verificabilidad de la condición que involucra la no negatividad de la segunda variación. Específicamente, en esta sección, introduciremos en la Definición 3.57 la noción fundamental de un *punto conjugado* y demostraremos que si x_0 es un mínimo débil no singular de (P), entonces no existen puntos conjugados a t_0 sobre x_0 en el intervalo abierto (t_0, t_1). La caracterización mencionada arriba se demostrará en la Sección 16.

Como vimos anteriormente la segunda variación $I''(x_0, \cdot)$ de I a lo largo de un arco admisible x_0 toma la forma

$$I''(x_0, y) = \int_{t_0}^{t_1} 2\omega(t, y(t), \dot{y}(t))dt,$$

donde

$$2\omega(t, y, \dot{y}) := \langle y, L_{xx}y\rangle + 2\langle \dot{y}, L_{\dot{x}x}y\rangle + \langle \dot{y}, L_{\dot{x}\dot{x}}\dot{y}\rangle$$

y los argumentos en las derivadas de L son $(t, x_0(t), \dot{x}_0(t))$. Además, se demostró que si x_0 es un mínimo débil de (P), entonces

$$I''(x_0, y) \geq 0 \tag{3.25}$$

en el espacio vectorial de *variaciones admisibles* Y. La condición (3.25) es una condición de segundo orden y es llamada la *condición necesaria de segundo orden en términos de variaciones para un mínimo.* La desigualdad (3.25) en Y es equivalente a la afirmación de que $y = 0$ es un mínimo global de $I''(x_0, \cdot)$ en Y. El problema de minimizar $I''(x_0, \cdot)$ en Y que se denotará por (PA) será llamado el *problema accesorio.*

De ahora en adelante, supondremos que cualquier arco x_0 es un mínimo débil de (P) y además éste no tiene esquinas y es no singular. Será conveniente designar a la segunda variación $I''(x_0, \cdot)$ por J. La integral

$$J(y) = \int_{t_0}^{t_1} 2\omega(t, y(t), \dot{y}(t))dt$$

es una forma cuadrática en y, en el sentido de que

$$J(y) = \int_{t_0}^{t_1} \left\langle \begin{pmatrix} y(t) \\ \dot{y}(t) \end{pmatrix}, \begin{pmatrix} L_{xx} & L_{x\dot{x}} \\ L_{\dot{x}x} & L_{\dot{x}\dot{x}} \end{pmatrix} \begin{pmatrix} y(t) \\ \dot{y}(t) \end{pmatrix} \right\rangle dt$$

donde nuevamente los argumentos de las derivadas de L son $(t, x_0(t), \dot{x}_0(t))$. La forma bilineal correspondiente es

$$\begin{aligned} J(z, y) &:= \int_{t_0}^{t_1} \left\langle \begin{pmatrix} z(t) \\ \dot{z}(t) \end{pmatrix}, \begin{pmatrix} L_{xx} & L_{x\dot{x}} \\ L_{\dot{x}x} & L_{\dot{x}\dot{x}} \end{pmatrix} \begin{pmatrix} y(t) \\ \dot{y}(t) \end{pmatrix} \right\rangle dt \\ &= \int_{t_0}^{t_1} \{\langle z(t), L_{xx}y(t) + L_{x\dot{x}}\dot{y}(t)\rangle + \langle \dot{z}(t), L_{\dot{x}x}y(t) + L_{\dot{x}\dot{x}}\dot{y}(t)\rangle\}dt \\ &= \int_{t_0}^{t_1} \{\omega_y(t, y(t), \dot{y}(t))z(t) + \omega_{\dot{y}}(t, y(t), \dot{y}(t))\dot{z}(t)\}dt. \end{aligned}$$

Es claro que

$$J(y,z) = J(z,y), \quad J(y,y) = J(y).$$

Obsérvese que para toda $(t,y,\dot{y}) \in T \times \mathbf{R}^n \times \mathbf{R}^n$,

$$\omega_{\dot{y}\dot{y}}(t,y,\dot{y}) = L_{\dot{x}\dot{x}}(t,x_0(t),\dot{x}_0(t)) > 0,$$

y en consecuencia la integral J es positiva regular y por lo tanto las soluciones de la forma integral de la ecuación de Euler para J

$$\omega_{\dot{y}}(t,y(t),\dot{y}(t)) = \int_{t_0}^{t} \omega_y(\tau,y(\tau),\dot{y}(\tau))d\tau + c^* \quad (t \in T)$$

no tienen esquinas. De esta manera, la ecuación anterior es equivalente a la ecuación

$$\frac{d}{dt}\omega_{\dot{y}}(t,y(t),\dot{y}(t)) = \omega_y(t,y(t),\dot{y}(t)) \quad (t \in T). \tag{3.26}$$

Esta ecuación es llamada la *ecuación de Jacobi*, y a sus soluciones se les llamará *extremos accesorio.*

Ahora, dado un extremo accesorio y, definamos $q\colon T \to \mathbf{R}^n$ por

$$q(t) := \omega^*_{\dot{y}}(t,y(t),\dot{y}(t)) = L_{\dot{x}x}y(t) + L_{\dot{x}\dot{x}}\dot{y}(t),$$

y como siempre los argumentos en las derivadas de L son $(t,x_0(t),\dot{x}_0(t))$. Puesto que

$$|L_{\dot{x}\dot{x}}(t,x_0(t),\dot{x}_0(t))| \neq 0 \quad (t \in T),$$

entonces de la ecuación anterior $\dot{y}(t)$ se puede despejar en términos de $y(t)$ y $q(t)$. Cuando esto se lleva a cabo se puede ver que la ecuación (3.26) es equivalente al sistema

$$\dot{y}(t) = A(t)y(t) + B(t)q(t), \quad \dot{q}(t) = C(t)y(t) + D(t)q(t) \quad (t \in T), \tag{3.27}$$

donde

$$A(t) := -L^{-1}_{\dot{x}\dot{x}}L_{\dot{x}x}, \quad B(t) := L^{-1}_{\dot{x}\dot{x}},$$
$$C(t) := L_{xx} - L_{x\dot{x}}L^{-1}_{\dot{x}\dot{x}}L_{\dot{x}x}, \quad D(t) := L_{x\dot{x}}L^{-1}_{\dot{x}\dot{x}},$$

y los argumentos en las derivadas de L son $(t,x_0(t),\dot{x}_0(t))$. Nótese que el sistema (3.27) es lineal en y y q y por lo tanto de la teoría de las ecuaciones diferenciales lineales aplicado a (3.27) obtenemos el siguiente resultado.

3.56 Lema: *Un extremo accesorio y está determinado de forma única por los valores de $y(t^0)$, $\dot{y}(t^0)$ o equivalentemente por los valores $y(t^0)$, $q(t^0)$ en un solo punto $t = t^0$. En particular, si $y(t^0) = \dot{y}(t^0) = 0$, entonces $y \equiv 0$ en T.*

3.57 Definición: *Un punto $t = s$ en $(t_0,t_1]$ se dice que es un punto conjugado a $t = t_0$ sobre x_0 si hay un extremo accesorio y tal que $y(t_0) = y(s) = 0$, y $y \not\equiv 0$ en (t_0,s).*

3.58 Observación: *Al trabajar con puntos conjugados es conveniente interpretar a un punto $t = s$ como un punto en $(t_0,t_1]$ o como el punto $(s,x_0(s))$ sobre x_0. De esta manera, si nos referimos a un punto $t = s$ sobre x_0, lo que estamos diciendo es el punto $(s,x_0(s))$.*

3.59 Teorema (Jacobi): *Si un arco no singular sin esquinas x_0 es un mínimo débil de (P), entonces no hay ningún punto $s \in (t_0,t_1)$ conjugado a $t = t_0$ sobre x_0.*

Demostración: Supongamos que hay un punto $t = s$ conjugado a $t = t_0$ sobre x_0 en (t_0,t_1). Sea y un extremo accesorio que satisfaga $y(t_0) = y(s) = 0$, $y \not\equiv 0$ en (t_0,s). Sea z un arco en Y definido por

$$z(t) := y(t) \quad (t_0 \leq t \leq s), \quad z(t) := 0 \quad (s \leq t \leq t_1).$$

Se tiene que

$$\begin{aligned}
J(z) &= \int_{t_0}^{s} 2\omega(\tau, z(\tau), \dot{z}(\tau))d\tau \\
&= \int_{t_0}^{s} \{\omega_y(\tau, y(\tau), \dot{y}(\tau))y(\tau) + \omega_{\dot{y}}(\tau, y(\tau), \dot{y}(\tau))\dot{y}(\tau)\}d\tau \\
&= \int_{t_0}^{s} \left\{ \left(\frac{d}{d\tau}\omega_{\dot{y}}(\tau, y(\tau), \dot{y}(\tau)) \right) y(\tau) + \omega_{\dot{y}}(\tau, y(\tau), \dot{y}(\tau))\dot{y}(\tau) \right\} d\tau \\
&= \int_{t_0}^{s} \frac{d}{d\tau}[\omega_{\dot{y}}(\tau, y(\tau), \dot{y}(\tau))y(\tau)]d\tau = 0.
\end{aligned}$$

En consecuencia, z es un mínimo global de (PA) y por lo tanto z es un extremo accesorio. Puesto que $z \equiv 0$ en $[s, t_1]$, entonces hay un punto $t = t^0$ en (s, t_1) tal que $z(t^0) = \dot{z}(t^0) = 0$. Se sigue del Lema 3.56 que $z \equiv 0$ en T, el cual no es el caso. ▮

3.60 Observación: *Obsérvese que la hipótesis de la no singularidad del arco x_0 en el Teorema 3.59 es fundamental. Para ilustrar este hecho, consideremos el problema (P) de minimizar la integral*

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

sujeta a

(a) $x\colon [0,1] \to \mathbf{R}$ *es* C^1 *a pedazos.*

(b) $x(0) = x(1) = 0$.

Aquí, $T = [0,1]$, $n = 1$, $L(t, x, \dot{x}) = x\dot{x}$, y $\xi_0 = \xi_1 = 0$. Claramente, $I(x) = 0$ para toda x admisible y por lo tanto cualquier x admisible es una solución global del problema (P). Además, puesto que $L_{\dot{x}\dot{x}}(t, x, \dot{x}) = 0$ para toda $(t, x, \dot{x}) \in T \times \mathbf{R} \times \mathbf{R}$, entonces cualquier arco x es 'singular'. Es fácil ver que para toda $(t, y, \dot{y}) \in T \times \mathbf{R} \times \mathbf{R}$, $2\omega(t, y, \dot{y}) = 2y\dot{y}$ y en consecuencia la ecuación de Jacobi toma la forma

$$\frac{d}{dt}y(t) = \dot{y}(t) \quad (t \in T).$$

De esta manera, dada una trayectoria x, cualquier punto $s \in (0,1)$ es un punto conjugado de $t = 0$ sobre x. En otras palabras, en general la no existencia de puntos conjugados sobre x a $t = 0$ en el intervalo $(0,1)$ no es una condición necesaria si x es singular.

3.61 Ejemplo: *Consideremos el problema (P) de minimizar*

$$I(x) = \int_0^b \tfrac{1}{2}\{\dot{x}^2(t) - x^2(t)\}dt \quad (b > \pi)$$

sujeta a

(a) $x\colon [0,b] \to \mathbf{R}$ *es* C^1 *a pedazos.*

(b) $x(0) = x(b) = 0$.

Aquí, $T = [0,b]$, $n = 1$, y $L(t, x, \dot{x}) = \frac{1}{2}\{\dot{x}^2 - x^2\}$. Además, no es difícil ver que para cualquier x que satisface la condición (a), la segunda variación $I''(x; \cdot)$ no depende de x. También, es fácil ver que

$$2\omega(t, y, \dot{y}) = \dot{y}^2 - y^2$$

y por lo tanto

$$J(y) = \int_0^b \{\dot{y}^2(t) - y^2(t)\}dt.$$

Puesto que en este ejemplo, los extremos accesorio son C^2 en T, no es difícil ver que la ecuación de Jacobi está dada por

$$\ddot{y}(t) + y(t) = 0 \quad (t \in T).$$

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

$$y(t) = \beta \operatorname{sen} t \quad (t \in T)$$

donde β es una constante. Los puntos conjugados a $t = 0$ sobre cualquier arco x son los puntos de la forma $t_n = n\pi$ ($n \in \mathbf{N}$). En consecuencia, el punto $s = \pi$ es conjugado a $t = 0$ sobre cualquier arco x. Por lo tanto, por el Teorema 3.59 ningún x admisible puede ser un mínimo débil del problema (P), esto es, el problema (P) no tiene solución. En particular, si x_0 es el arco constante igual a cero, puesto que $I(x_0) = 0$ y el valor anterior no es un valor mínimo, entonces para toda $\epsilon > 0$, existe x admisible con $\|x - x_0\|_1 < \epsilon$ tal que $I(x)$ es menor que $I(x_0)$, esto es, existe x admisible con $\|x\|_1 < \epsilon$ tal que $I(x) < 0$.

3.62 Observación: *Como veremos posteriormente si $b \in (0, \pi)$, entonces*

$$\int_0^b \{\dot{x}^2(t) - x^2(t)\} dt > 0$$

para toda x no nula que satisface (a) *y* (b) *del Ejemplo 3.61. Además,*

$$\int_0^\pi \{\dot{x}^2(t) - x^2(t)\} dt \geq 0$$

para toda x que satisface (a) *y* (b) *del Ejemplo 3.61.*

14. Teoremas de sumergimiento

En esta sección seguiremos en la dirección de demostrar la teoría de Jacobi. En particular, nuevamente con la ayuda de la teoría de las ecuaciones diferenciales y los resultados de la Sección 12, demostraremos que un extremo no singular está sumergido en una familia $2n$-paramétrica de extremos. De hecho, si estos extremos pasan por un punto de la gráfica de x_0, entonces x_0 estará sumergido en una familia n-paramétrica de éstos. Demostraremos también que si fijamos a $2n - 1$ de los parámetros y dejamos variar a uno solo de éstos, entonces al derivar la ecuación de Euler generada por estos extremos con respecto a esta variable real, dicha ecuación se convierte en la ecuación de Jacobi. Por lo tanto, en esta sección es crucial notar que los extremos accesorio, es decir, las soluciones de la ecuación de Jacobi, son de un cierto modo, *variaciones* que surgen del hecho de que un extremo no singular se encuentra sumergido continuamente en una familia de extremos.

Sean $\mathcal{F}$ una región en $\mathbf{R}^{1+n}$ y $f(t, x) \colon \mathcal{F} \to \mathbf{R}^n$ una función continua tal que

$$f_x(t, x) \colon \mathcal{F} \to \mathbf{R}^{n \times n} \text{ es continua.}$$

Supongamos que $x_0 \colon T \to \mathbf{R}^n$ es una solución de la ecuación diferencial

$$\dot{x}(t) = f(t, x(t)) \quad (t \in T). \tag{3.28}$$

Entonces, x_0 puede ser sumergido en una familia n-paramétrica de soluciones de esta ecuación. Este resultado se puede obtener al mantener α fija en el siguiente resultado.

3.63 Teorema (Teorema de sumergimiento): *Existen una constante $\rho > 0$ y una vecindad $\mathcal{F}_0$ de los puntos (t, x) sobre una solución x_0 de la ecuación diferencial (3.28) tal que a través de cada punto $(\alpha, \beta) = (\alpha, \beta^1, \ldots, \beta^n)$ en $\mathcal{F}_0$ pasa una única solución*

$$x(t, \alpha, \beta) \quad (t_0 - \rho \leq t \leq t_1 + \rho)$$

de la ecuación (3.28) tal que

$$x(t, \alpha, x_0(\alpha)) = x_0(t) \quad (t_0 \leq t \leq t_1).$$

Las funciones $x(t, \alpha, \beta)$, $\dot{x}(t, \alpha, \beta)$ son continuas y poseen derivadas continuas con respecto a β para toda (t, α, β) con t en el rango $t_0 - \rho \leq t \leq t_1 + \rho$ y (α, β) en $\mathcal{F}_0$. Además, el determinante

$$|x_\beta(t, \alpha, \beta)|$$

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

El Teorema 3.64 es una consecuencia del Teorema 3.63 y del hecho de que x_0 es un extremo no singular si y solo si x_0 junto con su variable canónica $p(\cdot) := L^*_{\dot{x}}(\cdot, x_0(\cdot), \dot{x}_0(\cdot))$ satisfacen una ecuación diferencial de la forma (3.28) en el espacion $2n$-dimensional.

3.64 Teorema: *Un extremo no singular x_0 es un miembro de una familia $2n$-paramétrica*

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

de extremos para los valores $\alpha = \alpha_0$, $\beta = \beta_0$, $t_0 \leq t \leq t_1$. Las funciones $x(t, \alpha, \beta)$ y

$$p(t, \alpha, \beta) = L^*_{\dot{x}}(t, x(t, \alpha, \beta), \dot{x}(t, \alpha, \beta))$$

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

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

es diferente de cero a lo largo de x_0.

Como demostraremos posteriormente, si $d(t, \alpha_0, \beta_0)$ es diferente de cero en un punto $t = t^0$, éste será diferente de cero en un intervalo $t_0 - \delta \leq t \leq t_1 + \delta$. Por ejemplo, si (α, β) se elige de tal forma que

$$\alpha = x(t^0, \alpha, \beta), \quad \beta = p(t^0, \alpha, \beta), \tag{3.29}$$

entonces $\tilde{d}(t^0, \alpha, \beta) = 1$, donde

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

De hecho, es fácil demostar que

$$\tilde{d}(t, \alpha, \beta) = |L_{\dot{x}\dot{x}}(t, x(t, \alpha, \beta), \dot{x}(t, \alpha, \beta))| d(t, \alpha, \beta). \tag{3.30}$$

En efecto, obsérvese que

$$\begin{aligned} p_\alpha(t, \alpha, \beta) &= L_{\dot{x}x} x_\alpha(t, \alpha, \beta) + L_{\dot{x}\dot{x}} \dot{x}_\alpha(t, \alpha, \beta), \\ p_\beta(t, \alpha, \beta) &= L_{\dot{x}x} x_\beta(t, \alpha, \beta) + L_{\dot{x}\dot{x}} \dot{x}_\beta(t, \alpha, \beta), \end{aligned} \tag{3.31}$$

donde los argumentos en las derivadas de L son $(t, x(t, \alpha, \beta), \dot{x}(t, \alpha, \beta))$. Claramente,

$$\begin{aligned} \tilde{d}(t, \alpha, \beta) = \begin{vmatrix} x_\alpha(t, \alpha, \beta) & x_\beta(t, \alpha, \beta) \\ p_\alpha(t, \alpha, \beta) & p_\beta(t, \alpha, \beta) \end{vmatrix} &= \begin{vmatrix} x_\alpha(t, \alpha, \beta) & x_\beta(t, \alpha, \beta) \\ L_{\dot{x}\dot{x}} \dot{x}_\alpha(t, \alpha, \beta) & L_{\dot{x}\dot{x}} \dot{x}_\beta(t, \alpha, \beta) \end{vmatrix} \\ &= \begin{vmatrix} I & 0 \\ 0 & L_{\dot{x}\dot{x}} \end{vmatrix} \cdot \begin{vmatrix} x_\alpha(t, \alpha, \beta) & x_\beta(t, \alpha, \beta) \\ \dot{x}_\alpha(t, \alpha, \beta) & \dot{x}_\beta(t, \alpha, \beta) \end{vmatrix} \\ &= |L_{\dot{x}\dot{x}}| \begin{vmatrix} x_\alpha(t, \alpha, \beta) & x_\beta(t, \alpha, \beta) \\ \dot{x}_\alpha(t, \alpha, \beta) & \dot{x}_\beta(t, \alpha, \beta) \end{vmatrix} \\ &= |L_{\dot{x}\dot{x}}| d(t, \alpha, \beta), \end{aligned}$$

donde los argumentos en $L_{\dot{x}\dot{x}}$ son $(t, x(t,\alpha,\beta), \dot{x}(t,\alpha,\beta))$ y (3.30) se verifica.

Por el Teorema 3.64, existe una constante $\delta > 0$ tal que

$$x(t,\alpha_0,\beta_0) \quad (t_0 - \delta \leq t \leq t_1 + \delta)$$

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

Si en la familia del Teorema 3.64 los parámetros (α, β) se eligen de tal forma que (3.29) se satisfaga, entonces fijando $\alpha = \alpha_0 = x_0(t^0)$ uno obtiene el siguiente.

3.65 Teorema: *Un extremo no singular x_0 es un miembro, para los valores $\beta = \beta_0$, $t_0 \leq t \leq t_1$, de una familia n-paramétrica de extremos*

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

que pasan a través de un punto sobre x_0 (o sobre una extensión de x_0). Esta familia tiene propiedades de continuidad y diferenciabilidad descritas en el Teorema 3.64. Además, la matriz

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

tiene rango n en cada punto de x_0.

Demostración: Definamos $x(t,\beta) := x(t,\alpha_0,\beta)$ y notemos que por (3.29) y la segunda igualdad en (3.31),

$$I = L_{\dot{x}\dot{x}}(t^0, x(t^0,\beta_0), \dot{x}(t^0,\beta_0))\dot{x}_\beta(t^0,\beta_0)$$

si y solo si

$$\dot{x}_\beta(t^0,\beta_0) = L^{-1}_{\dot{x}\dot{x}}(t^0, x_0(t^0), \dot{x}_0(t^0)).$$

De esta manera, $\dot{x}_\beta(t^0,\beta_0)$ tiene n filas de dimensión n linealmente independientes. Como habíamos mencionado anteriormente, demostraremos que si la matriz

$$\begin{pmatrix} x_\beta(t,\beta_0) \\ \dot{x}_\beta(t,\beta_0) \end{pmatrix}$$

tiene rango n en un solo punto $t = t^0$, entonces ésta tendrá rango n para todos los puntos $t \in [t_0 - \delta, t_1 + \delta]$, el intervalo correspondiente a una extensión de x_0. ∎

3.66 Teorema: *Sea un extremo x_0 un miembro para $\epsilon = 0$, $t_0 \leq t \leq t_1$, de una familia uniparamétrica de extremos*

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

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

$$y: \qquad y(t) = x_\epsilon(t,0) \quad (t \in T)$$

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

$$\frac{d}{dt}[L_{\dot{x}x}y(t) + L_{\dot{x}\dot{x}}\dot{y}(t)] = L_{xx}y(t) + L_{x\dot{x}}\dot{y}(t) \quad (t \in T). \tag{3.32}$$

Aquí, las derivadas de segundo orden de L están evaluadas en los puntos $(t, x_0(t), \dot{x}_0(t))$.

Demostración: Este resultado se obtiene de la relación

$$\frac{\partial}{\partial t}L_{\dot{x}}(t, x(t,\epsilon), \dot{x}(t,\epsilon)) = L_x(t, x(t,\epsilon), \dot{x}(t,\epsilon))$$

al derivar con respecto a ϵ en $\epsilon = 0$. ∎

3.67 Observación: *Obsérvese que la ecuación que aparece en (3.32) es la ecuación de Jacobi con respecto al arco x_0.*

15. Determinación de puntos conjugados

El propósito principal de esta sección es derivar algunos resultados técnicos que nos serán de utilidad en la Secciones 16 y 18. Con estos resultados en la mano en la Sección 16 concluiremos con los enunciados y las demostraciones que forman parte de la teoría de Jacobi.

3.68 Lema: *Un conjunto de r extremos accesorio $y_1, \dots, y_r$ es linealmente independiente si y solo si la matriz*

$$\begin{pmatrix} y_1(t) \cdots y_r(t) \\ \dot{y}_1(t) \cdots \dot{y}_r(t) \end{pmatrix}$$

tiene rango r en algún punto $t = t^0$. Si ésta tiene rango r en $t = t^0$, ésta tiene rango r para todos los valores de t en $t_0 \leq t \leq t_1$. Existen $2n$ extremos accesorio linealmente independientes. Si $y_1, \dots, y_{2n}$ son $2n$ extremos accesorio linealmente independientes, entonces todo extremo accesorio y se puede expresar por

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

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

Demostración: Supongamos que $y_1, \dots, y_r$ son linealmente independientes, y que existe un punto t^1 en T tal que

$$\sum_1^r \gamma_i (y_i(t^1), \dot{y}_i(t^1)) = 0,$$

donde las $\gamma_i's$ son constantes no todas cero. El arco $w = \sum_1^r \gamma_i y_i$ es un extremo accesorio con $w(t^1) = \dot{w}(t^1) = 0$. Entonces $w \equiv 0$ y por lo tanto $\gamma_i = 0$ $(i = 1, \dots, r)$ lo cual es una contradicción.

Supongamos que $y_1, \dots, y_r$ son linealmente dependientes. Entonces existe $\alpha = (\alpha_1, \dots, \alpha_r)^* \neq 0$ tal que $\sum_1^r \alpha_i y_i(t) = 0$ en T. Entonces $\sum_1^r \alpha_i \dot{y}_i(t) = 0$ en T. Esto implica que

$$\sum_1^r \alpha_i (y_i(t), \dot{y}_i(t)) = 0 \text{ en } T \text{ con } \alpha \neq 0,$$

y entonces la matriz del lema no tiene rango r en ningún punto de T.

Si la matriz del lema tiene rango r en algún punto $t = t^0$, entonces para toda $\beta = (\beta_1, \dots, \beta_r)^*$ distinta de cero, tenemos que

$$\begin{aligned} \sum_1^r \beta_i (y_i(t^0), \dot{y}_i(t^0)) &= \left(\sum_1^r \beta_i y_i(t^0), \sum_1^r \beta_i \dot{y}_i(t^0) \right) \\ &= (z(t^0; \beta), \dot{z}(t^0; \beta)) \neq 0 \end{aligned}$$

donde $z(t; \beta) = \sum_1^r \beta_i y_i(t)$ es un extremo accesorio y por lo tanto $(z(t; \beta), \dot{z}(t; \beta)) \neq 0$ para toda $t \in T$. De esta manera, la matriz tiene rango r para todos los valores de t en T.

Por el Teorema 3.63, escogiendo $\bar{t}$ fija en T, y por el sistema lineal

$$\dot{y}(t) = A(t)y(t) + B(t)q(t), \quad \dot{q}(t) = C(t)y(t) + D(t)q(t) \quad (t \in T), \tag{3.33}$$

sabemos que un extremo accesorio es un miembro de una familia $2n$-paramétrica de extremos accesorio $y(t, \alpha, \beta) = y(t, \alpha^1, \dots, \alpha^n, \beta^1, \dots, \beta^n)$. Entonces podemos tomar $2n$ vectores

$$(\alpha_i, \beta_i)^* = (\alpha_i^1, \dots, \alpha_i^n, \beta_i^1, \dots, \beta_i^n) \quad (i = 1, \dots, 2n)$$

linealmente independientes tales que a través de cada uno pasa una solución (y_i, q_i) de (3.33), i.e., existen $2n$ soluciones del sistema lineal (3.33) tales que $(y_i(\bar{t}), q_i(\bar{t})) = (\alpha_i, \beta_i)$ y en consecuencia la matriz

$$\begin{pmatrix} y_1(t) \cdots y_{2n}(t) \\ q_1(t) \cdots q_{2n}(t) \end{pmatrix}$$

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

$$\begin{aligned}
\begin{pmatrix} y_1(t) \cdots y_{2n}(t) \\ \dot{y}_1(t) \cdots \dot{y}_{2n}(t) \end{pmatrix} &= \begin{pmatrix} y_1(t) \cdots y_{2n}(t) \\ A(t)y_1(t) + B(t)q_1(t) \cdots A(t)y_{2n}(t) + B(t)q_{2n}(t) \end{pmatrix} \\
&= \begin{pmatrix} y_1(t) \cdots y_{2n}(t) \\ A(t)y_1(t) + L_{\dot{x}\dot{x}}^{-1}q_1(t) \cdots A(t)y_{2n}(t) + L_{\dot{x}\dot{x}}^{-1}q_{2n}(t) \end{pmatrix},
\end{aligned}$$

donde los argumentos en las derivadas de $L_{\dot{x}\dot{x}}^{-1}$ son $(t, x_0(t), \dot{x}_0(t))$. En consecuencia, las dos matrices tienen rango $2n$ en $t = \bar{t}$. Por lo tanto, $y_1, \ldots, y_{2n}$ son $2n$ extremos accesorio linealmente independientes.

Sean $y_1, \ldots, y_{2n}$ $2n$ extremos accesorio linealmente independientes, y sea y un extremo accesorio no nulo. Sea $c \in T$ tal que $y(c) \neq 0$. Entonces, existen constantes $b_i's$ $(i = 1, \ldots, 2n)$ no todas cero tal que $(y(c), \dot{y}(c)) = \sum_1^{2n} b_i(y_i(c), \dot{y}_i(c))$. El arco $z = \sum_1^{2n} b_i y_i$ es un extremo accesorio que cumple con $z(c) = y(c)$ y $\dot{z}(c) = \dot{y}(c)$. De esta manera, $y = \sum_1^{2n} b_i y_i$. ▮

3.69 Teorema: *Sean $y_1, \ldots, y_n, z_1, \ldots, z_n$ $2n$ extremos accesorio linealmente independientes y definamos*

$$D(t, t^0) := \begin{vmatrix} y_1(t) \cdots y_n(t) & z_1(t) \cdots z_n(t) \\ y_1(t^0) \cdots y_n(t^0) & z_1(t^0) \cdots z_n(t^0) \end{vmatrix}.$$

Los puntos $t = c$ conjugados a $t = t_0$ son los ceros $c \neq t_0$ de $D(t, t_0)$.

Demostración: Supongamos que c es conjugado a t_0. Sea y un extremo accesorio no nulo tal que $y(t_0) = y(c) = 0$. Seleccionamos las constantes α_j, β_j tales que

$$y = \sum_1^n \alpha_j y_j + \sum_1^n \beta_j z_j. \tag{3.34}$$

Puesto que $y \neq 0$, estas constantes no son todas iguales a cero. El hecho de que $y(t_0) = 0$ y $y(c) = 0$, nos dice que

$$\begin{pmatrix} y_1(c) \cdots y_n(c) & z_1(c) \cdots z_n(c) \\ y_1(t_0) \cdots y_n(t_0) & z_1(t_0) \cdots z_n(t_0) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{3.35}$$

En consecuencia, $D(c, t_0) = 0$. Recíprocamente, si $D(c, t_0) = 0$ $(c \neq t_0)$, existen constantes α_j, β_j no todas cero, tales que (3.35) se satisface. El arco definido por (3.34) es un extremo accesorio no nulo que satisface $y(t_0) = y(c) = 0$. Consecuentemente c es conjugado a t_0. ▮

3.70 Corolario: *Sea*

$$x(t, \alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_n)$$

una familia $2n$-paramétrica de extremos que contiene a x_0 para los valores $\alpha = \alpha_0$, $\beta = \beta_0$, $t_0 \leq t \leq t_1$ y que tiene las propiedades descritas en el Teorema 3.64. Un punto c es conjugado a t_0 sobre x_0 si y solo si $c \neq t_0$ y $D(c, t_0, \alpha_0, \beta_0) = 0$, donde

$$D(t, t^0, \alpha, \beta) := \begin{vmatrix} x_\alpha(t, \alpha, \beta) & x_\beta(t, \alpha, \beta) \\ x_\alpha(t^0, \alpha, \beta) & x_\beta(t^0, \alpha, \beta) \end{vmatrix}.$$

Demostración: Este resultado se obtiene de los Teoremas 3.64, 3.66, del Lema 3.68 y del Teorema 3.69. ▮

3.71 Teorema: *Sean $z_1, \ldots, z_n$ n extremos accesorio linealmente independientes que se anulan en $t = t_0$, y definamos*

$$\Delta(t) := |z_1(t) \cdots z_n(t)|.$$

Un punto c es conjugado a t_0 si y solo si $c \neq t_0$ y $\Delta(c) = 0$.

Demostración: Este resultado es una consecuencia del Teorema 3.69 al seleccionar n extremos accesorio $y_1, \ldots, y_n$ linealmente independientes tales que $|y_1(t_0) \cdots y_n(t_0)| = 1$, puesto que de esta manera tenemos

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

Nótese que la elección de los extremos accesorio $y_1, \ldots, y_n$ tales que $|y_1(t_0) \cdots y_n(t_0)| = 1$ se puede llevar a cabo de la siguiente manera. Sean (y_i, q_i) $(i = 1, \ldots, n)$, $(z_i, \bar{q}_i)$ $(i = 1, \ldots, n)$ $2n$ soluciones de la ecuación diferencial canónica

$$\dot{y}(t) = A(t)y(t) + B(t)q(t), \quad \dot{q}(t) = C(t)y(t) + D(t)q(t) \quad (t \in T)$$

tales que

$$(y_1(t_0), q_1(t_0)) = (1, \ldots, 0, 0, \ldots, 0)^*, \ldots, (y_n(t_0), q_n(t_0)) = (0, \ldots, 1, 0, \ldots, 0)^*,$$
$$(z_1(t_0), \bar{q}_1(t_0)) = (0, \ldots, 0, 1, \ldots, 0)^*, \ldots, (z_n(t_0), \bar{q}_n(t_0)) = (0, \ldots, 0, 0, \ldots, 1)^*.$$

Entonces, $y_1, \ldots, y_n, z_1, \ldots, z_n$ son $2n$ extremos accesorio linealmente independientes tales que

$$|y_1(t_0) \cdots y_n(t_0)| = 1$$

y $z_1, \ldots, z_n$ se anulan en $t = t_0$. ▮

3.72 Corolario: *Sea*

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

una familia n-paramétrica de extremos que pasan a través del punto inicial de x_0 para los valores $\beta = \beta_0$, $t_0 \leq t \leq t_1$ y que tienen las propiedades descritas en el Teorema 3.65. Un punto c es conjugado a t_0 sobre x_0 si y solo si $c \neq t_0$ y $\Delta(c, \beta_0) = 0$, donde

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

Demostración: De la relación

$$x_0(t_0) = x(t_0, \beta)$$

al derivar se obtiene que $x_\beta(t_0, \beta_0) = 0$. Por el Teorema 3.65 y por el Lema 3.68, las funciones

$$z_j(t) := x_{\beta_j}(t, \beta_0) \quad (j = 1, \ldots, n)$$

son extremos accesorio linealmente independientes que se anulan en $t = t_0$. El corolario es por lo tanto una consecuencia del Teorema 3.71. ▮

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

$$D(t, t^0) = \begin{vmatrix} y_1(t) \cdots y_{2n}(t) \\ y_1(t^0) \cdots y_{2n}(t^0) \end{vmatrix}.$$

Existe una constante $\delta > 0$ tal que $D(t, t^0) \neq 0$ para toda pareja de puntos distintos t, t^0 en $t_0 \leq t \leq t_1$ con $|t - t^0| \leq \delta$.

Demostración: Obsérvese que para toda $j = 1, \ldots, 2n$,

$$y_j(t) - y_j(t^0) = (t - t^0) A_j(t, t^0)$$

donde

$$A_j(t, t^0) = \int_0^1 \dot{y}_j(t^0 + \theta(t - t^0)) d\theta.$$

Por lo tanto,

$$D(t,t^0) = \begin{vmatrix} y_1(t) - y_1(t^0) \cdots y_{2n}(t) - y_{2n}(t^0) \\ y_1(t^0) \cdots y_{2n}(t^0) \end{vmatrix} = (t-t^0)^n \Delta(t,t^0),$$

donde

$$\Delta(t,t^0) = \begin{vmatrix} A_1(t,t^0) \cdots A_{2n}(t,t^0) \\ y_1(t^0) \cdots y_{2n}(t^0) \end{vmatrix}.$$

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

$$\Delta(t,t) = \begin{vmatrix} \dot{y}_1(t) \cdots \dot{y}_{2n}(t) \\ y_1(t) \cdots y_{2n}(t) \end{vmatrix} \neq 0 \tag{3.36}$$

en $t_0 \leq t \leq t_1$. Sea $\{(t_m, t_m^0)\}_{m \in \mathbf{N}}$ una sucesión en $T \times T$ tal que

$$\lim_{m \to \infty} (t_m, t_m^0) = (t, t^0).$$

Puesto que podemos elegir $K_j \in \mathbf{R}$ tal que $|\dot{y}_j(t^0 + \theta(t - t^0))| \leq K_j$ para toda $t, t^0 \in T$, para toda $\theta \in [0,1]$ y para toda $j = 1, \ldots, n$, entonces por el teorema de la convergencia acotada, para toda $j = 1, \ldots, n$, se tiene que

$$\begin{aligned} \lim_{m \to \infty} A_j(t_m, t_m^0) &= \lim_{m \to \infty} \int_0^1 \dot{y}_j(t_m^0 + \theta(t_m - t_m^0)) d\theta \\ &= \int_0^1 \lim_{m \to \infty} \dot{y}_j(t_m^0 + \theta(t_m - t_m^0)) d\theta \\ &= \int_0^1 \dot{y}_j(t^0 + \theta(t - t^0)) d\theta \\ &= A_j(t, t^0), \end{aligned}$$

y así, $\Delta(t,t^0)$ es continua en t y t^0. Por (3.36), sin pérdida de generalidad podemos suponer que existe $\epsilon_0 > 0$ tal que $\Delta(t^0,t^0) \geq \epsilon_0$ para toda $t^0 \in T$. Como $\Delta(t,t^0)$ es continua en t y t^0 se sigue que para toda $\epsilon \in (0, \epsilon_0)$, hay una constante $\delta_\epsilon > 0$ tal que

$$|t - t^0| \leq \delta_\epsilon \Longleftrightarrow |(t,t^0) - (t^0,t^0)| \leq \delta_\epsilon \Longrightarrow |\Delta(t,t^0) - \Delta(t^0,t^0)| < \epsilon.$$

De esta manera,

$$|t - t^0| \leq \delta_\epsilon \Longrightarrow \epsilon > |\Delta(t^0,t^0)| - |\Delta(t,t^0)|$$

y entonces

$$|t - t^0| \leq \delta_\epsilon \Longrightarrow |\Delta(t,t^0)| > |\Delta(t^0,t^0)| - \epsilon \geq \epsilon_0 - \epsilon > 0.$$

Consecuentemente, haciendo $\delta = \delta_\epsilon$, tenemos que $\Delta(t,t^0) \neq 0$ para toda t, t^0 en $t_0 \leq t \leq t_1$ con $|t - t^0| \leq \delta$. En consecuencia, $D(t,t^0) = (t-t^0)^n \Delta(t,t^0) \neq 0$ si $t \neq t^0$ y $|t - t^0| \leq \delta$. ∎

16. Positividad de la segunda variación

Como vimos en la Sección 9, la no negatividad de la segunda variación sobre el conjunto de variaciones admisibles es una condición necesaria de segundo orden. Es importante observar que esta condición es aplicable aun cuando la solución propuesta x_0 es singular. No obstante, así como sucede con la anulación de la primera variación sobre el conjunto de variaciones admisibles, en general, la verificabilidad de la condición que involucra a la segunda variación tiene el mismo grado de dificultad que aquel que se presenta al tratar de resolver el problema directamente. Con el propósito de resolver este asunto, Jacobi caracterizó a la no negatividad de la segunda variación sobre el conjunto de variaciones admisibles con las condiciones de Legendre y Jacobi. Por un lado, esta caracterización es muy satisfactoria puesto que la verificabilidad de las condiciones de Legendre y Jacobi es una tarea muy esperanzadora. Por otro lado, la teoría de Jacobi tiene la pequeña desventaja de que ésta únicamente es aplicable para soluciones no singulares de clase C^1. De esta manera, debemos de tener en mente que las condiciones necesarias de segundo en la teoría clásica del cálculo de variaciones tienen sus limitaciones debido a que éstas nos dan una información limitada cuando los candidatos a resolver el problema (P) tienen picos y éstos son singulares.

Puesto que un mínimo débil del problema (P) satisface la condición de Legendre, como podremos observar sin ninguna dificultad, en esta sección será suficiente suponer que x_0 es un extremo no singular que satisface la condición de Legendre.

Como x_0 es no singular, éste puede ser extendido a la izquierda de t_0 y a la derecha de t_1 para obtener un extremo más grande

$$\bar{x}_0: \qquad \bar{x}_0(t) \quad (t \in [t_0 - \epsilon, t_1 + \epsilon],\, \epsilon > 0).$$

Este arco será llamado una extensión de x_0.

3.74 Lema: *Si y y z son extremos accesorio, la expresión*

$$\omega_{\dot{y}}(t, y(t), \dot{y}(t))z(t) - \omega_{\dot{y}}(t, z(t), \dot{z}(t))y(t)$$

es constante en T.

Demostración: Obsérvese que

$$\begin{aligned}
&\frac{d}{dt}[\omega_{\dot{y}}(t, y(t), \dot{y}(t))z(t) - \omega_{\dot{y}}(t, z(t), \dot{z}(t))y(t)] \\
&= \omega_y(t, y(t), \dot{y}(t))z(t) + \omega_{\dot{y}}(t, y(t), \dot{y}(t))\dot{z}(t) \\
&\qquad - \omega_y(t, z(t), \dot{z}(t))y(t) - \omega_{\dot{y}}(t, z(t), \dot{z}(t))\dot{y}(t) \\
&= \langle L_{xx}y(t) + L_{x\dot{x}}\dot{y}(t), z(t)\rangle + \langle L_{\dot{x}x}y(t) + L_{\dot{x}\dot{x}}\dot{y}(t), \dot{z}(t)\rangle \\
&\qquad - \langle L_{xx}z(t) + L_{x\dot{x}}\dot{z}(t), y(t)\rangle - \langle L_{\dot{x}x}z(t) + L_{\dot{x}\dot{x}}\dot{z}(t), \dot{y}(t)\rangle \\
&= \langle L_{xx}y(t) + L_{x\dot{x}}\dot{y}(t), z(t)\rangle + \langle L_{\dot{x}x}y(t) + L_{\dot{x}\dot{x}}\dot{y}(t), \dot{z}(t)\rangle \\
&\qquad - \langle z(t), L_{xx}y(t) + L_{x\dot{x}}\dot{y}(t)\rangle - \langle \dot{z}(t), L_{\dot{x}x}y(t) + L_{\dot{x}\dot{x}}\dot{y}(t)\rangle \\
&= 0
\end{aligned}$$

en T. Aquí, las derivadas de segundo orden de L están evaluadas en los puntos $(t, x_0(t), \dot{x}_0(t))$. ∎

3.75 Lema: *Si no existe un punto $c \in (t_0, t_1]$ conjugado a t_0 sobre x_0, entonces hay un punto $\bar{t}$ a la izquierda de t_0 en una extensión de x_0 que no tiene ningún punto conjugado sobre x_0, es decir, para toda $c \in T$, no existe ningún extremo accesorio no nulo y tal que $y(\bar{t}) = y(c) = 0$.*

Demostración: Sean $y_1, \ldots, y_{2n}$ extremos accesorio linealmente independientes definidos en un intervalo $[t_0 - \epsilon, t_1 + \epsilon]$ y que pertenecen a una extensión $\bar{x}_0$ de x_0. Por el Teorema 3.73, existe una constante $\delta > 0$ tal que

$$D(t, t^0) = \begin{vmatrix} y_1(t) \cdots y_{2n}(t) \\ y_1(t^0) \cdots y_{2n}(t^0) \end{vmatrix}$$

es diferente de cero para todos los puntos distintos t, t^0 en $[t_0 - \epsilon, t_1 + \epsilon]$ que satisfacen la relación $|t - t^0| \leq \delta$. Seleccionemos $\delta \leq 2\epsilon$. Puesto que no existe ningún punto c conjugado a t_0 en $(t_0, t_1]$, tenemos, por el Teorema 3.69,

$$D(t, t_0) \neq 0 \quad (t \in (t_0, t_1]),$$

y en consecuencia

$$D(t, t_0) \neq 0 \quad (t \in [t_0 + \delta/2, t_1]).$$

Por continuidad hay un punto $\bar{t} \in [t_0 - \delta/2, t_0)$ tal que $D(t, \bar{t}) \neq 0$ $(t \in [t_0 + \delta/2, t_1])$. Puesto que un punto $t \in (\bar{t}, t_0 + \delta/2]$ difiere de $\bar{t}$ por a lo más δ, también se tiene que $D(t, \bar{t}) \neq 0$ $(t \in (\bar{t}, t_0 + \delta/2])$. Consecuentemente, $D(t, \bar{t}) \neq 0$ $(t \in (\bar{t}, t_1])$. De esta manera no existe ningún punto $c \in (\bar{t}, t_1]$ conjugado a $\bar{t}$ sobre $\bar{x}_0$. ∎

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

$$q_j(t) := \omega^*_{\dot{y}}(t, y_j(t), \dot{y}_j(t)) \quad (t \in T). \tag{3.37}$$

En virtud del Lema 3.74, para toda $(j, k = 1, \ldots, n)$ las expresiones

$$\langle y_j(t), q_k(t)\rangle - \langle q_j(t), y_k(t)\rangle \tag{3.38}$$

tienen los mismos valores para toda $t \in T$.

3.76 Definición: *Se dirá que los extremos accesorio $y_1, \ldots, y_n$ forman un 'sistema conjugado' si éstos son linealmente independientes y las expresiones (3.38) son todas cero.*

3.77 Teorema: *Si no existe ningún punto $c \in (t_0, t_1]$ sobre x_0 conjugado a t_0, entonces hay un sistema conjugado $y_1, \ldots, y_n$ de extremos accesorio cuyo determinante*

$$\Delta(t) = |y_1(t) \cdots y_n(t)| \tag{3.39}$$

es diferente de cero para toda $t \in T$.

Demostración: Sea $\bar{t}$ que esté relacionada a x_0 como en el Lema 3.75 y sean $y_1, \ldots, y_n$ extremos accesorio linealmente independientes que se anulan en $t = \bar{t}$. Puesto que las expresiones (3.38) se anulan en $t = \bar{t}$, estos extremos forman un sistema conjugado. El determinante correspondiente (3.39) se anula solamente en $t = \bar{t}$ y en los puntos conjugados de $\bar{t}$. Este determinante por lo tanto no se anula en T. ∎

3.78 Teorema: *La segunda variación $I''(x_0, y)$ de I a lo largo de x_0 es positiva para toda $y \neq 0$ en Y si y solo si hay un sistema conjugado $y_1, \ldots, y_n$ de extremos accesorio cuyo determinante (3.39) es diferente de cero en T.*

Demostración: Por la demostración del Teorema 3.59, si $I''(x_0, \cdot)$ es positiva en Y, no puede haber un punto c conjugado a t_0 sobre x_0 en (t_0, t_1). Si t_1 fuera conjugado a t_0, existiría un extremo accesorio no nulo y con $y(t_0) = y(t_1) = 0$. Para este arco uno tendría que $I''(x_0, y) = 0$ e $I''(x_0, \cdot)$ no podría ser positiva en Y. En consecuencia, si $I''(x_0, \cdot)$ es positiva en Y, no existe ningún punto c conjugado a t_0 en $(t_0, t_1]$. En este caso, por el Teorema 3.77, hay un sistema conjugado cuyo determinante (3.39) es diferente de cero en T.

Supongamos ahora que hay un sistema conjugado $y_1, \ldots, y_n$ con un determinante (3.39) distinto de cero en T. Dado un arco $y \in Y$, la ecuación

$$y(t) = \sum_1^n y_j(t) w^j \quad (t \in T) \tag{3.40}$$

tiene una única solución $w\colon T \to \mathbf{R}^n$ C^1 a pedazos en T que satisface $w(t_0) = w(t_1) = 0$. Se tiene que

$$\dot{y}(t) = \sum_1^n \dot{y}_j(t) w^j(t) + \kappa(t) \quad \left(\kappa = \sum_1^n y_j \dot{w}^j\right). \tag{3.41}$$

Demostraremos que

$$I''(x_0, y) = \int_{t_0}^{t_1} \langle L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t))\kappa(t), \kappa(t)\rangle dt. \tag{3.42}$$

Como un primer paso obsérvese que, puesto que y_j $(j = 1, \ldots, n)$ es un extremo accesorio, la variable canónica correspondiente q_j $(j = 1, \ldots, n)$ definida por (3.37) satisface que

$$\dot{q}_j(t) = \omega_y^*(t, y_j(t), \dot{y}_j(t)) \quad (t \in T, j = 1, \ldots, n).$$

Tenemos que para toda $t \in T$,

$$\begin{aligned}
\omega_y^*(t, y(t), \dot{y}(t)) &= L_{xx} y(t) + L_{x\dot{x}} \dot{y}(t) \\
&= L_{xx} \sum_1^n y_j(t) w^j(t) + L_{x\dot{x}} \left(\sum_1^n \dot{y}_j(t) w^j(t) + \kappa(t) \right) \\
&= \sum_1^n L_{xx} y_j(t) w^j(t) + \sum_1^n L_{x\dot{x}} \dot{y}_j(t) w^j(t) + L_{x\dot{x}} \kappa(t) \\
&= \sum_1^n (L_{xx} y_j(t) + L_{x\dot{x}} \dot{y}_j(t)) w^j(t) + L_{x\dot{x}} \kappa(t) \\
&= \sum_1^n \omega_y^*(t, y_j(t), \dot{y}_j(t)) w^j(t) + L_{x\dot{x}} \kappa(t) \\
&= \sum_1^n \dot{q}_j(t) w^j(t) + L_{x\dot{x}} \kappa(t),
\end{aligned}$$

$$\begin{aligned}
\omega_{\dot{y}}^*(t, y(t), \dot{y}(t)) &= L_{\dot{x}x} y(t) + L_{\dot{x}\dot{x}} \dot{y}(t) \\
&= L_{\dot{x}x} \sum_1^n y_j(t) w^j(t) + L_{\dot{x}\dot{x}} \left(\sum_1^n \dot{y}_j(t) w^j(t) + \kappa(t) \right) \\
&= \sum_1^n L_{\dot{x}x} y_j(t) w^j(t) + \sum_1^n L_{\dot{x}\dot{x}} \dot{y}_j(t) w^j(t) + L_{\dot{x}\dot{x}} \kappa(t) \\
&= \sum_1^n (L_{\dot{x}x} y_j(t) + L_{\dot{x}\dot{x}} \dot{y}_j(t)) w^j(t) + L_{\dot{x}\dot{x}} \kappa(t) \\
&= \sum_1^n \omega_{\dot{y}}^*(t, y_j(t), \dot{y}_j(t)) w^j(t) + L_{\dot{x}\dot{x}} \kappa(t) \\
&= \sum_1^n q_j(t) w^j(t) + L_{\dot{x}\dot{x}} \kappa(t),
\end{aligned}$$

donde los argumentos en las segundas derivadas de L son los puntos $(t, x_0(t), \dot{x}_0(t))$. En consecuencia, con la ayuda de (3.40) y (3.41),

$$
\begin{aligned}
2\omega(t,y(t),\dot y(t)) &= \omega_y(t,y(t),\dot y(t))y(t) + \omega_{\dot y}(t,y(t),\dot y(t))\dot y(t)\\
&= \left(\sum_1^n \dot q_j^*(t)w^j(t) + \kappa^*(t)L_{\dot x x}\right)y(t)\\
&+ \left(\sum_1^n q_j^*(t)w^j(t) + \kappa^*(t)L_{\dot x\dot x}\right)\dot y(t)\\
&= \left(\sum_1^n \dot q_j^*(t)w^j(t) + \kappa^*(t)L_{\dot x x}\right)\sum_1^n y_k(t)w^k(t)\\
&+ \left(\sum_1^n q_j^*(t)w^j(t) + \kappa^*(t)L_{\dot x\dot x}\right)\left(\sum_1^n \dot y_k(t)w^k(t) + \kappa(t)\right)\\
&= \sum_{j,k} \dot q_j^*(t)y_k(t)w^j(t)w^k(t) + \kappa^*(t)L_{\dot x x}\sum_1^n y_k(t)w^k(t)\\
&+ \sum_{j,k} q_j^*(t)\dot y_k(t)w^j(t)w^k(t) + \kappa^*(t)L_{\dot x\dot x}\sum_1^n \dot y_k(t)w^k(t)\\
&+ \left(\sum_1^n q_j^*(t)w^j(t)\right)\kappa(t) + \kappa^*(t)L_{\dot x\dot x}\kappa(t)\\
&= \sum_{j,k}(\dot q_j^*(t)y_k(t) + q_j^*(t)\dot y_k(t))w^j(t)w^k(t)\\
&+ \kappa^*(t)\sum_1^n (L_{\dot x x}y_k(t) + L_{\dot x\dot x}\dot y_k(t))w^k(t)\\
&+ \left(\sum_1^n q_j^*(t)w^j(t)\right)\left(\sum_1^n y_k(t)\dot w^k(t)\right) + \langle L_{\dot x\dot x}\kappa(t),\kappa(t)\rangle\\
&= \sum_{j,k}(\dot q_j^*(t)y_k(t) + q_j^*(t)\dot y_k(t))w^j(t)w^k(t)\\
&+ \kappa^*(t)\sum_1^n \omega_{\dot y}^*(t,y_k(t),\dot y_k(t))w^k(t)\\
&+ \sum_{j,k} q_j^*(t)y_k(t)w^j(t)\dot w^k(t) + \langle L_{\dot x\dot x}\kappa(t),\kappa(t)\rangle\\
&= \sum_{j,k}(\dot q_j^*(t)y_k(t) + q_j^*(t)\dot y_k(t))w^j(t)w^k(t) + \kappa^*(t)\sum_1^n q_k(t)w^k(t)\\
&+ \sum_{j,k} q_j^*(t)y_k(t)w^j(t)\dot w^k(t) + \langle L_{\dot x\dot x}\kappa(t),\kappa(t)\rangle\\
&= \sum_{j,k}(\dot q_j^*(t)y_k(t) + q_j^*(t)\dot y_k(t))w^j(t)w^k(t) + \sum_{j,k} y_j^*(t)q_k(t)\dot w^j(t)w^k(t)\\
&+ \sum_{j,k} q_j^*(t)y_k(t)w^j(t)\dot w^k(t) + \langle L_{\dot x\dot x}\kappa(t),\kappa(t)\rangle\\
&= \sum_{j,k}(\dot q_j^*(t)y_k(t) + q_j^*(t)\dot y_k(t))w^j(t)w^k(t)\\
&+ 2\sum_{j,k} q_j^*(t)y_k(t)w^j(t)\dot w^k(t) + \langle L_{\dot x\dot x}\kappa(t),\kappa(t)\rangle,
\end{aligned}
$$

donde los argumentos en las segundas derivadas de L son los puntos $(t, x_0(t), \dot x_0(t))$. Puesto que $y_1, \dots, y_n$

es un sistema conjugado, se tiene que

$$q_j^* y_k = q_k^* y_j \quad (j, k = 1, \ldots, n).$$

Consecuentemente,

$$\begin{aligned}
&\frac{d}{dt}\left[\sum_{j,k} q_j^*(t) y_k(t) w^j(t) w^k(t)\right] + \langle L_{\dot{x}\dot{x}}\kappa(t), \kappa(t)\rangle \\
&= \sum_{j,k} \dot{q}_j^*(t) y_k(t) w^j(t) w^k(t) + \sum_{j,k} q_j^*(t) \frac{d}{dt}(y_k(t) w^j(t) w^k(t)) + \langle L_{\dot{x}\dot{x}}\kappa(t), \kappa(t)\rangle \\
&= \sum_{j,k} \dot{q}_j^*(t) y_k(t) w^j(t) w^k(t) + \sum_{j,k} q_j^*(t)\left(\dot{y}_k(t) w^j(t) w^k(t) + y_k(t)\frac{d}{dt}[w^j(t) w^k(t)]\right) + \langle L_{\dot{x}\dot{x}}\kappa(t), \kappa(t)\rangle \\
&= \sum_{j,k} \dot{q}_j^*(t) y_k(t) w^j(t) w^k(t) + \sum_{j,k} q_j^*(t)(\dot{y}_k(t) w^j(t) w^k(t) + y_k(t)[\dot{w}^j(t) w^k(t) + w^j(t)\dot{w}^k(t)]) + \langle L_{\dot{x}\dot{x}}\kappa(t), \kappa(t)\rangle \\
&= \sum_{j,k} \dot{q}_j^*(t) y_k(t) w^j(t) w^k(t) + \sum_{j,k} q_j^*(t) \dot{y}_k(t) w^j(t) w^k(t) \\
&+ \sum_{j,k} q_j^*(t) y_k(t) \dot{w}^j(t) w^k(t) + \sum_{j,k} q_j^*(t) y_k(t) w^j(t) \dot{w}^k(t) + \langle L_{\dot{x}\dot{x}}\kappa(t), \kappa(t)\rangle \\
&= \sum_{j,k} (\dot{q}_j^*(t) y_k(t) + q_j^*(t)\dot{y}_k(t)) w^j(t) w^k(t) \\
&+ \sum_{j,k} q_k^*(t) y_j(t) \dot{w}^j(t) w^k(t) + \sum_{j,k} q_j^*(t) y_k(t) w^j(t) \dot{w}^k(t) + \langle L_{\dot{x}\dot{x}}\kappa(t), \kappa(t)\rangle \\
&= \sum_{j,k} (\dot{q}_j^*(t) y_k(t) + q_j^*(t)\dot{y}_k(t)) w^j(t) w^k(t) + 2\sum_{j,k} q_j^*(t) y_k(t) w^j(t) \dot{w}^k(t) + \langle L_{\dot{x}\dot{x}}\kappa(t), \kappa(t)\rangle \\
&= 2\omega(t, y(t), \dot{y}(t)),
\end{aligned}$$

donde los argumentos en las segundas derivadas de L son los puntos $(t, x_0(t), \dot{x}_0(t))$.

Utilizando las relaciones $w(t_0) = w(t_1) = 0$, obtenemos la fórmula (3.42). En virtud de que x_0 es no singular y de que satisface la condición de Legendre, se sigue de (3.42) que $I''(x_0, y) > 0$ al menos que

$$\kappa = \sum_1^n y_j \dot{w}^j = 0.$$

Puesto que $|y_1(t) \cdots y_n(t)| \neq 0$ en T y $w(t_0) = 0$ esto puede ocurrir solamente si $w = 0$, y en consecuencia solamente si $y = 0$. ▮

La ecuación (3.40) se llama la *transformación de Legendre.*

Si x_0 es un extremo no singular tal que

$$s \in (t_0, t_1) \Longrightarrow s \text{ no es conjugado a } t_0 \text{ sobre } x_0,$$

entonces se dice que x_0 satisface la *condición de Jacobi.* Similarmente, si x_0 es un extremo no singular tal que

$$s \in (t_0, t_1] \Longrightarrow s \text{ no es conjugado a } t_0 \text{ sobre } x_0,$$

entonces se dice que x_0 satisface la *condición reforzada de Jacobi.*

En los Teoremas 3.82 y 3.83 se encuentran las *dos* caracterizaciones *más importantes* de la teoría de Jacobi. El Teorema 3.79 es una consecuencia inmediata de los Teoremas 3.77 y 3.78. Como veremos en el Teorema 3.83, los Teoremas 3.79 y 3.81 implicarán que si x_0 es un extremo no singular, entonces las condiciones reforzadas de Legendre y Jacobi son equivalentes a la positividad de la segunda variación sobre el conjunto de variaciones admisibles no nulas, es decir, los Teoremas 3.79 y 3.81 implicarán la segunda caracterización más importante de la teoría de Jacobi.

3.79 Teorema: *La desigualdad $I''(x_0, y) > 0$ se cumple para toda $y \neq 0$ en Y si y solo si no existe ningún punto $c \in (t_0, t_1]$ conjugado a t_0 sobre x_0.*

Los Teoremas 3.80 y 3.81 implicarán que si x_0 es un extremo no singular, entonces las condiciones de Legendre y Jacobi son equivalentes a la no negatividad de la segunda variación sobre el conjunto de variaciones admisibles, es decir, los Teoremas 3.80 y 3.81 implicarán la primera caracterización más importante de la teoría de Jacobi.

3.80 Teorema: *La desigualdad $I''(x_0, y) \geq 0$ se cumple para toda y en Y si y solo si no existe ningún punto $c \in (t_0, t_1)$ conjugado a t_0 sobre x_0.*

Demostración:

'$\Rightarrow$': La demostración se encuentra dentro de la prueba del Teorema 3.59.

'$\Leftarrow$': Supongamos que

$$c \in (t_0, t_1) \Longrightarrow c \text{ no es un punto conjugado a } t_0 \text{ sobre } x_0, \tag{3.43}$$

y que además existe $y \in Y$ tal que $I''(x_0, y) < 0$. Por continuidad, existe $\bar{t} \in (t_0, t_1)$ y $\bar{y}\colon [t_0, \bar{t}] \to \mathbf{R}^n$ C^1 a pedazos con $\bar{y}(t_0) = \bar{y}(\bar{t}) = 0$ tal que

$$\int_{t_0}^{\bar{t}} 2\omega(\tau, \bar{y}(\tau), \dot{\bar{y}}(\tau))d\tau < 0.$$

Por el Teorema 3.79, existe un punto $c \in (t_0, \bar{t}] \subset (t_0, t_1)$ el cual es un punto conjugado a t_0 sobre x_0, lo cual contradice (3.43). ∎

El Teorema 3.81 es la última herramienta que necesitaremos para obtener las caracterizaciones más importantes de la teoría de Jacobi.

3.81 Teorema: *Si $I''(x_0, y) \geq 0$ para toda $y \in Y$, entonces $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) \geq 0$ para toda $t \in T$.*

Demostración: Supongamos que existe una $\bar{t} \in (t_0, t_1)$ la cual no es un punto esquina de x_0 tal que

$$L_{\dot{x}\dot{x}}(\bar{t}, \bar{x}, \bar{u}) < 0 \tag{3.44}$$

donde $\bar{x} := x_0(\bar{t})$ y $\bar{u} := \dot{x}_0(\bar{t})$. Definamos $\rho\colon \mathbf{R} \to \mathbf{R}$ por

$$\rho(\tau) := \begin{cases} \exp(-(1-\tau^2)^{-1}) & \text{si } \tau \in (-1, 1), \\ 0 & \text{en otro caso.} \end{cases}$$

Nótese que $\rho \in C_c^\infty(\mathbf{R})$, esto es, ρ es C^∞ con soporte compacto en $\mathbf{R}$ y de hecho el soporte de ρ y de todas sus derivadas es el intervalo $[-1, 1]$. Definamos $\bar{\rho}\colon \mathbf{R} \to \mathbf{R}^n$ por $\bar{\rho}(\tau) := (\rho(\tau), \ldots, \rho(\tau))^*$ $(\tau \in \mathbf{R})$. Puesto que ρ es $C_c^\infty(\mathbf{R})$, por (3.44) se tiene que para toda $\epsilon > 0$ suficientemente pequeña tal que $(s, \tau) \in (\bar{t}-\epsilon, \bar{t}+\epsilon) \times (-1, 1)$,

$$\langle \bar{\rho}'(\tau), L_{\dot{x}\dot{x}}(\tilde{x}_0(s))\bar{\rho}'(\tau) \rangle < 0, \tag{3.45}$$

donde $\tilde{x}_0(t) := (t, x_0(t), \dot{x}_0(t))$ $(t \in T)$. Ahora, para toda $\epsilon > 0$ suficientemente pequeña, definamos $y_\epsilon\colon T \to \mathbf{R}^n$ por

$$y_\epsilon(t) := \bar{\rho}((t - \bar{t})/\epsilon).$$

Para toda $\epsilon > 0$ suficientemente pequeña, tenemos que

$$\begin{aligned}
I''(x_0, y_\epsilon) &= \int_{\bar{t}-\epsilon}^{\bar{t}+\epsilon} \{\langle y_\epsilon(t), L_{xx}(\tilde{x}_0(t))y_\epsilon(t)\rangle + 2\langle y_\epsilon(t), L_{x\dot{x}}(\tilde{x}_0(t))\dot{y}_\epsilon(t)\rangle + \langle \dot{y}_\epsilon(t), L_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{y}_\epsilon(t)\rangle\}dt \\
&= \int_{-1}^{1} \{\langle \bar{\rho}(\tau), L_{xx}(\tilde{x}_0(\bar{t}+\epsilon\tau))\bar{\rho}(\tau)\rangle + 2\langle \bar{\rho}(\tau), L_{x\dot{x}}(\tilde{x}_0(\bar{t}+\epsilon\tau))\bar{\rho}'(\tau)/\epsilon\rangle \\
&\qquad + \langle \bar{\rho}'(\tau)/\epsilon, L_{\dot{x}\dot{x}}(\tilde{x}_0(\bar{t}+\epsilon\tau))\bar{\rho}'(\tau)/\epsilon\rangle\}\epsilon d\tau \\
&= \int_{-1}^{1} \{\langle \bar{\rho}(\tau), L_{xx}(\tilde{x}_0(\bar{t}+\epsilon\tau))\bar{\rho}(\tau)\rangle\epsilon + 2\langle \bar{\rho}(\tau), L_{x\dot{x}}(\tilde{x}_0(\bar{t}+\epsilon\tau))\bar{\rho}'(\tau)\rangle\}d\tau \\
&\qquad + \frac{1}{\epsilon}\int_{-1}^{1} \langle \bar{\rho}'(\tau), L_{\dot{x}\dot{x}}(\tilde{x}_0(\bar{t}+\epsilon\tau))\bar{\rho}'(\tau)\rangle d\tau.
\end{aligned}$$

Por (3.45),

$$\int_{-1}^{1} \langle \bar{\rho}'(\tau), L_{\dot{x}\dot{x}}(\tilde{x}_0(\bar{t}+\epsilon\tau))\bar{\rho}'(\tau)\rangle d\tau < 0$$

y por lo tanto para $\epsilon > 0$ suficientemente pequeña obtenemos que $I''(x_0, y_\epsilon) < 0$. Puesto que para toda $\epsilon > 0$ suficientemente pequeña $y_\epsilon \in Y$ el teorema se sigue por contraposición. ▮

Los Teoremas 3.82 y 3.83 son una componente fundamental de las condiciones necesarias y suficientes de segundo orden en la teoría clásica del cálculo de variaciones. Como mencionamos anteriormente estos teoremas se le atribuyen a Carl Gustav Jakob Jabobi (1804-1851).

En los Teoremas 3.82 y 3.83, únicamente asumamos que x_0 es un extremo no singular, esto es, x_0 es un arco sin esquinas que satisface la ecuación de Euler y $|L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t))| \neq 0$ $(t \in T)$.

3.82 Teorema: *La desigualdad $I''(x_0, y) \geq 0$ se cumple para toda y en Y si y solo si no existe ningún punto $c \in (t_0, t_1)$ conjugado a t_0 sobre x_0 y $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) \geq 0$ para toda $t \in T$.*

3.83 Teorema: *La desigualdad $I''(x_0, y) > 0$ se cumple para toda $y \neq 0$ en Y si y solo si no existe ningún punto $c \in (t_0, t_1]$ conjugado a t_0 sobre x_0 y $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ para toda $t \in T$.*

17. Las variaciones como diferenciales

En esta sección establecemos los primeros teoremas de suficiencia de este trabajo. Estos teoremas que proporcionan condiciones suficientes para óptimos locales y globales surgen del hecho de que la primera y segunda variaciones de I son el primer y segundo diferenciales de I relativos a la topología de X generada por la norma $\|\cdot\|_1$. Cabe señalar que estos teoremas de suficiencia no requieren de la hipótesis de la *no singularidad* ni de la *suavidad* del arco óptimo bajo consideración.

Recordemos que

$$X = \{x: T \to \mathbf{R}^n \mid x \text{ es } C^1 \text{ a pedazos}\},$$

$$Y = \{y \in X \mid y(t_0) = y(t_1) = 0\},$$

$$\|x\|_1 = \sup_{t \in T}(|x(t)| + |\dot{x}(t)|).$$

Dadas $\epsilon > 0$ y $x \in X$, definamos

$$\mathcal{T}_1(x;\epsilon) := \{(t, x, \dot{x}) \in T \times \mathbf{R}^n \times \mathbf{R}^n : |x - x(t)| < \epsilon,\ |\dot{x} - \dot{x}(t)| < \epsilon\}.$$

Dado un arco $x_0 \in X$ y dada $\epsilon > 0$, obsérvese que L y sus derivadas parciales de primer y segundo orden con respecto a x y a $\dot{x}$ son uniformemente continuas en la cerradura de $\mathcal{T}_1(x_0;\epsilon)$. Se sigue que si denotamos por P tanto a L como a cualquiera de sus derivadas de primer y segundo orden, entonces

$$\lim_{\|x - x_0\|_1 \to 0} P(t, x(t), \dot{x}(t)) = P(t, x_0(t), \dot{x}_0(t))$$

uniformemente en T. Utilizando este hecho, se sigue de la fórmula

$$I'(x,y) = \int_{t_0}^{t_1} \{L_x(t,x(t),\dot{x}(t))y(t) + L_{\dot{x}}(t,x(t),\dot{x}(t))\dot{y}(t)\}dt$$

que para toda $\epsilon > 0$ existe $\delta > 0$ tal que

$$|I'(x,y) - I'(x_0,y)| \leq \epsilon \|y\|_1 \tag{3.46}$$

para todos los arcos $x \in X$ tales que $\|x - x_0\|_1 < \delta$. Similarmente, utilizando la fórmula correspondiente para la segunda variación $I''(x,y)$, se puede ver fácilmente que si $\delta > 0$ es suficientemente pequeña, también tenemos que

$$|I''(x,y) - I''(x_0,y)| \leq \epsilon \|y\|_1^2 \tag{3.47}$$

siempre que $\|x - x_0\|_1 < \delta$ y x pertenezca a X. Las desigualdades (3.46) y (3.47) son equivalentes a las afirmaciones

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

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

Aplicando el teorema de Taylor a la función

$$F(\theta) := I(x_0 + \theta(x - x_0)) \quad (0 \leq \theta \leq 1),$$

obtenemos las fórmulas de Taylor

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

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

donde, al definir $y = x - x_0$,

$$\begin{aligned} R_1(x_0,y) &= \int_0^1 [I'(x_0 + \theta y, y) - I'(x_0,y)]d\theta \\ &= \int_0^1 (1-\theta) I''(x_0 + \theta y, y) d\theta, \\ R_2(x_0,y) &= \int_0^1 (1-\theta)[I''(x_0 + \theta y, y) - I''(x_0,y)]d\theta. \end{aligned} \tag{3.48}$$

En vista de (3.46) y (3.47), se sigue que

$$\lim_{x \to x_0} \frac{R_1(x_0, x - x_0)}{\|x - x_0\|_1} = 0 \quad \text{y} \quad \lim_{x \to x_0} \frac{R_2(x_0, x - x_0)}{\|x - x_0\|_1^2} = 0, \tag{3.49}$$

esto es, $I'(x_0,y)$ e $I''(x_0,y)$ son el primer y segundo diferenciales de I en $x = x_0$.

El Teorema 3.84 nos da condiciones suficientes para un mínimo débil del problema (P). Como mencionamos con anterioridad este teorema no requiere de hipótesis de no singularidad ni de suavidad de la trayectoria seleccionada para ser un óptimo local. Sin embargo, es importante hacer notar que la positividad de la segunda variación de I debe de ser a lo largo de x con x en una vecindad suficientemente pequeña de la topología inducida por la norma $\|\cdot\|_1$ y no necesariamente a lo largo de x_0. Adicionalmente, debería de observarse que la anulación de la primera variación a lo largo de x_0 sobre el conjunto de variaciones admisibles puede ser reemplazada por la forma integral de la ecuación de Euler.

3.84 Teorema: *Sea x_0 un arco admisible y supongamos que existe una constante $\delta > 0$ tal que*

$$I'(x_0,y) = 0 \quad e \quad I''(x,y) > 0 \tag{3.50}$$

para toda $y \in Y$, $y \neq 0$ y toda x admisible que satisface $\|x - x_0\|_1 < \delta$. Entonces la desigualdad

$$I(x) > I(x_0)$$

se cumple para toda $x \neq x_0$ con x admisible y $\|x - x_0\|_1 < \delta$. En particular, x_0 es un mínimo estricto débil de (P).

Demostración: Por (3.48) y (3.50), se tiene que

$$I(x) - I(x_0) = R_1(x_0, x - x_0) > 0$$

si $x \neq x_0$ con x admisible y $\|x - x_0\|_1 < \delta$. ∎

Aunque las relaciones dadas en (3.49) no son satisfechas si la norma $\|\cdot\|_1$ es reemplazada por la norma $\|\cdot\|_0$, como uno puede verificar fácilmente, no tenemos ninguna dificultad para demostrar el siguiente teorema que proporciona condiciones suficientes para mínimos locales fuertes.

3.85 Teorema: *Sea x_0 un arco admisible y supongamos que existe una constante $\delta > 0$ tal que*

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

para toda $y \in Y$, $y \neq 0$ y toda x admisible que satisface $\|x - x_0\|_0 < \delta$. Entonces la desigualdad

$$I(x) > I(x_0)$$

se cumple para toda $x \neq x_0$ con x admisible y $\|x - x_0\|_0 < \delta$. En particular, x_0 es un mínimo estricto fuerte de (P).

De hecho, también se puede verificar sin dificultad la veracidad del siguiente teorema que proporciona condiciones suficientes para mínimos globales.

3.86 Teorema: *Sea x_0 un arco admisible y supongamos que*

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

para toda $y \in Y$, $y \neq 0$ y toda x admisible. Entonces la desigualdad

$$I(x) > I(x_0)$$

se cumple para toda $x \neq x_0$ con x admisible. En particular, x_0 es un mínimo estricto global de (P).

18. Teoremas fundamentales de suficiencia en cálculo de variaciones

En esta sección estableceremos los teoremas clásicos de suficiencia del cálculo de variaciones. Demostraremos que si x_0 satisface las condiciones reforzadas de Euler, Legendre y Jacobi, entonces x_0 es un mínimo débil estricto de (P). Adicionalmente a estas condiciones, si x_0 satisface la condición reforzada de Weierstrass, entonces x_0 es un mínimo fuerte estricto de (P).

Recordemos que una *región* $\mathcal{F}$ en un espacio vectorial de dimensión finita es un conjunto abierto conexo.

3.87 Definición: *Definimos un campo de Mayer $\mathcal{F}$ como una región $\mathcal{F}$ en el espacio-tx y una función $U\colon \mathcal{F} \to \mathbf{R}^n$ tales que*

(i) *La función U es de clase C^1 en $\mathcal{F}$.*

(ii) *La integral de Hilbert*

$$\tilde{I} := \int \{L(t, x, U(t, x))dt + L_{\dot{x}}(t, x, U(t, x))(dx - U(t, x)dt)\}$$

es independiente de la trayectoria en $\mathcal{F}$.

La integral $\tilde{I}$ es de la forma

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

donde

$$P(t,x) := L^*_{\dot{x}}(t,x,U(t,x)) \quad \text{y} \quad Q(t,x) := L(t,x,U(t,x)) - P^*(t,x)U(t,x).$$

A lo largo de arcos no paramétricos

$$x: \quad t(s) = s, \ x(s) \quad (\alpha \le s \le \beta)$$

que caen en $\mathcal{F}$ la integral $\tilde{I}$ toma la forma

$$\tilde{I}(x) = \int_\alpha^\beta \tilde{L}(t,x(t),\dot{x}(t))dt$$

donde

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

Puesto que todo arco no paramétrico x en $\mathcal{F}$ minimiza a $\tilde{I}$ en la clase de todos los arcos no paramétricos que unen sus puntos finales, la ecuación

$$\begin{aligned}
0 &= \frac{d}{dt}\tilde{L}_{\dot{x}}(t,x(t),\dot{x}(t)) - \tilde{L}_x(t,x(t),\dot{x}(t)) \\
&= \frac{d}{dt}L_{\dot{x}}(t,x(t),U(t,x(t))) - \dot{x}^*(t)P_x(t,x(t)) - Q_x(t,x(t)) \\
&= \frac{d}{dt}L_{\dot{x}}(t,x(t),U(t,x(t))) - L_x(t,x(t),U(t,x(t))) - [\dot{x}^*(t) - U^*(t,x(t))]P_x(t,x(t))
\end{aligned} \tag{3.51}$$

se debe de cumplir a lo largo de dicho arco.

A lo largo de una solución de la ecuación

$$\dot{x}(t) = U(t,x(t)) \tag{3.52}$$

la ecuación (3.51) se reduce a la ecuación de Euler para L. Esto nos entrega el siguiente resultado.

3.88 Lema: *Una solución x de la ecuación diferencial (3.52) es un extremo de la integral*

$$I(x) = \int_{t_0}^{t_1} L(t,x(t),\dot{x}(t))dt.$$

siempre que $[t_0,t_1]$ esté contenido en el dominio de x. Un arco x que satisface (3.52) será llamado un extremo del campo $\mathcal{F}$. Si un arco admisible x es un extremo del campo $\mathcal{F}$, entonces $\tilde{I}(x) = I(x)$. A través de cada punto de $\mathcal{F}$ pasa uno y solamente un extremo del campo.

3.89 Observación: *Como veremos más adelante la utilidad principal de los campos de Mayer es que su existencia nos entrega una herramienta fundamental para la obtención de algunos teoremas que proporcionan condiciones suficientes para mínimos locales débiles y fuertes del problema (P). Como observaremos posteriormente no perderemos generalidad si la región $\mathcal{F}$ que involucra la definición de un campo de Mayer se escoge como $\mathcal{F} = B((t_0,\xi_0);\epsilon) \cup \mathcal{T}_0(x_0;\epsilon) \cup B((t_1,\xi_1);\epsilon)$ para alguna $\epsilon > 0$ y x_0 un arco de clase C^1, donde*

$$\mathcal{T}_0(x_0;\epsilon) := \{(t,x) \in T \times \mathbf{R}^n : |x - x_0(t)| < \epsilon\}.$$

Más aún, dado un arco paramétrico

$$X: \quad t(s), \ X(s) \quad (s_1 \le s \le s_2)$$

que cae en $\mathcal{F}$, como podremos observar posteriormente no perderemos generalidad si suponemos que $t^0 \leq t(s)$ para toda $s \in [s_1, s_2]$, donde t^0 será un punto que se encuentra a la izquierda de t_0 y que corresponde a una extensión de x_0.

3.90 Teorema (Teorema fundamental de suficiencia): *Sea x_0 admisible un extremo de un campo de Mayer*

$$\mathcal{F} = B((t_0, \xi_0); \epsilon) \cup \mathcal{T}_0(x_0; \epsilon) \cup B((t_1, \xi_1); \epsilon)$$

para alguna $\epsilon > 0$. Supongamos que $E(t, x, U(t, x), \dot{x}) > 0$ para $(t, x, \dot{x})$ con $(t, x) \in \mathcal{T}_0(x_0; \epsilon)$ y $U(t, x) \neq \dot{x}$. Entonces

$$I(x_0) < I(x) \text{ para toda } x \text{ admisible, } x \neq x_0,\ \|x - x_0\|_0 < \epsilon,$$

esto es, x_0 es un mínimo estricto fuerte de (P).

Demostración: Sea x admisible tal que $\|x - x_0\|_0 < \epsilon$. Obsérvese que

$$\tilde{I}(x) = \tilde{I}(x_0) = I(x_0).$$

Por hipótesis,

$$I(x) = \tilde{I}(x) + \int_{t_0}^{t_1} E(t, x(t), U(t, x(t)), \dot{x}(t))dt \geq I(x_0).$$

La igualdad se cumple sólo si

$$\dot{x}(t) = U(t, x(t)) \quad (t \in T).$$

Pero, puesto que $\xi_0 = x(t_0) = x_0(t_0)$, entonces $x = x_0$. Por lo tanto,

$$I(x_0) < I(x) \text{ para toda } x \text{ admisible, } x \neq x_0,\ \|x - x_0\|_0 < \epsilon. \blacksquare$$

3.91 Teorema (Teorema fundamental de suficiencia): *Sea x_0 admisible un extremo de un campo de Mayer*

$$\mathcal{F} = B((t_0, \xi_0); \epsilon) \cup \mathcal{T}_0(x_0; \epsilon) \cup B((t_1, \xi_1); \epsilon)$$

para alguna $\epsilon > 0$. Supongamos que $E(t, x, U(t, x), \dot{x}) > 0$ para $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$ y $U(t, x) \neq \dot{x}$. Entonces

$$I(x_0) < I(x) \text{ para toda } x \text{ admisible, } x \neq x_0,\ \|x - x_0\|_1 < \epsilon,$$

esto es, x_0 es un mínimo estricto débil de (P).

Demostración: Sea x admisible tal que $\|x - x_0\|_1 < \epsilon$. Obsérvese que

$$\tilde{I}(x) = \tilde{I}(x_0) = I(x_0).$$

Por hipótesis,

$$I(x) = \tilde{I}(x) + \int_{t_0}^{t_1} E(t, x(t), U(t, x(t)), \dot{x}(t))dt \geq I(x_0).$$

La igualdad se cumple sólo si

$$\dot{x}(t) = U(t, x(t)) \quad (t \in T).$$

Pero, puesto que $\xi_0 = x(t_0) = x_0(t_0)$, entonces $x = x_0$. Por lo tanto,

$$I(x_0) < I(x) \text{ para toda } x \text{ admisible, } x \neq x_0,\ \|x - x_0\|_1 < \epsilon. \blacksquare$$

El Teorema 3.92 nos asegura que los campos de Mayer en realidad sí existen. Específicamente, si x_0 satisface las condiciones reforzadas de Legendre, Euler y Jacobi, entonces existe un campo de Mayer que contiene a la gráfica de x_0. Este resultado combinado con los Teoremas 3.90 y 3.91 nos darán las herramientas necesarias para obtener los teoremas de suficiencia clásicos de la teoría del cálculo de variaciones. La demostración del Teorema 3.92 se dará un poco más adelante.

3.92 Teorema: *Sea x_0 admisible de clase C^1. Si*

a. $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

b. $\frac{d}{dt} L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = L_x(t, x_0(t), \dot{x}_0(t))$ $(t \in T)$.

c. $c \in (t_0, t_1] \Longrightarrow c$ *no es un punto conjugado a t_0 sobre x_0.*

Entonces x_0 es un extremo de un campo de Mayer $\mathcal{F}$.

Sea $S \subset \mathbf{R}^n$. Una *vecindad* de un conjunto S es un conjunto abierto que contiene a S. Una ϵ*-vecindad* de S es el conjunto de puntos que caen en una bola de radio ϵ con centro en algún punto de S. Obsérvese que si $S \subset \mathbf{R}^n$ es un conjunto cerrado no vacío y f es una función real continua definida en una vecindad de S tal que $f(x) > m$ para toda $x \in S$, entonces existe una vecindad N de S tal que $f(x) > m$ para toda $x \in N$. Si S es compacto, entonces N puede ser reemplazada por una ϵ-vecindad de S.

3.93 Lema: *Sea O un subconjunto abierto de $\mathbf{R}^q$ el cual es una vecindad de un conjunto compacto K. Sea $A(z)\colon O \to \mathbf{R}^{n\times n}$ una función continua tal que $A(z)$ es una matriz simétrica para toda $z \in O$. Sea*

$$Q(z; x) := \langle A(z)x, x\rangle \quad (z \in O, x \in \mathbf{R}^n).$$

Si $Q(z;x) > 0$ para $x \neq 0$ y para $z \in K$, entonces existe una vecindad N de K, $N \subset O$, y un número positivo m tal que la desigualdad

$$Q(z; x) \geq m|x|^2$$

se cumple para $x \in \mathbf{R}^n$ y z en N.

Demostración: El conjunto de puntos (x, z) con $|x| = 1$ y $z \in K$ es un conjunto compacto en el que $Q(z; x)$ es positiva. Por continuidad, $Q(z; x)$ es positiva en una vecindad de este conjunto, y en consecuencia en un conjunto compacto de la forma

$$|x| = 1, \quad z \in K_\epsilon \quad (\epsilon > 0),$$

donde K_ϵ es la cerradura de una ϵ-vecindad de K. El mínimo m de $Q(z; x)$ en este conjunto por lo tanto es positivo. Entonces,

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

de donde

$$Q(z; x) \geq m|x|^2$$

para toda $x \in \mathbf{R}^n$ y $z \in N := \operatorname{int} K_\epsilon$. ∎

3.94 Lema: *Sea $x_0 \in C^1$. Supongamos que*

a. $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

b. *Para alguna $\epsilon > 0$, $E(t, x, \dot{x}, u) \geq 0$ para $(t, x, \dot{x}, u)$ con $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$.*

Entonces,

$$E(t, x, \dot{x}, u) > 0 \text{ para } (t, x, \dot{x}, u) \text{ con } (t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon) \text{ y } \dot{x} \neq u.$$

Demostración: Por el Lema 3.93, podemos disminuir $\epsilon > 0$, si es necesario, de tal forma que

$$L_{\dot{x}\dot{x}}(t, x, \dot{x}) > 0 \quad ((t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)). \tag{3.53}$$

Supongamos que $E(t, x, \dot{x}, u) = 0$ para alguna $(t, x, \dot{x}, u)$ con $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$ y $u \neq \dot{x}$. Definamos $g\colon \mathbf{R}^n \to \mathbf{R}$ por $g(v) := E(t, x, v, u)$. Claramente $g(\dot{x}) = 0$ y en consecuencia g tiene un mínimo local en $v = \dot{x}$. En consecuencia, $0 = g'(\dot{x}) = -(u - \dot{x})^* L_{\dot{x}\dot{x}}(t, x, \dot{x})$. Por (3.53), $u = \dot{x}$ lo cual es una contradicción. ∎

El Teorema 3.95 es el teorema de suficiencia más importante de la teoría clásica del cálculo de variaciones para mínimos locales fuertes. Este teorema asegura que las condiciones reforzadas de Legendre, Euler, Weierstrass y Jacobi son suficientes para un mínimo fuerte estricto de (P).

3.95 Teorema: *Sea x_0 admisible de clase C^1. Supongamos que*

a. $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

b. $\frac{d}{dt} L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = L_x(t, x_0(t), \dot{x}_0(t))$ $(t \in T)$.

c. *Para alguna* $\epsilon > 0$, $E(t, x, \dot{x}, u) \geq 0$ *para* $(t, x, \dot{x}, u)$ *con* $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$.

d. $c \in (t_0, t_1] \Longrightarrow c$ *no es un punto conjugado a* t_0 *sobre* x_0.

Entonces x_0 es un mínimo fuerte estricto de (P).

Demostración: Por el Teorema 3.92, x_0 es un extremo de un campo de Mayer

$$\mathcal{F} = B((t_0, \xi_0); \delta) \cup \mathcal{T}_0(x_0; \delta) \cup B((t_1, \xi_1); \delta)$$

para alguna $\delta > 0$. En virtud del Lema 3.94,

$$E(t, x, \dot{x}, u) > 0 \text{ para } (t, x, \dot{x}, u) \text{ con } (t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon) \text{ y } \dot{x} \neq u.$$

Por la continuidad de U, si $\delta > 0$ se disminuye apropiadamente, los elementos $(t, x, U(t, x))$ caerán en $\mathcal{T}_1(x_0; \epsilon)$ siempre que $(t, x) \in \mathcal{T}_0(x_0; \delta)$. Entonces,

$$E(t, x, U(t, x), \dot{x}) > 0 \text{ para } (t, x, \dot{x}) \text{ con } (t, x) \in \mathcal{T}_0(x_0; \delta),\ U(t, x) \neq \dot{x}.$$

Por el Teorema 3.90,

$$I(x_0) < I(x) \text{ para toda } x \text{ admisible, } x \neq x_0,\ \|x - x_0\|_0 < \delta,$$

esto es, x_0 es un mínimo fuerte estricto de (P). ∎

El siguiente resultado que proporciona condiciones suficientes para mínimos estrictos fuertes del problema (P) es una consecuencia inmediata de los Teoremas 3.95 y 3.83.

3.96 Corolario: *Sea x_0 admisible de clase C^1. Supongamos que*

a. $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

b. $\frac{d}{dt} L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = L_x(t, x_0(t), \dot{x}_0(t))$ $(t \in T)$.

c. *Para alguna* $\epsilon > 0$, $E(t, x, \dot{x}, u) \geq 0$ *para* $(t, x, \dot{x}, u)$ *con* $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$.

d. $I''(x_0, y) > 0$ *para* $y \in Y$, $y \neq 0$.

Entonces x_0 es un mínimo fuerte estricto de (P).

El Teorema 3.97 es el teorema de suficiencia más importante de la teoría clásica del cálculo de variaciones para mínimos locales débiles. Este teorema asegura que las condiciones reforzadas de Legendre, Euler y Jacobi son suficientes para un mínimo débil estricto de (P).

3.97 Teorema: *Sea x_0 admisible de clase C^1. Supongamos que*

a. $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

b. $\frac{d}{dt} L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = L_x(t, x_0(t), \dot{x}_0(t))$ $(t \in T)$.

c. $c \in (t_0, t_1] \Longrightarrow c$ *no es un punto conjugado a* t_0 *sobre* x_0.

Entonces x_0 es un mínimo débil estricto de (P).

Demostración: Por el Teorema 3.92, x_0 es un extremo de un campo de Mayer

$$\mathcal{F} = B((t_0, \xi_0); \delta) \cup \mathcal{T}_0(x_0; \delta) \cup B((t_1, \xi_1); \delta)$$

para alguna $\delta > 0$. Por (a) y el Lema 3.93, existen $m, \epsilon > 0$ tales que

$$\langle L_{\dot{x}\dot{x}}(t, x, \dot{x})h, h\rangle \geq m|h|^2 \quad (h \in \mathbf{R}^n, (t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)).$$

Puesto que

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

se sigue que

$$E(t, x, \dot{x}, u) \geq 0 \text{ para } (t, x, \dot{x}, u) \text{ con } (t, x, \dot{x}) \text{ y } (t, x, u) \text{ en } \mathcal{T}_1(x_0; \epsilon).$$

Por el Lema 3.94,

$$E(t, x, \dot{x}, u) > 0 \text{ para } (t, x, \dot{x}, u) \text{ con } (t, x, \dot{x}) \text{ y } (t, x, u) \text{ en } \mathcal{T}_1(x_0; \epsilon),\ \dot{x} \neq u.$$

Por la continuidad de U, si $\delta > 0$ es disminuida apropiadamente, los elementos $(t, x, U(t, x))$ caerán en $\mathcal{T}_1(x_0; \epsilon)$ siempre que $(t, x) \in \mathcal{T}_0(x_0; \delta)$. Entonces,

$$E(t, x, U(t, x), \dot{x}) > 0 \text{ para } (t, x, \dot{x}) \in \mathcal{T}_1(x_0; \delta),\ U(t, x) \neq \dot{x}.$$

Por el Teorema 3.91,

$$I(x_0) < I(x) \text{ para toda } x \text{ admisible, } x \neq x_0,\ \|x - x_0\|_1 < \delta,$$

esto es, x_0 es un mínimo débil estricto de (P). ∎

El siguiente resultado que proporciona condiciones suficientes para mínimos estrictos débiles del problema (P) es una consecuencia inmediata de los Teoremas 3.97 y 3.83.

3.98 Corolario: *Sea x_0 admisible de clase C^1. Supongamos que*

a. $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

b. $\frac{d}{dt}L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = L_x(t, x_0(t), \dot{x}_0(t))$ $(t \in T)$.

c. $I''(x_0, y) > 0$ *para* $y \in Y$, $y \neq 0$.

Entonces x_0 es un mínimo débil estricto de (P).

Los Lemas 3.99, 3.100 y 3.101 nos ayudarán a demostrar el Teorema 3.92. Como uno puede verificar fácilmente, estos lemas son una consecuencia directa de las Secciones 14 y 15.

3.99 Lema: *Un extremo no singular x_0 es un miembro, para un valor vectorial $\beta = \beta_0$ y $t \in T$, de una familia n-paramétrica de extremos*

$$x(t, \beta)$$

los cuales pasan a través de un punto de la gráfica de x_0 (o de una extensión de x_0). Las funciones $x(t, \beta)$ y $\dot{x}(t, \beta)$ son de clase C^{m-1} si L es de clase C^m $(m \geq 2)$.

3.100 Lema: *Sea*

$$x(t, \beta)$$

una familia n-paramétrica de extremos que pasan a través del punto inicial de x_0, que contienen a x_0 para un valor $\beta = \beta_0$, $t \in T$, y tales que $x(t, \beta)$ y $\dot{x}(t, \beta)$ son de clase C^{m-1} si L es de clase C^m $(m \geq 2)$. Un punto c es conjugado a $t = t_0$ sobre x_0 si y solo si $c \neq t_0$ y $\Delta(c, \beta_0) = 0$, donde

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

3.101 Lema: *Si no existe un punto c conjugado a $t = t_0$ sobre x_0 en $(t_0, t_1]$, entonces existe un punto t^0 a la izquierda de t_0 en una extensión de x_0 que no tiene ningún punto conjugado sobre x_0.*

Demostración del Teorema 3.92: Por simplicidad en la notación, en la demostración de este teorema cuando nos refiramos a una extensión $\bar{x}_0$ de x_0 haremos el convenio de denotar a $\bar{x}_0$ simplemente como x_0.

Por (a), (b) y (c) del Teorema 3.92 y el Lema 3.101, existe un punto t^0 correspondiente a una extensión de x_0 y que se encuentra a la izquierda de t_0 tal que

$$c \in (t^0, t_1] \implies c \text{ no es un punto conjugado a } t = t^0 \text{ sobre } x_0.$$

Por (a) y (b) del Teorema 3.92, existe

$$x(t, \beta)$$

una familia de extremos que pasan por el punto $(t^0, x_0(t^0))$ y que contienen a x_0 para los valores $t \in (t^0, t_1]$ y $\beta = \beta_0$. Por el Lema 3.100, el determinante

$$\Delta(t, \beta_0) = |x_\beta(t, \beta_0)| \neq 0 \quad (t \in (t^0, t_1]).$$

Por los teoremas de la Sección 14, sabemos que existe $\gamma > 0$ tal que el dominio de las funciones $x(t, \beta)$ puede ser extendido a la izquierda de t_0 y a la derecha de t_1 de tal forma que $x(t, \beta)$ sigue siendo un extremo para $t \in [t_0 - \gamma, t_1 + \gamma]$, $\beta \in B(\beta_0; \gamma)$, y $x(t, \beta_0)$ es una extensión de x_0 que sigue estando sumergida en la familia n-paramétrica de extremos $x(t, \beta)$ que pasan por el punto $(t^0, x_0(t^0))$. Nótese que sin pérdida de generalidad, podemos suponer que $t^0 \in [t_0 - \gamma, t_1 + \gamma]$.

Ahora definamos $S := (t_0 - \gamma, t_1 + \gamma) \times \mathbf{R}^n \times B(\beta_0; \gamma)$ y $F: S \to \mathbf{R}^n$ por

$$F(t, x, \beta) := x(t, \beta) - x.$$

Notemos que S es un abierto de $\mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^n$, F es continua en S y $F_\beta: S \to \mathbf{R}^{n \times n}$ dada por $F_\beta(t, x, \beta) = x_\beta(t, \beta)$ también es continua en S. Adicionalmente, si $K_0 := \{(t, x_0(t)) \mid t \in T\}$ y $g: K_0 \to \mathbf{R}^n$ la definimos por

$$g(t, x_0(t)) := \beta_0,$$

entonces g es continua en el compacto K_0, $(t, x_0(t)) \in K_0$ implica que $(t, x_0(t), \beta_0) \in S$, $F(t, x_0(t), \beta_0) = 0$, y $|F_\beta(t, x_0(t), \beta_0)| = |x_\beta(t, \beta_0)| \neq 0$. Obsérvese además que $F_t(t, x, \beta) = \dot{x}(t, \beta)$ y $F_x(t, x, \beta) = -I_{n \times n}$ son continuas en S, es decir, F es de clase C^1 en S.

Por el Teorema 3.54, existen $\epsilon \in (0, \gamma)$, $\mathcal{F} = B((t_0, \xi_0); \epsilon) \cup \mathcal{T}_0(x_0; \epsilon) \cup B((t_1, \xi_1); \epsilon)$, y $\beta: \mathcal{F} \to \mathbf{R}^n$ de clase C^1 tales que

$$F(t, x, \beta(t, x)) = 0 \quad ((t, x) \in \mathcal{F})$$

si y solo si

$$x(t, \beta(t, x)) = x \quad ((t, x) \in \mathcal{F}). \tag{3.54}$$

La función

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

es la función pendiente de estos extremos y por el Lema 3.99, ésta es de clase C^1 en $\mathcal{F}$. Nuevamente, sin pérdida de generalidad, podemos suponer que el punto $(t^0, x_0(t^0))$ pertenece a $\mathcal{F}$.

Ahora consideremos un arco paramétrico continuo

$$X: \quad t(s),\ X(s) \quad (s_1 \leq s \leq s_2)$$

en $\mathcal{F}$ y que tiene derivadas continuas a pedazos. Las funciones

$$t(s), \quad \beta(s) := \beta(t(s), X(s)) \quad (s_1 \leq s \leq s_2)$$

determinan una familia uni-paramétrica de extremos

$$x(s): \quad x(t,s) := x(t,\beta(s)) \quad (t^0 \leq t \leq t(s)).$$

Por (3.54), obsérvese que

$$x(t(s),s) = x(t(s),\beta(s)) = x(t(s),\beta(t(s),X(s))) = X(s) \quad (s_1 \leq s \leq s_2).$$

En consecuencia, es importante notar que los puntos finales de $x(s)$ forman el arco X.

Utilizando las notaciones

$$\delta x := x_s(t,s)ds, \quad dx := \dot{x}(t,s)dt + \delta x,$$

y la ecuación de Euler, no es difícil ver que la función

$$F(s) := I(x(s)) = \int_{t^0}^{t(s)} L(t,x(t,s),\dot{x}(t,s))dt$$

es C^1 a pedazos en $[s_1,s_2]$ y que su diferencial toma la forma

$$\begin{aligned}
dF &= L(t(s),x(t(s),s),\dot{x}(t(s),s))dt \\
&+ \int_{t^0}^{t(s)} \{L_x(t,x(t,s),\dot{x}(t,s))\delta x + L_{\dot{x}}(t,x(t,s),\dot{x}(t,s))\delta\dot{x}\}dt \\
&= L(t(s),x(t(s),s),\dot{x}(t(s),s))dt \\
&+ \int_{t^0}^{t(s)} \left\{ \left(\frac{\partial}{\partial t} L_{\dot{x}}(t,x(t,s),\dot{x}(t,s)) \right) \delta x + L_{\dot{x}}(t,x(t,s),\dot{x}(t,s)) \frac{\partial}{\partial s}\frac{\partial}{\partial t} x(t,s)ds \right\} dt \\
&= L(t(s),x(t(s),s),\dot{x}(t(s),s))dt \\
&+ \int_{t^0}^{t(s)} \left\{ \left(\frac{\partial}{\partial t} L_{\dot{x}}(t,x(t,s),\dot{x}(t,s)) \right) x_s(t,s)ds + L_{\dot{x}}(t,x(t,s),\dot{x}(t,s)) \frac{\partial}{\partial t} x_s(t,s)ds \right\} dt \\
&= L(t(s),x(t(s),s),\dot{x}(t(s),s))dt + \int_{t^0}^{t(s)} \frac{\partial}{\partial t} L_{\dot{x}}(t,x(t,s),\dot{x}(t,s)) x_s(t,s)dtds \\
&= L(t(s),X(s),U(t(s),X(s)))dt + L_{\dot{x}}(t(s),X(s),U(t(s),X(s)))x_s(t(s),s)ds \\
&= L(t(s),X(s),U(t(s),X(s)))dt \\
&+ L_{\dot{x}}(t(s),X(s),U(t(s),X(s)))[dX - U(t(s),X(s))dt] \\
&= d\tilde{I} \text{ a lo largo de } X.
\end{aligned}$$

Integrando desde s_1 hasta s_2 se sigue que

$$F(s_2) - F(s_1) = I(x(s_2)) - I(x(s_1)) = \tilde{I}(X).$$

Ahora consideremos otro arco paramétrico continuo

$$\bar{X}: \quad \bar{t}(s), \quad \bar{X}(s) \quad (\bar{s}_1 \leq s \leq \bar{s}_2)$$

en $\mathcal{F}$, que tiene derivadas continuas a pedazos y que satisface que

$$(\bar{t}(\bar{s}_i), \bar{X}(\bar{s}_i)) = (t(s_i), X(s_i)) \quad (i = 1,2). \tag{3.55}$$

Las funciones

$$\bar{t}(s), \quad \bar{\beta}(s) := \beta(\bar{t}(s), \bar{X}(s)) \quad (\bar{s}_1 \leq s \leq \bar{s}_2)$$

determinan una familia uni-paramétrica de extremos

$$\bar{x}(s): \quad \bar{x}(t,s) := x(t,\bar{\beta}(s)) \quad (t^0 \leq t \leq \bar{t}(s)).$$

Por (3.54), obsérvese que

$$\bar{x}(\bar{t}(s), s) = x(\bar{t}(s), \bar{\beta}(s)) = x(\bar{t}(s), \beta(\bar{t}(s), \mathcal{X}(s))) = \bar{X}(s) \quad (\bar{s}_1 \leq s \leq \bar{s}_2).$$

Así, los puntos finales de $\bar{x}(s)$ forman el arco $\bar{X}$.

Utilizando las notaciones

$$\delta\bar{x} := \bar{x}_s(t, s)ds, \quad d\bar{x} := \dot{\bar{x}}(t, s)dt + \delta\bar{x},$$

y la ecuación de Euler, no es difícil ver que la diferencial de la función

$$\bar{F}(s) := I(\bar{x}(s)) = \int_{t^0}^{\bar{t}(s)} L(t, \bar{x}(t, s), \dot{\bar{x}}(t, s))dt$$

toma la forma

$$\begin{aligned}
d\bar{F} &= L(\bar{t}(s), \bar{x}(\bar{t}(s), s), \dot{\bar{x}}(\bar{t}(s), s))d\bar{t} \\
&+ \int_{t^0}^{\bar{t}(s)} \{L_x(t, \bar{x}(t, s), \dot{\bar{x}}(t, s))\delta\bar{x} + L_{\dot{x}}(t, \bar{x}(t, s), \dot{\bar{x}}(t, s))\delta\dot{\bar{x}}\}dt \\
&= L(\bar{t}(s), \bar{X}(s), U(\bar{t}(s), \bar{X}(s)))d\bar{t} + L_{\dot{x}}(\bar{t}(s), \bar{X}(s), U(\bar{t}(s), \bar{X}(s)))\bar{x}_s(\bar{t}(s), s)ds \\
&= L(\bar{t}(s), \bar{X}(s), U(\bar{t}(s), \bar{X}(s)))d\bar{t} \\
&+ L_{\dot{x}}(\bar{t}(s), \bar{X}(s), U(\bar{t}(s), \bar{X}(s)))[d\bar{X} - U(\bar{t}(s), \bar{X}(s))d\bar{t}\,] \\
&= d\tilde{I} \text{ a lo largo de } \bar{X}.
\end{aligned}$$

Integrando desde $\bar{s}_1$ hasta $\bar{s}_2$ se sigue que

$$\bar{F}(\bar{s}_2) - \bar{F}(\bar{s}_1) = I(\bar{x}(\bar{s}_2)) - I(\bar{x}(\bar{s}_1)) = \tilde{I}(\bar{X}).$$

Por (3.55), para $i = 1, 2$

$$x(t, \beta(s_i)) = x(t, \bar{\beta}(\bar{s}_i)) \quad (t \in [t^0, t_1]). \tag{3.56}$$

Por (3.56), para $i = 1, 2$

$$x(s_i) = x(\cdot, s_i) = x(\cdot, \beta(s_i)) = x(\cdot, \bar{\beta}(\bar{s}_i)) = \bar{x}(\cdot, \bar{s}_i) = \bar{x}(\bar{s}_i). \tag{3.57}$$

Por (3.57), para $i = 1, 2$

$$\dot{x}(\cdot, s_i) = \dot{\bar{x}}(\cdot, \bar{s}_i). \tag{3.58}$$

Por (3.57) y (3.58), para $i = 1, 2$

$$F(s_i) = I(x(s_i)) = I(\bar{x}(\bar{s}_i)) = \bar{F}(\bar{s}_i)$$

de donde se obtiene que

$$\tilde{I}(X) = \tilde{I}(\bar{X}).$$

De esta manera, la integral $\tilde{I}$ es independiente de la trayectoria en $\mathcal{F}$. ▮

3.102 Observación: *Nótese que en la demostración del Teorema 3.92 ha sido crucial la hipótesis de que L sea de clase C^2 con respecto a todas sus variables t, x y $\dot{x}$. Por lo tanto, la técnica de campos de Mayer nos permite aplicar los teoremas de suficiencia 3.95 y 3.97 únicamente cuando L es C^2 en todas sus variables.*

3.103 Observación: *Dado un arco x_0 a las condiciones* (a), (b), (c) *y* (d) *del Teorema 3.95 se les conoce como las condiciones 'reforzadas de Legendre, Euler, Weierstrass, y Jacobi' respectivamente. Vale la pena mencionar que el reforzamiento de la condición de Weierstrass no consiste en reemplazar la no negatividad de la función exceso E dada en la condición necesaria por su positividad así como se lleva a cabo* *en la condición reforzada de Legendre. De hecho, en el ejemplo siguiente demostramos que si un arco x_0 satisface*

a. $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ *($t \in T$).*

b. $\frac{d}{dt} L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = L_x(t, x_0(t), \dot{x}_0(t))$ *($t \in T$).*

c. $E(t, x_0(t), \dot{x}_0(t), u) > 0$ *para* $(t, u) \in T \times \mathbf{R}^n$, $u \neq \dot{x}_0(t)$.

d. $I''(x_0, y) > 0$ *para* $y \in Y$, $y \neq 0$.

Entonces x_0 no necesariamente es un mínimo fuerte de (P).

3.104 Ejemplo: *Consideremos el problema (P) de minimizar*

$$I(x) = \int_0^1 \{\dot{x}^2(t) - 4x(t)\dot{x}^3(t) + 2t\dot{x}^4(t)\} dt$$

sujeta a

(a) $x\colon [0,1] \to \mathbf{R}$ *es C^1 a pedazos.*

(b) $x(0) = x(1) = 0$.

En este caso $T = [0,1]$, $n = 1$, $\xi_0 = \xi_1 = 0$, y $L(t, x, \dot{x}) = \dot{x}^2 - 4x\dot{x}^3 + 2t\dot{x}^4$. La ecuación de Euler

$$\frac{d}{dt}[2\dot{x}(t) - 12x(t)\dot{x}^2(t) + 8t\dot{x}^3(t)] = -4\dot{x}^3(t) \quad (t \in T)$$

es satisfecha por el arco admisible $x_0(t) = 0$ ($t \in T$) y por lo tanto 3.103(b) se verifica. Puesto que

$$L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) = 2 \quad (t \in T),$$

entonces, la condición 3.103(a) también se cumple. Puesto que

$$E(t, x_0(t), \dot{x}_0(t), u) = u^2 + 2tu^4,$$

se tiene que x_0 satisface la condición 3.103(c). La segunda variación de I a lo largo de x_0 está dada por

$$I''(x_0, y) = \int_0^1 2\dot{y}^2(t) dt$$

la cual es positiva para toda $y \in Y$, $y \neq 0$. De esta manera, la condición 3.103(d) también se cumple.

Para toda $k, h > 0$, consideremos los arcos $x_k^h\colon T \to \mathbf{R}$ definidos por

$$x_k^h(t) := \begin{cases} \frac{k}{h} t & \text{si } t \in [0, h], \\ \frac{k}{1-h}(1-t) & \text{si } t \in [h, 1]. \end{cases}$$

Claramente, para toda $k, h > 0$, los arcos x_k^h son admisibles. Además,

$$\begin{aligned} I(x_k^h) &= \int_0^h \left[\frac{k^2}{h^2} - \frac{2k^4}{h^4} t\right] dt + \int_h^1 \left[\frac{k^2}{(1-h)^2} + \frac{k^4}{(1-h)^4}(4-2t)\right] dt \\ &= -\frac{k^4}{h^2} + \frac{k^2}{h} + \frac{k^2}{1-h} + \frac{k^4}{(1-h)^3}(3-h). \end{aligned}$$

Tenemos que $|x_k^h(t)| \leq k$ ($t \in T, k, h > 0$). Para cada k fija se puede ver que $I(x_k^h) < 0 = I(x_0)$ siempre que h sea suficientemente pequeña. En consecuencia, I no alcanza un mínimo fuerte en el arco admisible x_0 para el problema (P).

19. Una propiedad de la función exceso de Weierstrass

En esta sección presentamos una propiedad muy relevante de la función exceso de Weierstrass. Esta propiedad se encuentra en el Teorema 3.105 y nos dice que si las condiciones reforzadas de Legendre y Weierstrass son satisfechas alrededor de un arco C^1 a pedazos, entonces la función exceso de Weierstrass no solo es estrictamente positiva si $u \neq \dot{x}$ como lo asegura el Corolario 3.94 sino que además dicha función exceso está acotada inferiormente por una proporción de una función que es muy parecida a la función cuadrática alrededor del cero y muy parecida a la función valor absoluto alrededor de $-\infty$ y $+\infty$.

Definamos $V: \mathbf{R}^n \to \mathbf{R}$ por

$$V(c) := (1 + |c|^2)^{1/2} - 1,$$

y notemos que para toda $c \in \mathbf{R}^n$,

$$V(c) \leq \frac{|c|^2}{2} \quad \text{y} \quad V(c) \leq |c|.$$

3.105 Lema: *Supongamos que $x_0: T \to \mathbf{R}^n$ es C^1 a pedazos y que satisface*

i. $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

ii. *Para alguna $\epsilon > 0$, $E(t, x, \dot{x}, u) \geq 0$ para $(t, x, \dot{x}, u)$ con $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$.*

Entonces existen $\delta, h > 0$ tales que, para toda $(t, x, \dot{x}, u)$ con $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \delta)$,

$$E(t, x, \dot{x}, u) \geq hV(u - \dot{x}).$$

Demostración: Por el Lema 3.93, existen $\epsilon_0, h_0 > 0$ tales que

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

para toda $c \in \mathbf{R}^n$ y toda $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon_0)$. Supongamos que $\epsilon_0 < \epsilon$. Si $(t, x, \dot{x})$ y (t, x, u) pertenecen a $\mathcal{T}_1(x_0; \epsilon_0)$, tenemos que

$$E(t, x, \dot{x}, u) = \int_0^1 (1 - \lambda) \langle u - \dot{x}, L_{\dot{x}\dot{x}}(t, x, \dot{x} + \lambda[u - \dot{x}])(u - \dot{x}) \rangle d\lambda \geq \tfrac{1}{2} h_0 |u - \dot{x}|^2.$$

Sea $\delta > 0$ tal que la cerradura de $\mathcal{T}_1(x_0; \delta)$ esté contenida en $\mathcal{T}_1(x_0; \epsilon_0)$, y sea $\rho > 0$ tal que $(t, x, \dot{x} + c) \in \mathcal{T}_1(x_0; \epsilon_0)$ para toda $c \in \mathbf{R}^n$ con $|c| \leq \rho$ y $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \delta)$.

Ahora, tomemos $(t, x, \dot{x}, u)$ con $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \delta)$. Si $(t, x, u) \in \mathcal{T}_1(x_0; \epsilon_0)$, el resultado se sigue con $h = h_0$, puesto que

$$E(t, x, \dot{x}, u) \geq \tfrac{1}{2} h_0 |u - \dot{x}|^2 \geq h_0 V(u - \dot{x}).$$

Si $(t, x, u) \notin \mathcal{T}_1(x_0; \epsilon_0)$, sea $k := |u - \dot{x}|/\rho$ and $c := (u - \dot{x})/k$, y observemos que $|c| = \rho$ y $k > 1$. En consecuencia,

$$\begin{aligned} E(t, x, \dot{x}, u) =& E(t, x, \dot{x}, \dot{x} + kc) \\ =& E(t, x, \dot{x} + c, \dot{x} + kc) + kE(t, x, \dot{x}, \dot{x} + c) + (k - 1)E(t, x, \dot{x} + c, \dot{x}) \\ \geq& kE(t, x, \dot{x}, \dot{x} + c) \geq \tfrac{1}{2} k h_0 |c|^2 = \tfrac{1}{2} h_0 |c||kc| \geq \tfrac{1}{2} h_0 \rho V(kc) = \tfrac{1}{2} h_0 \rho V(u - \dot{x}). \end{aligned}$$

El resultado se sigue eligiendo $h = \min\{h_0, h_0\rho/2\}$. ▮

20. Lemas auxiliares

En esta sección demostraremos dos resultados auxiliares que al ser combinados con el Teorema 3.105 nos darán unas herramientas fundamentales para demostrar los teoremas de suficiencia que se presentan en las Secciones 21, 22, 23 y 24.

Definimos

$$\mathcal{X} := \{x: T \to \mathbf{R}^n \mid x \text{ es absolutamente continua en } T\},$$

y definamos las funciones $V\colon \mathbf{R}^n \to \mathbf{R}$, $D\colon \mathcal{X} \to \mathbf{R}$ por

$$V(c) := (1+|c|^2)^{1/2} - 1 \quad \text{y} \quad D(x) := V(x(t_0)) + \int_{t_0}^{t_1} V(\dot{x}(t))dt.$$

En los siguientes dos lemas, asumiremos que tenemos dadas una $x_0 \in \mathcal{X}$ y una sucesión $\{x_q\}$ en $\mathcal{X}$ tal que

$$\lim_{q\to\infty} D(x_q - x_0) = 0 \quad \text{y} \quad d_q := [2D(x_q - x_0)]^{1/2} > 0 \quad (q \in \mathbf{N}).$$

Para toda $q \in \mathbf{N}$ y $t \in T$, definimos

$$y_q(t) := \frac{x_q(t) - x_0(t)}{d_q}.$$

Dadas una sucesión $\{u_q\}$ en $L^1(T;\mathbf{R}^n)$ y $u_0 \in L^1(T;\mathbf{R}^n)$, diremos que $u_q(t) \to u_0(t)$ casi uniformemente en T, si dada $\epsilon > 0$, existe un conjunto medible $S_\epsilon \subset T$ con $m(S_\epsilon) < \epsilon$ tal que $u_q(t) \to u_0(t)$ uniformemente en $T \setminus S_\epsilon$.

3.106 Lema: *Existen una subsucesión de $\{x_q\}$, nuevamente denotada por $\{x_q\}$, y $y_0 \in \mathcal{X}$ con $\dot{y}_0 \in L^2(T;\mathbf{R}^n)$ tales que $\dot{x}_q(t) \to \dot{x}_0(t)$ casi uniformemente en T, $y_q(t) \to y_0(t)$ uniformemente en T, y la sucesión $\{\dot{y}_q\}$ converge débil en $L^1(T;\mathbf{R}^n)$ a $\dot{y}_0$.*

Demostración: Para toda $q \in \mathbf{N}$ y para casi toda $t \in T$, definamos

$$c_q := [1 + \tfrac{1}{2}V(x_q(t_0) - x_0(t_0))]^{1/2},$$

$$W_q(t) := [1 + \tfrac{1}{2}V(\dot{x}_q(t) - \dot{x}_0(t))]^{1/2}.$$

Obsérvese que $V(c)(2+V(c)) = |c|^2$ para toda $c \in \mathbf{R}^n$. Entonces para toda $q \in \mathbf{N}$,

$$\frac{|y_q(t_0)|^2}{c_q^2} + \int_{t_0}^{t_1} \frac{|\dot{y}_q(t)|^2}{W_q^2(t)}dt = 1.$$

Claramente, $\lim_{q\to\infty} c_q = 1$. Por lo tanto, existe una subsucesión de $\{x_q\}$, nuevamente denotada por $\{x_q\}$, una $\bar{y}_0 \in \mathbf{R}^n$ y una $\sigma_0 \in L^2(T;\mathbf{R}^n)$ tal que

$$\lim_{q\to\infty} \frac{y_q(t_0)}{c_q} = \lim_{q\to\infty} y_q(t_0) = \bar{y}_0,$$

$$\{\dot{y}_q/W_q\} \text{ converge débil en } L^2(T;\mathbf{R}^n) \text{ a } \sigma_0.$$

Ahora, definamos

$$y_0(t) := \int_{t_0}^{t} \sigma_0(s)ds + \bar{y}_0 \quad (t \in T),$$

y notemos que $y_0 \in \mathcal{X}$ con $\dot{y}_0 \in L^2(T;\mathbf{R}^n)$ puesto que $\dot{y}_0(t) = \sigma_0(t)$ (c.s. en T).

Como para toda $q \in \mathbf{N}$, $W_q^2(t) \geq W_q(t) \geq 1$ (c.s. en T), tenemos que

$$0 \leq \int_{t_0}^{t_1} [W_q(t) - 1]dt \leq \int_{t_0}^{t_1} [W_q^2(t) - 1]dt \leq \tfrac{1}{2}D(x_q - x_0).$$

Observemos también que

$$0 \leq \int_{t_0}^{t_1} [W_q(t) - 1]^2 dt = \int_{t_0}^{t_1} [W_q^2(t) - 1]dt - 2\int_{t_0}^{t_1} [W_q(t) - 1]dt.$$

De esta manera, $\|W_q u - u\|_2 \to 0$ para toda $u \in L^\infty(T; \mathbf{R}^n)$. Entonces,

$$\lim_{q\to\infty} \int_{t_0}^{t_1} \langle \dot{y}_q(t), u(t)\rangle dt = \lim_{q\to\infty} \int_{t_0}^{t_1} \left\langle \frac{\dot{y}_q(t)}{W_q(t)}, W_q(t)u(t)\right\rangle dt = \int_{t_0}^{t_1} \langle \dot{y}_0(t), u(t)\rangle dt$$

para toda $u \in L^\infty(T; \mathbf{R}^n)$. Por lo tanto, $\{\dot{y}_q\}$ converge débil en $L^1(T; \mathbf{R}^n)$ a $\dot{y}_0$. Entonces, la sucesión $\{\dot{y}_q\}$ es equi-integrable en T, lo cual implica que $\{y_q\}$ es equi-continua en T. Puesto que para toda $q \in \mathbf{N}$, $y_q(t) = \int_{t_0}^{t} \dot{y}_q(s)ds + y_q(t_0)$ $(t \in T)$, deducimos que $y_q(t) \to y_0(t)$ uniformemente en T.

Sea $x \in \mathcal{X}$ y definamos $W(t) := [1 + \frac{1}{2}V(\dot{x}(t))]^{1/2}$ (c.s. en T). Observemos que

$$\int_{t_0}^{t_1} W^2(t)dt \leq t_1 - t_0 + \tfrac{1}{2}D(x) \quad \text{y} \quad \int_{t_0}^{t_1} \frac{|\dot{x}(t)|^2}{W^2(t)}dt \leq 2D(x).$$

Entonces,

$$\begin{aligned} \|\dot{x}\|_1^2 &= \left(\int_{t_0}^{t_1} |\dot{x}(t)|dt\right)^2 \\ &\leq \left(\int_{t_0}^{t_1} \frac{|\dot{x}(t)|^2}{W^2(t)}dt\right)\left(\int_{t_0}^{t_1} W^2(t)dt\right) \\ &\leq D(x)[2(t_1 - t_0) + D(x)]. \end{aligned}$$

Consecuentemente,

$$\lim_{q\to\infty} \|\dot{x}_q - \dot{x}_0\|_1 = 0,$$

lo cual implica que existe una subsucesión de $\{\dot{x}_q\}$, nuevamente denotada por $\{x_q\}$, tal que $\dot{x}_q(t) \to \dot{x}_0(t)$ casi uniformemente en T. ▮

3.107 Lema: *Sea $S \subset T$ un conjunto medible, $R_0 \in L^\infty(S; \mathbf{R}^{n\times n})$ y $\{R_q\}$ una sucesión en $L^\infty(S; \mathbf{R}^{n\times n})$. Si $\dot{x}_q(t) \to \dot{x}_0(t)$ uniformemente en S, $R_q(t) \to R_0(t)$ uniformemente en S, y $R_0(t) \geq 0$ $(t \in S)$, entonces*

$$\liminf_{q\to\infty} \int_S \langle \dot{y}_q(t), R_q(t)\dot{y}_q(t)\rangle dt \geq \int_S \langle \dot{y}_0(t), R_0(t)\dot{y}_0(t)\rangle dt.$$

Demostración: Recordando la definición de W_q dada en la demostración del Lema 3.106, obsérvese que como $\dot{x}_q(t) \to \dot{x}_0(t)$ uniformemente en S, entonces $W_q(t) \to 1$ uniformemente en S. De esta manera, para toda $h \in L^2(S; \mathbf{R}^n)$,

$$\lim_{q\to\infty} \int_S \langle \dot{y}_q(t), h(t)\rangle dt = \lim_{q\to\infty} \int_S \left\langle \frac{\dot{y}_q(t)}{W_q(t)}, W_q(t)h(t)\right\rangle dt = \int_S \langle \dot{y}_0(t), h(t)\rangle dt,$$

esto es, $\{\dot{y}_q\}$ converge débil en $L^2(S; \mathbf{R}^n)$ a $\dot{y}_0$. Puesto que $R_0(t) \geq 0$ $(t \in S)$, la función

$$v \mapsto \int_S \langle v(t), R_0(t)v(t)\rangle dt$$

es convexa en $L^2(S; \mathbf{R}^n)$ y puesto que ésta es fuertemente continua en $L^2(S; \mathbf{R}^n)$, entonces dicha función es débilmente semicontinua inferior en $L^2(S; \mathbf{R}^n)$. De esta manera,

$$\liminf_{q\to\infty} \int_S \langle \dot{y}_q(t), R_0(t)\dot{y}_q(t)\rangle dt \geq \int_S \langle \dot{y}_0(t), R_0(t)\dot{y}_0(t)\rangle dt.$$

Como $R_q(t) \to R_0(t)$ uniformemente en S, se sigue que

$$\liminf_{q\to\infty} \int_S \langle \dot{y}_q(t), R_q(t)\dot{y}_q(t)\rangle dt \geq \int_S \langle \dot{y}_0(t), R_0(t)\dot{y}_0(t)\rangle dt. \ ▮$$

21. Una demostración de suficiencia: mínimos débiles

En esta sección, el Teorema 3.108 da nuevas condiciones suficientes para un mínimo débil estricto del problema (P). Los rasgos principales del Teorema 3.108 son que el arco óptimo propuesto no es necesariamente C^1 sino solamente esencialmente acotado y la condición reforzada de Legendre tampoco es necesariamente satisfecha sino únicamente la condición necesaria correspondiente. Además, otra componente fundamental del Teorema 3.108 es que éste puede detectar soluciones *singulares*. Adicionalmente, el Teorema 3.108 nos dice que si un arco x_0 satisface sus condiciones, entonces no solo x_0 es un mínimo estricto débil de (P), sino que la diferencia entre los costos admisibles $I(x)$ y el costo óptimo $I(x_0)$ está estimada por una proporción de una funcional que juega el papel del cuadrado de una norma.

Supongamos que tenemos dados un intervalo $T := [t_0, t_1]$ en $\mathbf{R}$, dos puntos fijos ξ_0, ξ_1 en $\mathbf{R}^n$, y una función L que mapea $T \times \mathbf{R}^n \times \mathbf{R}^n$ a $\mathbf{R}$, y denotemos por $\mathcal{X}$ al espacio vectorial de las funciones absolutamente continuas que mapean T a $\mathbf{R}^n$.

En esta sección asumiremos que la función L es de clase C^2 con respecto a x y a $\dot{x}$ en $T \times \mathbf{R}^n \times \mathbf{R}^n$.

Consideremos el problema (P) de minimizar la funcional

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt$$

sobre todas las $x \in \mathcal{X}$ con $\dot{x} \in L^\infty(T; \mathbf{R}^n)$ que satisfacen las restricciones

$$\begin{cases} L(t, x(t), \dot{x}(t)) \text{ es integrable en } T. \\ x(t_0) = \xi_0,\ x(t_1) = \xi_1. \end{cases}$$

A los elementos de $\mathcal{X}$ les llamaremos *arcos* o *trayectorias*, y una trayectoria x es *admisible* si $\dot{x} \in L^\infty(T; \mathbf{R}^n)$ y ésta satisface las restricciones. Un arco x *resuelve* (P), si x es admisible e $I(x) \leq I(y)$ para todos los arcos admisibles y. Un arco x es un *mínimo débil* de (P), si x es admisible y para alguna $\epsilon > 0$, $I(x) \leq I(y)$ para todos los arcos admisibles y que satisfacen $\|y - x\|_1 < \epsilon$, donde para toda x admisible,

$$\|x\|_1 := \|x\|_0 + \|\dot{x}\|_\infty,$$

y como siempre $\|x\|_0 = \sup_{t \in T} |x(t)|$.

Con el propósito de establecer el primer teorema de suficiencia de esta sección, introduzcamos las siguientes definiciones.

- Dadas $x, x_0 \in \mathcal{X}$, definimos

$$\tilde{x}(t) := (t, x(t), \dot{x}(t)) \quad \text{y} \quad \tilde{x}_0(t) := (t, x_0(t), \dot{x}_0(t)) \quad \text{(c.s. en } T).$$

- Para cualquier $x \in \mathcal{X}$ con $\dot{x} \in L^\infty(T; \mathbf{R}^n)$, y cualquier $y \in \mathcal{X}$, la *primera variación* de I a lo largo de x en la dirección y está dada por

$$I'(x, y) := \int_{t_0}^{t_1} \{L_x(\tilde{x}(t))y(t) + L_{\dot{x}}(\tilde{x}(t))\dot{y}(t)\}dt.$$

Además, si $y \in \mathcal{X}$ con $\dot{y} \in L^2(T; \mathbf{R}^n)$, definimos la *segunda variación* de I a lo largo de x en la dirección y por

$$I''(x, y) := \int_{t_0}^{t_1} 2\omega(t, y(t), \dot{y}(t))dt$$

donde para toda $(t, y, \dot{y}) \in T \times \mathbf{R}^n \times \mathbf{R}^n$,

$$2\omega(t, y, \dot{y}) := \langle y, L_{xx}(\tilde{x}(t))y \rangle + 2\langle y, L_{x\dot{x}}(\tilde{x}(t))\dot{y} \rangle + \langle \dot{y}, L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y} \rangle.$$

• Definamos
$$\mathcal{Y} := \{y \in \mathcal{X} \mid \dot{y} \in L^2(T; \mathbf{R}^n),\ y(t_0) = y(t_1) = 0\}.$$

• La *función exceso de Weierstrass* de L, $E: T \times \mathbf{R}^{3n} \to \mathbf{R}$, está dada por

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

• Definimos las funciones $V: \mathbf{R}^n \to \mathbf{R}$, $D: \mathcal{X} \to \mathbf{R}$ por

$$V(a) := (1 + |a|^2)^{1/2} - 1 \quad \text{y} \quad D(x) := V(x(t_0)) + \int_{t_0}^{t_1} V(\dot{x}(t))dt.$$

3.108 Teorema: *Sea x_0 un arco admisible. Supongamos que existen dos números positivos h, ϵ, y una constante $c \in \mathbf{R}^n$, tales que lo siguiente se satisface:*

(i) $L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^{t} L_x(s, x_0(s), \dot{x}_0(s))ds + c^*$ *(c.s. en T).*

(ii) $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) \geq 0$ *(c.s. en T).*

(iii) $I''(x_0, y) > 0$ *para toda $y \in \mathcal{Y}$, $y \neq 0$.*

(iv) $\int_{t_0}^{t_1} E(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq hD(x - x_0)$ *siempre que x sea admisible y $\|x - x_0\|_1 < \epsilon$.*

Entonces, para alguna $\mu, \nu > 0$ y para todo arco admisible x con $\|x - x_0\|_1 < \nu$, se tiene que

$$I(x) \geq I(x_0) + \mu D(x - x_0).$$

En particular, x_0 es un mínimo estricto débil de (P).

Demostración: La demostración se hará por contraposición, esto es, vamos a suponer que para toda $\mu, \nu > 0$, existe un arco admisible x tal que

$$\|x - x_0\|_1 < \nu \quad \text{e} \quad I(x) < I(x_0) + \mu D(x - x_0). \tag{3.59}$$

También, supondremos que (i), (ii) y (iv) del Teorema 3.108 son satisfechas, y vamos a obtener la negación de la condición (iii) del Teorema 3.108.

Primero que todo, notemos que para toda x admisible,

$$I(x) = I(x_0) + I'(x_0, x - x_0) + \mathcal{K}(x) + \mathcal{E}(x) \tag{3.60}$$

donde
$$\mathcal{E}(x) := \int_{t_0}^{t_1} E(t, x(t), \dot{x}_0(t), \dot{x}(t))dt,$$

$$\mathcal{K}(x) := \int_{t_0}^{t_1} \{M(t, x(t)) + \langle \dot{x}(t) - \dot{x}_0(t), N(t, x(t))\rangle\}dt,$$

y las funciones M y N están dadas por

$$M(t, x) := L(t, x, \dot{x}_0(t)) - L(\tilde{x}_0(t)) - L_x(\tilde{x}_0(t))(x - x_0(t)),$$

$$N(t, x) := L^*_{\dot{x}}(t, x, \dot{x}_0(t)) - L^*_{\dot{x}}(\tilde{x}_0(t)).$$

Tenemos que,

$$M(t, x) = \tfrac{1}{2}\langle x - x_0(t), P(t, x)(x - x_0(t))\rangle, \quad N(t, x) = Q(t, x)(x - x_0(t)),$$

donde
$$P(t, x) := 2\int_0^1 (1 - \lambda)L_{xx}(t, x_0(t) + \lambda(x - x_0(t)), \dot{x}_0(t))d\lambda,$$

$$Q(t,x) := \int_0^1 L_{\dot{x}x}(t, x_0(t) + \lambda(x - x_0(t)), \dot{x}_0(t))d\lambda.$$

Ahora, por (3.59), para toda $q \in \mathbf{N}$ existe x_q admisible tal que

$$\|x_q - x_0\|_1 < 1/q, \quad I(x_q) - I(x_0) < \frac{1}{q}D(x_q - x_0). \tag{3.61}$$

Consecuentemente, por la segunda desigualdad en (3.61),

$$d_q := [2D(x_q - x_0)]^{1/2} > 0 \quad (q \in \mathbf{N}),$$

y por la primera desigualdad en (3.61),

$$\lim_{q\to\infty} D(x_q - x_0) = 0.$$

Para toda $t \in T$ y $q \in \mathbf{N}$, definamos

$$y_q(t) := \frac{x_q(t) - x_0(t)}{d_q}.$$

Por el Lema 3.106, existen una subsucesión de $\{x_q\}$, nuevamente denotada por $\{x_q\}$, y $y_0 \in \mathcal{X}$ con $\dot{y}_0 \in L^2(T;\mathbf{R}^n)$, tales que $y_q(t) \to y_0(t)$ uniformemente en T, y la sucesión $\{\dot{y}_q\}$ converge débilmente en $L^1(T;\mathbf{R}^n)$ a $\dot{y}_0$.

Probemos que $y_0 \in \mathcal{Y}$, $y_0 \neq 0$ e $I''(x_0, y_0) \leq 0$ con lo cual concluiríamos la demostración del teorema. El hecho de que $y_0(t_0) = y_0(t_1) = 0$ se sigue de la definición de y_q, la admisibilidad de x_q y puesto que $y_q(t) \to y_0(t)$ uniformemente en T.

Para toda $q \in \mathbf{N}$, tenemos que

$$\frac{\mathcal{K}(x_q)}{d_q^2} = \int_{t_0}^{t_1} \left\{ \frac{M(t, x_q(t))}{d_q^2} + \left\langle \dot{y}_q(t), \frac{N(t, x_q(t))}{d_q} \right\rangle \right\} dt.$$

Nuevamente, por el Lema 3.106,

$$\lim_{q\to\infty} \frac{M(t, x_q(t))}{d_q^2} = \tfrac{1}{2}\langle y_0(t), L_{xx}(\tilde{x}_0(t))y_0(t)\rangle,$$

$$\lim_{q\to\infty} \frac{N(t, x_q(t))}{d_q} = L_{\dot{x}x}(\tilde{x}_0(t))y_0(t)$$

ambos uniformemente en T y, puesto que $\{\dot{y}_q\}$ converge débil a $\dot{y}_0$ en $L^1(T;\mathbf{R}^n)$,

$$\tfrac{1}{2}I''(x_0, y_0) = \lim_{q\to\infty} \frac{\mathcal{K}(x_q)}{d_q^2} + \tfrac{1}{2}\int_{t_0}^{t_1} \langle \dot{y}_0(t), L_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{y}_0(t)\rangle dt. \tag{3.62}$$

Se tiene que

$$\liminf_{q\to\infty} \frac{\mathcal{E}(x_q)}{d_q^2} \geq \tfrac{1}{2}\int_{t_0}^{t_1} \langle \dot{y}_0(t), L_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{y}_0(t)\rangle dt. \tag{3.63}$$

En efecto, para casi toda $t \in T$ y $q \in \mathbf{N}$,

$$\frac{1}{d_q^2}E(t, x_q(t), \dot{x}_0(t), \dot{x}_q(t)) = \tfrac{1}{2}\langle \dot{y}_q(t), R_q(t)\dot{y}_q(t)\rangle$$

donde

$$R_q(t) := 2\int_0^1 (1-\lambda)L_{\dot{x}\dot{x}}(t, x_q(t), \dot{x}_0(t) + \lambda[\dot{x}_q(t) - \dot{x}_0(t)])d\lambda.$$

Claramente,
$$\lim_{q\to\infty} R_q(t) = R_0(t) := L_{\dot{x}\dot{x}}(\tilde{x}_0(t)) \quad \text{uniformemente en } T.$$

Por el Teorema 3.108(ii), $R_0(t) \geq 0$ (c.s. en T), y de esta manera (3.63) se sigue del Lema 3.107. Además, por el Teorema 3.108(i), se sigue que $I'(x_0, y) = 0$ para toda $y \in \mathcal{X}$ con $y(t_0) = y(t_1) = 0$. Con esto en mente, (3.60), (3.61), (3.62) y (3.63),

$$\tfrac{1}{2} I''(x_0, y_0) \leq \lim_{q\to\infty} \frac{\mathcal{K}(x_q)}{d_q^2} + \liminf_{q\to\infty} \frac{\mathcal{E}(x_q)}{d_q^2} = \liminf_{q\to\infty} \frac{I(x_q) - I(x_0)}{d_q^2} \leq 0.$$

Adicionalmente, si $y_0 = 0$, entonces
$$\lim_{q\to\infty} \frac{\mathcal{K}(x_q)}{d_q^2} = 0$$

y en consecuencia, por el Teorema 3.108(iv),

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

contradiciendo la positividad de h. ▮

El Corolario 3.109 es una consecuencia inmediata del Teorema 3.108. Es importante que el lector note las similitudes del Corolario 3.109 y el Corolario 3.98.

3.109 Corolario: *Sea x_0 un arco admisible C^1 a pedazos en T. Supongamos que existe una constante $c \in \mathbf{R}^n$ tal que lo siguiente se satisface:*

(i) $L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^{t} L_x(s, x_0(s), \dot{x}_0(s))ds + c^*$ *($t \in T$).*

(ii) $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ *($t \in T$).*

(iii) $I''(x_0, y) > 0$ *para toda $y \in \mathcal{Y}$, $y \neq 0$.*

Entonces, para alguna $\mu, \nu > 0$ y para todo arco admisible x con $\|x - x_0\|_1 < \nu$, se tiene que

$$I(x) \geq I(x_0) + \mu D(x - x_0).$$

En particular, x_0 es un mínimo estricto débil de (P).

Demostración: Por el Corolario 3.109(ii) y el Lema 3.93, podemos elegir $\epsilon > 0$ de tal forma que

$$L_{\dot{x}\dot{x}}(t, x, \dot{x}) > 0 \quad ((t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)).$$

De esta manera, si $(t, x, \dot{x}, u)$ es tal que $(t, x, \dot{x})$ y (t, x, u) pertenecen a $\mathcal{T}_1(x_0; \epsilon)$, entonces,

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

Por el Lema 3.105, disminuyendo $\epsilon > 0$ si es necesario, tenemos que existe $h > 0$ tal que

$$E(t, x, \dot{x}, u) \geq hV(u - \dot{x})$$

para $(t, x, \dot{x}, u)$ con $(t, x, \dot{x})$ y (t, x, u) en $\mathcal{T}_1(x_0; \epsilon)$. Por lo tanto, obtenemos que

$$\int_{t_0}^{t_1} E(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq hD(x - x_0)$$

siempre que x sea admisible y $\|x - x_0\|_1 < \epsilon$. La conclusión del corolario se sigue del Teorema 3.108. ▮

3.110 Observación: *Nótese que como en la teoría clásica, el Corolario 3.109 es aplicable para arcos no singulares y que una de sus novedades principales es que el arco x_0 es C^1 a pedazos y no necesariamente C^1 en T. Algo que podría ser natural preguntarnos es si la positividad de la segunda variación sobre el conjunto de variaciones admisibles no nulas se puede reemplazar por la condición reforzada de Jacobi en el Corolario 3.109. La respuesta no es sencilla puesto que la teoría de Jacobi depende fuertemente de la hipótesis de que x_0 sea C^1 en T. Si x_0 es C^1 en T en el Corolario 3.109, la respuesta a la misma pregunta también puede ser complicada puesto que el Corolario 3.109 requiere que la segunda variación a lo largo de x_0 sea positiva para todas las $y \in \mathcal{Y}$ distintas de cero y la condición reforzada de Jacobi junto con (i) y (ii) del Corolario 3.109 sólo implican que la segunda variación a lo largo de x_0 es positiva para todas las variaciones admisibles no nulas que son C^1 a pedazos en T. Adicionalmente, nótese que otra de las ventajas del Corolario 3.109 es que si x_0 satisface todas sus condiciones, entonces I alcanza un mínimo débil estricto en x_0 en un conjunto más amplio de arcos admisibles que el de los C^1 a pedazos, específicamente, I alcanza un mínimo débil estricto en x_0 en el conjunto de trayectorias admisibles absolutamente continuas con derivada esencialmente acotada.*

22. Una demostración de suficiencia: mínimos fuertes

En esta sección, el Teorema 3.111 da nuevas condiciones suficientes para un mínimo fuerte estricto del problema (P). Las componentes principales del Teorema 3.111 son que el arco óptimo propuesto no es necesariamente C^1 sino solamente esencialmente acotado y éste detecta soluciones *singulares*. Adicionalmente, como lo hace el Teorema 3.108 la conclusión del Teorema 3.111 nos dice que la diferencia entre los costos admisibles $I(x)$ y el costo óptimo $I(x_0)$ está estimada por una proporción de una funcional que juega el papel del cuadrado de una norma.

Supongamos que tenemos dados un intervalo $T := [t_0, t_1]$ en $\mathbf{R}$, dos puntos fijos ξ_0, ξ_1 en $\mathbf{R}^n$, y una función L que mapea $T \times \mathbf{R}^n \times \mathbf{R}^n$ a $\mathbf{R}$, y denotemos por $\mathcal{X}$ al espacio vectorial de las funciones absolutamente continuas que mapean T a $\mathbf{R}^n$.

En esta sección asumiremos que la función L es de clase C^2 con respecto a x y a $\dot{x}$ en $T \times \mathbf{R}^n \times \mathbf{R}^n$.

Consideremos el problema (P) de minimizar la funcional

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt$$

sobre todas las $x \in \mathcal{X}$ que satisfacen las restricciones

$$\begin{cases} L(t, x(t), \dot{x}(t)) \text{ es integrable en } T. \\ x(t_0) = \xi_0, \ x(t_1) = \xi_1. \end{cases}$$

A los elementos de $\mathcal{X}$ les llamaremos *arcos* o *trayectorias*, y una trayectoria x es *admisible* si ésta satisface las restricciones. Un arco admisible x *resuelve* (P), si $I(x) \le I(y)$ para todos los arcos admisibles y. Un arco admisible x es un *mínimo fuerte* de (P), si para alguna $\epsilon > 0$, $I(x) \le I(y)$ para todos los arcos admisibles y que satisfacen $\|y - x\|_0 < \epsilon$, donde para toda $x \in \mathcal{X}$,

$$\|x\|_0 := \sup_{t \in T} |x(t)|.$$

Con el propósito de establecer el primer teorema de suficiencia de esta sección, introduzcamos las siguientes definiciones.

• Dadas $x, x_0 \in \mathcal{X}$, definimos

$$\tilde{x}(t) := (t, x(t), \dot{x}(t)) \quad \text{y} \quad \tilde{x}_0(t) := (t, x_0(t), \dot{x}_0(t)) \quad \text{(c.s. en } T).$$

• Para cualquier $x \in \mathcal{X}$ con $\dot{x} \in L^\infty(T; \mathbf{R}^n)$, y cualquier $y \in \mathcal{X}$, la *primera variación* de I a lo largo de x en la dirección y está dada por

$$I'(x, y) := \int_{t_0}^{t_1} \{L_x(\tilde{x}(t))y(t) + L_{\dot{x}}(\tilde{x}(t))\dot{y}(t)\}dt.$$

Además, si $y \in \mathcal{X}$ con $\dot{y} \in L^2(T; \mathbf{R}^n)$, definimos la *segunda variación* de I a lo largo de x en la dirección y por

$$I''(x, y) := \int_{t_0}^{t_1} 2\omega(t, y(t), \dot{y}(t)) dt$$

donde para toda $(t, y, \dot{y}) \in T \times \mathbf{R}^n \times \mathbf{R}^n$,

$$2\omega(t, y, \dot{y}) := \langle y, L_{xx}(\tilde{x}(t))y\rangle + 2\langle y, L_{x\dot{x}}(\tilde{x}(t))\dot{y}\rangle + \langle \dot{y}, L_{\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}\rangle.$$

- Definamos
$$\mathcal{Y} := \{y \in \mathcal{X} \mid \dot{y} \in L^2(T; \mathbf{R}^n),\ y(t_0) = y(t_1) = 0\}.$$

- La *función exceso de Weierstrass* de L, $E\colon T \times \mathbf{R}^{3n} \to \mathbf{R}$, está dada por

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

- Definimos las funciones $V\colon \mathbf{R}^n \to \mathbf{R}$, $D\colon \mathcal{X} \to \mathbf{R}$ por

$$V(a) := (1 + |a|^2)^{1/2} - 1 \quad \text{y} \quad D(x) := V(x(t_0)) + \int_{t_0}^{t_1} V(\dot{x}(t))dt.$$

3.111 Teorema: *Sea x_0 un arco admisible con $\dot{x}_0 \in L^\infty(T; \mathbf{R}^n)$. Supongamos que existen dos números positivos h, ϵ, y una constante $c \in \mathbf{R}^n$, tales que lo siguiente se satisface:*

(i) $L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^{t} L_x(s, x_0(s), \dot{x}_0(s))ds + c^*$ *(c.s. en T).*

(ii) $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) \geq 0$ *(c.s. en T).*

(iii) $I''(x_0, y) > 0$ *para toda $y \in \mathcal{Y}$, $y \neq 0$.*

(iv) $E(t, x(t), \dot{x}_0(t), \dot{x}(t)) \geq 0$ *(c.s. en T) siempre que x sea admisible y $\|x - x_0\|_0 < \epsilon$.*

(v) $\int_{t_0}^{t_1} E(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq hD(x - x_0)$ *siempre que x sea admisible y $\|x - x_0\|_0 < \epsilon$.*

Entonces, para alguna $\mu, \nu > 0$ y para todo arco admisible x con $\|x - x_0\|_0 < \nu$, se tiene que

$$I(x) \geq I(x_0) + \mu D(x - x_0).$$

En particular, x_0 es un mínimo estricto fuerte de (P).

Demostración: La demostración se hará por contraposición, esto es, vamos a suponer que para toda $\mu, \nu > 0$, existe un arco admisible x tal que

$$\|x - x_0\|_0 < \nu \quad \text{e} \quad I(x) < I(x_0) + \mu D(x - x_0). \tag{3.64}$$

También, supondremos que (i), (ii), (iv) y (v) del Teorema 3.111 son satisfechas, y vamos a obtener la negación de la condición (iii) del Teorema 3.111.

Primero que todo, notemos que para toda x admisible,

$$I(x) = I(x_0) + I'(x_0, x - x_0) + \mathcal{K}(x) + \mathcal{E}(x) \tag{3.65}$$

donde

$$\mathcal{E}(x) := \int_{t_0}^{t_1} E(t, x(t), \dot{x}_0(t), \dot{x}(t))dt,$$

$$\mathcal{K}(x) := \int_{t_0}^{t_1} \{M(t, x(t)) + \langle \dot{x}(t) - \dot{x}_0(t), N(t, x(t))\rangle\}dt,$$

y las funciones M y N están dadas por

$$M(t, x) := L(t, x, \dot{x}_0(t)) - L(\tilde{x}_0(t)) - L_x(\tilde{x}_0(t))(x - x_0(t)),$$

$$N(t,x) := L_{\dot{x}}^*(t,x,\dot{x}_0(t)) - L_{\dot{x}}^*(\tilde{x}_0(t)).$$

Tenemos que,

$$M(t,x) = \tfrac{1}{2}\langle x - x_0(t), P(t,x)(x - x_0(t))\rangle, \quad N(t,x) = Q(t,x)(x - x_0(t)),$$

donde

$$P(t,x) := 2\int_0^1 (1-\lambda)L_{xx}(t, x_0(t) + \lambda(x - x_0(t)), \dot{x}_0(t))d\lambda,$$

$$Q(t,x) := \int_0^1 L_{\dot{x}x}(t, x_0(t) + \lambda(x - x_0(t)), \dot{x}_0(t))d\lambda.$$

Demostremos primero que existe $c_1 > 0$ tal que, para toda x admisible con $\|x - x_0\|_0 < 1$,

$$|\mathcal{K}(x)| \leq c_1\|x - x_0\|_0[1 + D(x - x_0)]. \tag{3.66}$$

En efecto, por el hecho de que $\dot{x}_0 \in L^\infty(T;\mathbf{R}^n)$, para algunas constantes $\alpha, k > 0$ y toda x admisible con $\|x - x_0\|_0 < 1$, tenemos que para casi toda $t \in T$,

$$\begin{aligned}
&|M(t,x(t)) + \langle \dot{x}(t) - \dot{x}_0(t), N(t,x(t))\rangle| \\
&= |\langle x(t) - x_0(t), \tfrac{1}{2}P(t,x(t))(x - x_0(t)) + Q^*(t,x(t))(\dot{x}(t) - \dot{x}_0(t))\rangle| \\
&\leq |x(t) - x_0(t)|(\tfrac{1}{2}|P(t,x(t))||x(t) - x_0(t)| + |Q^*(t,x(t))||\dot{x}(t) - \dot{x}_0(t)|) \\
&\leq \alpha|x(t) - x_0(t)|(|x(t) - x_0(t)| + |\dot{x}(t) - \dot{x}_0(t)|) \\
&\leq k|x(t) - x_0(t)|(|x(t) - x_0(t)|^2 + |\dot{x}(t) - \dot{x}_0(t)|^2)^{1/2} \\
&\leq k|x(t) - x_0(t)|(1 + |\dot{x}(t) - \dot{x}_0(t)|^2)^{1/2}.
\end{aligned}$$

Definamos $c_1 := \text{máx}\{k, k(t_1 - t_0)\}$. Entonces, para toda x admisible con $\|x - x_0\|_0 < 1$,

$$|\mathcal{K}(x)| \leq c_1\|x - x_0\|_0\int_{t_0}^{t_1}[1 + V(\dot{x}(t) - \dot{x}_0(t))]dt \leq c_1\|x - x_0\|_0[1 + D(x - x_0)]$$

y por lo tanto (3.66) se cumple.

Ahora, por (3.64), para toda $q \in \mathbf{N}$ existe x_q admisible tal que

$$\|x_q - x_0\|_0 < \text{mín}\{\epsilon, 1/q\}, \quad I(x_q) - I(x_0) < \frac{1}{q}D(x_q - x_0). \tag{3.67}$$

Consecuentemente, por la segunda desigualdad en (3.67),

$$d_q := [2D(x_q - x_0)]^{1/2} > 0 \quad (q \in \mathbf{N}).$$

Por el Teorema 3.111(i) se sigue que $I'(x_0, y) = 0$ para toda $y \in \mathcal{X}$ con $y(t_0) = y(t_1) = 0$. Con esto en mente, por (3.65), el Teorema 3.111(v), (3.66) y la primera desigualdad de (3.67),

$$I(x_q) - I(x_0) = \mathcal{K}(x_q) + \mathcal{E}(x_q) \geq -c_1\|x_q - x_0\|_0 + D(x_q - x_0)(h - c_1\|x_q - x_0\|_0).$$

Por (3.67), obtenemos que

$$D(x_q - x_0)\left(h - \frac{1}{q} - \frac{c_1}{q}\right) < \frac{c_1}{q}$$

y en consecuencia,

$$\lim_{q\to\infty} D(x_q - x_0) = 0.$$

Para toda $t \in T$ y $q \in \mathbf{N}$, definamos

$$y_q(t) := \frac{x_q(t) - x_0(t)}{d_q}.$$

Por el Lema 3.106, existen una subsucesión de $\{x_q\}$, nuevamente denotada por $\{x_q\}$, y $y_0 \in \mathcal{X}$ con $\dot{y}_0 \in L^2(T;\mathbf{R}^n)$, tales que $y_q(t) \to y_0(t)$ uniformemente en T, y la sucesión $\{\dot{y}_q\}$ converge débilmente en $L^1(T;\mathbf{R}^n)$ a $\dot{y}_0$.

Probemos que $y_0 \in \mathcal{Y}$, $y_0 \neq 0$ e $I''(x_0, y_0) \leq 0$ con lo cual concluiríamos la demostración del teorema. El hecho de que $y_0(t_0) = y_0(t_1) = 0$ se sigue de la definición de y_q, la admisibilidad de x_q y puesto que $y_q(t) \to y_0(t)$ uniformemente en T.

Para toda $q \in \mathbf{N}$, tenemos que

$$\frac{\mathcal{K}(x_q)}{d_q^2} = \int_{t_0}^{t_1} \left\{ \frac{M(t, x_q(t))}{d_q^2} + \left\langle \dot{y}_q(t), \frac{N(t, x_q(t))}{d_q} \right\rangle \right\} dt.$$

Nuevamente, por el Lema 3.106,

$$\lim_{q\to\infty} \frac{M(t, x_q(t))}{d_q^2} = \tfrac{1}{2}\langle y_0(t), L_{xx}(\tilde{x}_0(t))y_0(t)\rangle,$$

$$\lim_{q\to\infty} \frac{N(t, x_q(t))}{d_q} = L_{\dot{x}x}(\tilde{x}_0(t))y_0(t)$$

ambos uniformemente en T y, puesto que $\{\dot{y}_q\}$ converge débil a $\dot{y}_0$ en $L^1(T;\mathbf{R}^n)$,

$$\tfrac{1}{2}I''(x_0, y_0) = \lim_{q\to\infty} \frac{\mathcal{K}(x_q)}{d_q^2} + \tfrac{1}{2}\int_{t_0}^{t_1} \langle \dot{y}_0(t), L_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{y}_0(t)\rangle dt. \tag{3.68}$$

Se tiene que,

$$\liminf_{q\to\infty} \frac{\mathcal{E}(x_q)}{d_q^2} \geq \tfrac{1}{2}\int_{t_0}^{t_1} \langle \dot{y}_0(t), L_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{y}_0(t)\rangle dt. \tag{3.69}$$

En efecto, por el Lema 3.106, podemos escoger $S \subset T$ medible tal que $\dot{x}_q(t) \to \dot{x}_0(t)$ uniformemente en S. Además, para toda $t \in S$ y $q \in \mathbf{N}$,

$$\frac{1}{d_q^2} E(t, x_q(t), \dot{x}_0(t), \dot{x}_q(t)) = \tfrac{1}{2}\langle \dot{y}_q(t), R_q(t)\dot{y}_q(t)\rangle$$

donde

$$R_q(t) := 2\int_0^1 (1-\lambda)L_{\dot{x}\dot{x}}(t, x_q(t), \dot{x}_0(t) + \lambda[\dot{x}_q(t) - \dot{x}_0(t)])d\lambda.$$

Claramente,

$$\lim_{q\to\infty} R_q(t) = R_0(t) := L_{\dot{x}\dot{x}}(\tilde{x}_0(t)) \quad \text{uniformemente en } S.$$

Por el Teorema 3.111(ii), $R_0(t) \geq 0$ $(t \in S)$. Además, por el Teorema 3.111(iv), y el Lema 3.107, hay una subsucesión de $\{x_q\}$, nuevamente denotada por $\{x_q\}$, tal que

$$\liminf_{q\to\infty} \frac{\mathcal{E}(x_q)}{d_q^2} \geq \tfrac{1}{2}\int_S \langle \dot{y}_0(t), L_{\dot{x}\dot{x}}(\dot{x}_0(t))\dot{y}_0(t)\rangle dt.$$

Puesto que S se puede escoger de tal forma que difiera de T por un conjunto de medida arbitrariamente pequeña, y la función

$$t \mapsto \langle \dot{y}_0(t), L_{\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{y}_0(t)\rangle$$

pertenece a $L^1(T;\mathbf{R})$, esta desigualdad se cumple cuando $S = T$ y esto establece (3.69). Entonces, por (3.65), (3.67), (3.68) y (3.69),

$$\tfrac{1}{2}I''(x_0, y_0) \leq \lim_{q\to\infty} \frac{\mathcal{K}(x_q)}{d_q^2} + \liminf_{q\to\infty} \frac{\mathcal{E}(x_q)}{d_q^2} = \liminf_{q\to\infty} \frac{I(x_q) - I(x_0)}{d_q^2} \leq 0.$$

Adicionalmente, si $y_0 = 0$, entonces

$$\lim_{q\to\infty} \frac{\mathcal{K}(x_q)}{d_q^2} = 0$$

y en consecuencia, por el Teorema 3.111(v),

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

contradiciendo la positividad de h. ∎

El Corolario 3.112 es una consecuencia inmediata del Teorema 3.111. Es importante notar las similitudes del Corolario 3.112 y el Corolario 3.96.

3.112 Corolario: *Sea x_0 un arco admisible C^1 a pedazos en T. Supongamos que existe $\epsilon > 0$ y una constante $c \in \mathbf{R}^n$ tal que lo siguiente se satisface:*

(i) $L_{\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^{t} L_x(s, x_0(s), \dot{x}_0(s))ds + c^*$ $(t \in T)$.

(ii) $L_{\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

(iii) $E(t, x, \dot{x}, u) \geq 0$ *para* $(t, x, \dot{x}, u)$ *con* $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$.

(iv) $I''(x_0, y) > 0$ *para toda* $y \in \mathcal{Y}$, $y \neq 0$.

Entonces, para alguna $\mu, \nu > 0$ y para todo arco admisible x con $\|x - x_0\|_0 < \nu$, se tiene que

$$I(x) \geq I(x_0) + \mu D(x - x_0).$$

En particular, x_0 es un mínimo fuerte estricto de (P).

Demostración: Por el Lema 3.106, el Corolario 3.112(ii) y (iii), disminuyendo $\epsilon > 0$ si es necesario, existe $h > 0$ tal que

$$E(t, x, \dot{x}, u) \geq hV(u - \dot{x})$$

para $(t, x, \dot{x}, u)$ con $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$. Por lo tanto, $E(t, x(t), \dot{x}_0(t), \dot{x}(t)) \geq 0$ (c.s. en T) y

$$\int_{t_0}^{t_1} E(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq hD(x - x_0)$$

siempre que x sea admisible y $\|x - x_0\|_0 < \epsilon$. La conclusión del corolario se sigue del Teorema 3.111. ∎

3.113 Observación: *Como en la Observación 3.110, obsérvese que en el Teorema 3.111 y en el Corolario 3.112, si x_0 satisface todas sus condiciones, entonces I alcanza un mínimo fuerte estricto en x_0 en un conjunto más amplio de arcos admisibles que el de los C^1 a pedazos, específicamente, I alcanza un mínimo fuerte estricto en x_0 en el conjunto de trayectorias admisibles absolutamente continuas.*

23. Una demostración de suficiencia para problemas isoperimétricos: mínimos débiles

En esta sección el Teorema 3.114 nos entrega nuevas condiciones suficientes para un mínimo débil de un problema que tiene restricciones isoperimétricas con igualdades. Puesto que un problema isoperimétrico, en general, no se puede transformar en un problema de Lagrange como el problema (P) sin restricciones como el que hemos estudiado mayormente en este trabajo, el Teorema 3.114 amplia el rango de aplicabilidad de la teoría del cálculo de variaciones. Nuevamente, como en la Sección 21, el Teorema 3.114 detecta soluciones *singulares* y los mínimos débiles no solo son estrictos sino que la desviación entre costos admisibles y costos óptimos está estimada por la funcional D cuyo papel es muy similar al cuadrado de la norma del espacio de Banach $L^1(T;\mathbf{R}^n)$.

Supongamos que tenemos dados un intervalo $T := [t_0, t_1]$ en $\mathbf{R}$, dos puntos fijos ξ_0, ξ_1 en $\mathbf{R}^n$, funciones $L, L_1, \ldots, L_r$ que mapean $T \times \mathbf{R}^n \times \mathbf{R}^n$ a $\mathbf{R}$, y denotemos por $\mathcal{X}$ al espacio vectorial de las funciones absolutamente continuas que mapean T a $\mathbf{R}^n$.

En esta sección asumiremos que las funciones $L, L_1, \ldots, L_r$ son de clase C^2 con respecto a x y a $\dot{x}$ en $T \times \mathbf{R}^n \times \mathbf{R}^n$.

Consideremos el problema *isoperimétrico* (P) de minimizar la funcional

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt$$

sobre todas las $x \in \mathcal{X}$ con $\dot{x} \in L^\infty(T;\mathbf{R}^n)$ que satisfacen las restricciones

$$\begin{cases} L(t, x(t), \dot{x}(t)), L_i(t, x(t), \dot{x}(t)) \text{ son integrables en } T \ (i = 1, \ldots, r). \\ x(t_0) = \xi_0, \ x(t_1) = \xi_1. \\ I_i(x) := \int_{t_0}^{t_1} L_i(t, x(t), \dot{x}(t))dt = 0 \ (i = 1, \ldots, r). \end{cases}$$

A los elementos de $\mathcal{X}$ les llamaremos *arcos* o *trayectorias*, y una trayectoria x es *admisible* si $\dot{x} \in L^\infty(T;\mathbf{R}^n)$ y ésta satisface las restricciones. Un arco x *resuelve* (P), si x es admisible e $I(x) \leq I(y)$ para todos los arcos admisibles y. Un arco x es un *mínimo débil* de (P), si x es admisible y para alguna $\epsilon > 0$, $I(x) \leq I(y)$ para todos los arcos admisibles y que satisfacen $\|y - x\|_1 < \epsilon$.

Con el propósito de establecer el primer teorema de suficiencia de esta sección, introduzcamos las siguientes definiciones.

• Dadas $x, x_0 \in \mathcal{X}$, definimos

$$\tilde{x}(t) := (t, x(t), \dot{x}(t)) \quad \text{y} \quad \tilde{x}_0(t) := (t, x_0(t), \dot{x}_0(t)) \quad \text{(c.s. en } T).$$

• Dados r números reales $\lambda_1, \ldots, \lambda_r$ consideremos la funcional I_0 definida para todo arco admisible x por

$$I_0(x) := I(x) + \sum_1^r \lambda_i I_i(x) = \int_{t_0}^{t_1} L_0(\tilde{x}(t))dt,$$

donde $L_0: T \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ está dada por

$$L_0(t, x, \dot{x}) := L(t, x, \dot{x}) + \sum_1^r \lambda_i L_i(t, x, \dot{x}).$$

• Decimos que un arco C^1 a pedazos x_0 es *no singular* si $|L_{0\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t))| \neq 0$ para toda $t \in T$.

• Para cualquier $x \in \mathcal{X}$ con $\dot{x} \in L^\infty(T;\mathbf{R}^n)$, y cualquier $y \in \mathcal{X}$, la *primera variación* de I_i $(i = 0, 1, \ldots, r)$ a lo largo de x en la dirección y está dada por

$$I_i'(x, y) := \int_{t_0}^{t_1} \{L_{ix}(\tilde{x}(t))y(t) + L_{i\dot{x}}(\tilde{x}(t))\dot{y}(t)\}dt.$$

Además, si $y \in \mathcal{X}$ con $\dot{y} \in L^2(T;\mathbf{R}^n)$, definimos la *segunda variación* de I_0 a lo largo de x en la dirección y por

$$I_0''(x, y) := \int_{t_0}^{t_1} 2\omega_0(t, y(t), \dot{y}(t))dt$$

donde para toda $(t, y, \dot{y}) \in T \times \mathbf{R}^n \times \mathbf{R}^n$,

$$2\omega_0(t, y, \dot{y}) := \langle y, L_{0xx}(\tilde{x}(t))y\rangle + 2\langle y, L_{0x\dot{x}}(\tilde{x}(t))\dot{y}\rangle + \langle \dot{y}, L_{0\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}\rangle.$$

- Dada $x \in \mathcal{X}$, definamos a $\mathcal{Y}(x)$ como el conjunto de todas las $y \in \mathcal{X}$ con $\dot{y} \in L^2(T; \mathbf{R}^n)$ que satisfacen

$$y(t_0) = y(t_1) = 0 \quad \text{e} \quad I_i'(x, y) = 0 \ (i = 1, \ldots, r).$$

- Las *funciones exceso de Weierstrass* de L_i $(i = 0, 1, \ldots, r)$, $E_i: T \times \mathbf{R}^{3n} \to \mathbf{R}$, están dadas por

$$E_i(t, x, \dot{x}, u) := L_i(t, x, u) - L_i(t, x, \dot{x}) - L_{i\dot{x}}(t, x, \dot{x})(u - \dot{x}).$$

- Definimos las funciones $V: \mathbf{R}^n \to \mathbf{R}$, $D: \mathcal{X} \to \mathbf{R}$ por

$$V(a) := (1 + |a|^2)^{1/2} - 1 \quad \text{y} \quad D(x) := V(x(t_0)) + \int_{t_0}^{t_1} V(\dot{x}(t))dt.$$

3.114 Teorema: *Sea x_0 un arco admisible. Supongamos que existen dos números positivos h, ϵ, una constante $c \in \mathbf{R}^n$, y multiplicadores $\lambda_1, \ldots, \lambda_r$ tales que si*

$$L_0(t, x, \dot{x}) := L(t, x, \dot{x}) + \sum_1^r \lambda_i L_i(t, x, \dot{x}),$$

lo siguiente se satisface:

(i) $L_{0\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^t L_{0x}(s, x_0(s), \dot{x}_0(s))ds + c^*$ *(c.s. en T).*

(ii) $L_{0\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) \geq 0$ *(c.s. en T).*

(iii) $I_0''(x_0, y) > 0$ *para toda $y \in \mathcal{Y}(x_0)$, $y \neq 0$.*

(iv) $\int_{t_0}^{t_1} E_0(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq hD(x - x_0)$ *siempre que x sea admisible y $\|x - x_0\|_1 < \epsilon$.*

(v) $\int_{t_0}^{t_1} E_0(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq h|\int_{t_0}^{t_1} E_i(t, x(t), \dot{x}_0(t), \dot{x}(t))dt|$ $(i = 1, \ldots, r)$ *siempre que x sea admisible y $\|x - x_0\|_1 < \epsilon$.*

Entonces, para alguna $\mu, \nu > 0$ y para todo arco admisible x con $\|x - x_0\|_1 < \nu$, se tiene que

$$I(x) \geq I(x_0) + \mu D(x - x_0).$$

En particular, x_0 es un mínimo estricto débil de (P).

Demostración: La demostración se hará por contraposición, esto es, vamos a suponer que para toda $\mu, \nu > 0$, existe un arco admisible x tal que

$$\|x - x_0\|_1 < \nu \quad \text{e} \quad I(x) < I(x_0) + \mu D(x - x_0). \tag{3.70}$$

También, supondremos que (i), (ii), (iv) y (v) del Teorema 3.114 son satisfechas, y vamos a obtener la negación de la condición (iii) del Teorema 3.114. Primero que todo, obsérvese que como $I(x) = I_0(x)$ para toda x admisible, entonces (3.70) implica que para toda $\mu, \nu > 0$, existe x admisible con $\|x - x_0\|_1 < \nu$ e

$$I_0(x) < I_0(x_0) + \mu D(x - x_0). \tag{3.71}$$

Ahora notemos que para toda x admisible y para toda $i = 0, 1, \ldots, r$,

$$I_i(x) = I_i(x_0) + I_i'(x_0, x - x_0) + \mathcal{K}_i(x) + \mathcal{E}_i(x) \tag{3.72}$$

donde

$$\mathcal{E}_i(x) := \int_{t_0}^{t_1} E_i(t, x(t), \dot{x}_0(t), \dot{x}(t))dt,$$

$$\mathcal{K}_i(x) := \int_{t_0}^{t_1} \{M_i(t, x(t)) + \langle \dot{x}(t) - \dot{x}_0(t), N_i(t, x(t)) \rangle\}dt,$$

y las funciones M_i y N_i están dadas por

$$M_i(t,x) := L_i(t,x,\dot x_0(t)) - L_i(\tilde x_0(t)) - L_{ix}(\tilde x_0(t))(x - x_0(t)),$$

$$N_i(t,x) := L^*_{i\dot x}(t,x,\dot x_0(t)) - L^*_{i\dot x}(\tilde x_0(t)).$$

Tenemos que para toda $i = 0,1,\dots,r$,

$$M_i(t,x) = \tfrac{1}{2}\langle x - x_0(t), P_i(t,x)(x - x_0(t))\rangle, \quad N_i(t,x) = Q_i(t,x)(x - x_0(t)),$$

donde

$$P_i(t,x) := 2\int_0^1 (1-\lambda)L_{ixx}(t, x_0(t) + \lambda(x - x_0(t)), \dot x_0(t))d\lambda,$$

$$Q_i(t,x) := \int_0^1 L_{i\dot x x}(t, x_0(t) + \lambda(x - x_0(t)), \dot x_0(t))d\lambda.$$

Ahora, por (3.71), para toda $q \in \mathbf{N}$ existe x_q admisible tal que

$$\|x_q - x_0\|_1 < \text{mín}\{\epsilon, 1/q\}, \quad I_0(x_q) - I_0(x_0) < \frac{1}{q}D(x_q - x_0). \tag{3.73}$$

Consecuentemente, por la segunda desigualdad en (3.73),

$$d_q := [2D(x_q - x_0)]^{1/2} > 0 \quad (q \in \mathbf{N}),$$

y por la primera desigualdad en (3.73),

$$\lim_{q\to\infty} D(x_q - x_0) = 0.$$

Para toda $t \in T$ y $q \in \mathbf{N}$, definamos

$$y_q(t) := \frac{x_q(t) - x_0(t)}{d_q}.$$

Por el Lema 3.106, existen una subsucesión de $\{x_q\}$, nuevamente denotada por $\{x_q\}$, y $y_0 \in \mathcal{X}$ con $\dot y_0 \in L^2(T;\mathbf{R}^n)$, tales que $y_q(t) \to y_0(t)$ uniformemente en T, y la sucesión $\{\dot y_q\}$ converge débilmente en $L^1(T;\mathbf{R}^n)$ a $\dot y_0$.

Probemos que $y_0 \in \mathcal{Y}(x_0)$, $y_0 \neq 0$ e $I_0''(x_0, y_0) \le 0$ con lo cual concluiríamos la demostración del teorema. El hecho de que $y_0(t_0) = y_0(t_1) = 0$ se sigue de la definición de y_q, la admisibilidad de x_q y puesto que $y_q(t) \to y_0(t)$ uniformemente en T.

Para toda $q \in \mathbf{N}$ e $i = 0,1,\dots,r$, tenemos que

$$\frac{\mathcal{K}_i(x_q)}{d_q^2} = \int_{t_0}^{t_1}\left\{\frac{M_i(t,x_q(t))}{d_q^2} + \left\langle \dot y_q(t), \frac{N_i(t,x_q(t))}{d_q}\right\rangle\right\}dt.$$

Nuevamente, por el Lema 3.106, para toda $i = 0,1,\dots,r$,

$$\lim_{q\to\infty}\frac{M_i(t,x_q(t))}{d_q^2} = \tfrac{1}{2}\langle y_0(t), L_{ixx}(\tilde x_0(t))y_0(t)\rangle,$$

$$\lim_{q\to\infty}\frac{N_i(t,x_q(t))}{d_q} = L_{i\dot x x}(\tilde x_0(t))y_0(t)$$

todos uniformemente en T y, puesto que $\{\dot y_q\}$ converge débil a $\dot y_0$ en $L^1(T;\mathbf{R}^n)$,

$$\tfrac{1}{2}I_0''(x_0,y_0) = \lim_{q\to\infty}\frac{\mathcal{K}_0(x_q)}{d_q^2} + \tfrac{1}{2}\int_{t_0}^{t_1}\langle \dot y_0(t), L_{0\dot x\dot x}(\tilde x_0(t))\dot y_0(t)\rangle dt. \tag{3.74}$$

Se tiene que,

$$\liminf_{q\to\infty} \frac{\mathcal{E}_0(x_q)}{d_q^2} \geq \tfrac{1}{2}\int_{t_0}^{t_1} \langle \dot{y}_0(t), L_{0\dot{x}\dot{x}}(\tilde{x}_0(t))\dot{y}_0(t)\rangle dt. \tag{3.75}$$

En efecto, para casi toda $t \in T$ y $q \in \mathbf{N}$,

$$\frac{1}{d_q^2} E_0(t, x_q(t), \dot{x}_0(t), \dot{x}_q(t)) = \tfrac{1}{2}\langle \dot{y}_q(t), R_q(t)\dot{y}_q(t)\rangle$$

donde

$$R_q(t) := 2\int_0^1 (1-\lambda)L_{0\dot{x}\dot{x}}(t, x_q(t), \dot{x}_0(t) + \lambda[\dot{x}_q(t) - \dot{x}_0(t)])d\lambda.$$

Claramente,

$$\lim_{q\to\infty} R_q(t) = R_0(t) := L_{0\dot{x}\dot{x}}(\tilde{x}_0(t)) \quad \text{uniformemente en } T.$$

Por el Teorema 3.114(ii), $R_0(t) \geq 0$ (c.s. en T), y de esta manera (3.75) se sigue del Lema 3.107. Además, por el Teorema 3.114(i), se sigue que $I_0'(x_0, y) = 0$ para toda $y \in \mathcal{X}$ con $y(t_0) = y(t_1) = 0$. Con esto en mente, (3.72), (3.73), (3.74) y (3.75),

$$\tfrac{1}{2}I_0''(x_0, y_0) \leq \lim_{q\to\infty} \frac{\mathcal{K}_0(x_q)}{d_q^2} + \liminf_{q\to\infty} \frac{\mathcal{E}_0(x_q)}{d_q^2} = \liminf_{q\to\infty} \frac{I_0(x_q) - I_0(x_0)}{d_q^2} \leq 0.$$

Adicionalmente, si $y_0 = 0$, entonces

$$\lim_{q\to\infty} \frac{\mathcal{K}_0(x_q)}{d_q^2} = 0$$

y en consecuencia, por el Teorema 3.114(iv),

$$\tfrac{1}{2}h \leq \liminf_{q\to\infty} \frac{\mathcal{E}_0(x_q)}{d_q^2} \leq 0,$$

contradiciendo la positividad de h.

El teorema estará demostrado si establecemos que

$$I_i'(x_0, y_0) = 0 \quad (i = 1, \ldots, r).$$

Para obtener este resultado, observemos primero que por el Lema 3.106,

$$\lim_{q\to\infty} \frac{\mathcal{K}_i(x_q)}{d_q} = \lim_{q\to\infty} \frac{\mathcal{K}_0(x_q)}{d_q} = 0 \quad (i = 1, \ldots, r). \tag{3.76}$$

Con esto en mente, por el Teorema 3.114(i), (3.72) y (3.73),

$$0 \geq \limsup_{q\to\infty} \frac{I_0(x_q) - I_0(x_0)}{d_q} = \limsup_{q\to\infty} \frac{\mathcal{E}_0(x_q)}{d_q}.$$

Por la primera desigualdad en (3.73) y el Teorema 3.114(iv), para toda $q \in \mathbf{N}$, $\mathcal{E}_0(x_q) \geq 0$, y por lo tanto

$$\lim_{q\to\infty} \frac{\mathcal{E}_0(x_q)}{d_q} = 0,$$

y en consecuencia, por el Teorema 3.114(v),

$$\lim_{q\to\infty} \frac{\mathcal{E}_i(x_q)}{d_q} = 0 \quad (i = 1, \ldots, r). \tag{3.77}$$

Puesto que $I_i(x_q) = I_i(x_0) = 0$ $(i = 1, \ldots, r)$, tenemos por (3.72), (3.76) y (3.77),

$$0 = \lim_{q\to\infty} I_i'(x_0, y_q) + \lim_{q\to\infty} \frac{\mathcal{K}_i(x_q)}{d_q} + \lim_{q\to\infty} \frac{\mathcal{E}_i(x_q)}{d_q} = \lim_{q\to\infty} I_i'(x_0, y_q),$$

para toda $i = 1, \ldots, r$. En virtud del Lema 3.106,

$$0 = \lim_{q\to\infty} I_i'(x_0, y_q) = I_i'(x_0, y_0) \quad (i = 1, \ldots, r). \blacksquare$$

El Corolario 3.115 da condiciones suficientes para extremos no singulares del problema (P) de esta sección. Cabe señalar que el arco óptimo propuesto no es necesariamente C^1 sino únicamente C^1 a pedazos en T, en contraste, con todos los teoremas de suficiencia clásicos que obtuvimos en la Sección 18.

3.115 Corolario: *Sea x_0 un arco admisible C^1 a pedazos en T. Supongamos que existen dos números positivos h, ϵ, una constante $c \in \mathbf{R}^n$, y multiplicadores $\lambda_1, \ldots, \lambda_r$ tales que si*

$$L_0(t, x, \dot{x}) := L(t, x, \dot{x}) + \sum_1^r \lambda_i L_i(t, x, \dot{x}),$$

lo siguiente se satisface:

(i) $L_{0\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^t L_{0x}(s, x_0(s), \dot{x}_0(s))ds + c^*$ $(t \in T)$.

(ii) $L_{0\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

(iii) $I_0''(x_0, y) > 0$ *para toda* $y \in \mathcal{Y}(x_0)$, $y \neq 0$.

(iv) $\int_{t_0}^{t_1} E_0(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq h|\int_{t_0}^{t_1} E_i(t, x(t), \dot{x}_0(t), \dot{x}(t))dt|$ $(i = 1, \ldots, r)$ *siempre que x sea admisible y $\|x - x_0\|_1 < \epsilon$.*

Entonces, para alguna $\mu, \nu > 0$ y para todo arco admisible x con $\|x - x_0\|_1 < \nu$, se tiene que

$$I(x) \geq I(x_0) + \mu D(x - x_0).$$

En particular, x_0 es un mínimo estricto débil de (P).

Demostración: Por el Corolario 3.115(ii), podemos disminuir $\epsilon > 0$ tal que

$$L_{0\dot{x}\dot{x}}(t, x, \dot{x}) > 0 \quad ((t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)).$$

De esta manera, si $(t, x, \dot{x}, u)$ es tal que $(t, x, \dot{x})$ y (t, x, u) pertenecen a $\mathcal{T}_1(x_0; \epsilon)$, entonces,

$$E_0(t, x, \dot{x}, u) = \int_0^1 (1 - \lambda)\langle u - \dot{x}, L_{0\dot{x}\dot{x}}(t, x, \dot{x} + \lambda[u - \dot{x}])(u - \dot{x})\rangle d\lambda \geq 0.$$

Por el Lema 3.105, disminuyendo $\epsilon > 0$ si es necesario, tenemos que existe $h_0 > 0$ tal que

$$E_0(t, x, \dot{x}, u) \geq h_0 V(u - \dot{x})$$

para $(t, x, \dot{x}, u)$ con $(t, x, \dot{x})$ y (t, x, u) en $\mathcal{T}_1(x_0; \epsilon)$. Si es necesario, disminuimos al número positivo h dado en la condición (iv) del Corolario 3.115 de tal forma que h sea igual que h_0 y obtenemos que

$$\int_{t_0}^{t_1} E_0(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq hD(x - x_0)$$

siempre que x sea admisible y $\|x - x_0\|_1 < \epsilon$. La conclusión del corolario se sigue del Teorema 3.114. $\blacksquare$

24. Una demostración de suficiencia para problemas isoperimétricos: mínimos fuertes

En esta sección el Teorema 3.116 nos entrega nuevas condiciones suficientes para un mínimo fuerte estricto del problema isoperimétrico (P). Nuevamente, una de las propiedades más importantes del Teorema 3.116 es que éste detecta soluciones *singulares*.

Supongamos que tenemos dados un intervalo $T := [t_0, t_1]$ en $\mathbf{R}$, dos puntos fijos ξ_0, ξ_1 en $\mathbf{R}^n$, funciones $L, L_1, \ldots, L_r$ que mapean $T \times \mathbf{R}^n \times \mathbf{R}^n$ a $\mathbf{R}$, y denotemos por $\mathcal{X}$ al espacio vectorial de las funciones absolutamente continuas que mapean T a $\mathbf{R}^n$.

En esta sección asumiremos que las funciones $L, L_1, \ldots, L_r$ son de clase C^2 con respecto a x y a $\dot{x}$ en $T \times \mathbf{R}^n \times \mathbf{R}^n$.

Consideremos el problema *isoperimétrico* (P) de minimizar la funcional

$$I(x) := \int_{t_0}^{t_1} L(t, x(t), \dot{x}(t))dt$$

sobre todas las $x \in \mathcal{X}$ que satisfacen las restricciones

$$\begin{cases} L(t, x(t), \dot{x}(t)), L_i(t, x(t), \dot{x}(t)) \text{ son integrables en } T \ (i = 1, \ldots, r). \\ x(t_0) = \xi_0, \ x(t_1) = \xi_1. \\ I_i(x) := \int_{t_0}^{t_1} L_i(t, x(t), \dot{x}(t))dt = 0 \ (i = 1, \ldots, r). \end{cases}$$

A los elementos de $\mathcal{X}$ les llamaremos *arcos* o *trayectorias*, y una trayectoria x es *admisible* si ésta satisface las restricciones. Un arco admisible x *resuelve* (P), si $I(x) \leq I(y)$ para todos los arcos admisibles y. Un arco admisible x es un *mínimo fuerte* de (P), si para alguna $\epsilon > 0$, $I(x) \leq I(y)$ para todos los arcos admisibles y que satisfacen $\|y - x\|_0 < \epsilon$.

Con el propósito de establecer el primer teorema de suficiencia de esta sección, introduzcamos las siguientes definiciones.

• Dadas $x, x_0 \in \mathcal{X}$, definimos

$$\tilde{x}(t) := (t, x(t), \dot{x}(t)) \quad \text{y} \quad \tilde{x}_0(t) := (t, x_0(t), \dot{x}_0(t)) \quad \text{(c.s. en } T).$$

• Dados r números reales $\lambda_1, \ldots, \lambda_r$ consideremos la funcional I_0 definida para todo arco admisible x por

$$I_0(x) := I(x) + \sum_1^r \lambda_i I_i(x) = \int_{t_0}^{t_1} L_0(\tilde{x}(t))dt,$$

donde $L_0 \colon T \times \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ está dada por

$$L_0(t, x, \dot{x}) := L(t, x, \dot{x}) + \sum_1^r \lambda_i L_i(t, x, \dot{x}).$$

• Para cualquier $x \in \mathcal{X}$, y cualquier $y \in \mathcal{X}$, la *primera variación* de I_i $(i = 0, 1, \ldots, r)$ a lo largo de x en la dirección y está dada por

$$I_i'(x, y) := \int_{t_0}^{t_1} \{L_{ix}(\tilde{x}(t))y(t) + L_{i\dot{x}}(\tilde{x}(t))\dot{y}(t)\}dt.$$

Además, si $y \in \mathcal{X}$ con $\dot{y} \in L^2(T; \mathbf{R}^n)$, definimos la *segunda variación* de I_0 a lo largo de x en la dirección y por

$$I_0''(x, y) := \int_{t_0}^{t_1} 2\omega_0(t, y(t), \dot{y}(t))dt$$

donde para toda $(t, y, \dot{y}) \in T \times \mathbf{R}^n \times \mathbf{R}^n$,

$$2\omega_0(t, y, \dot{y}) := \langle y, L_{0xx}(\tilde{x}(t))y\rangle + 2\langle y, L_{0x\dot{x}}(\tilde{x}(t))\dot{y}\rangle + \langle \dot{y}, L_{0\dot{x}\dot{x}}(\tilde{x}(t))\dot{y}\rangle.$$

- Dada $x \in \mathcal{X}$, definamos a $\mathcal{Y}(x)$ como el conjunto de todas las $y \in \mathcal{X}$ con $\dot{y} \in L^2(T; \mathbf{R}^n)$ que satisfacen

$$y(t_0) = y(t_1) = 0 \quad \text{e} \quad I_i'(x, y) = 0 \ (i = 1, \dots, r).$$

- Las *funciones exceso de Weierstrass* de L_i $(i = 0, 1, \dots, r)$, $E_i: T \times \mathbf{R}^{3n} \to \mathbf{R}$, están dadas por

$$E_i(t, x, \dot{x}, u) := L_i(t, x, u) - L_i(t, x, \dot{x}) - L_{i\dot{x}}(t, x, \dot{x})(u - \dot{x}).$$

- Definimos las funciones $V: \mathbf{R}^n \to \mathbf{R}$, $D: \mathcal{X} \to \mathbf{R}$ por

$$V(a) := (1 + |a|^2)^{1/2} - 1 \quad \text{y} \quad D(x) := V(x(t_0)) + \int_{t_0}^{t_1} V(\dot{x}(t))dt.$$

3.116 Teorema: *Sea x_0 un arco admisible con $\dot{x}_0 \in L^\infty(T; \mathbf{R}^n)$. Supongamos que existen dos números positivos h, ϵ, una constante $c \in \mathbf{R}^n$, y multiplicadores $\lambda_1, \dots, \lambda_r$ tales que si*

$$L_0(t, x, \dot{x}) := L(t, x, \dot{x}) + \sum_1^r \lambda_i L_i(t, x, \dot{x}),$$

lo siguiente se satisface:

(i) $L_{0\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^t L_{0x}(s, x_0(s), \dot{x}_0(s))ds + c^*$ *(c.s. en T).*

(ii) $L_{0\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) \geq 0$ *(c.s. en T).*

(iii) $I_0''(x_0, y) > 0$ *para toda $y \in \mathcal{Y}(x_0)$, $y \neq 0$.*

(iv) $E_0(t, x(t), \dot{x}_0(t), \dot{x}(t)) \geq 0$ *siempre que x sea admisible y $\|x - x_0\|_0 < \epsilon$.*

(v) $\int_{t_0}^{t_1} E_0(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq hD(x - x_0)$ *siempre que x sea admisible y $\|x - x_0\|_0 < \epsilon$.*

(vi) $\int_{t_0}^{t_1} E_0(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq h|\int_{t_0}^{t_1} E_i(t, x(t), \dot{x}_0(t), \dot{x}(t))dt|$ $(i = 1, \dots, r)$ *siempre que x sea admisible y $\|x - x_0\|_0 < \epsilon$.*

Entonces, para alguna $\mu, \nu > 0$ y para todo arco admisible x con $\|x - x_0\|_0 < \nu$, se tiene que

$$I(x) \geq I(x_0) + \mu D(x - x_0).$$

En particular, x_0 es un mínimo fuerte estricto de (P).

Demostración: La demostración se hará por contraposición, esto es, vamos a suponer que para toda $\mu, \nu > 0$, existe un arco admisible x tal que

$$\|x - x_0\|_0 < \nu \quad \text{e} \quad I(x) < I(x_0) + \mu D(x - x_0). \tag{3.78}$$

También, supondremos que (i), (ii), (iv), (v) y (vi) del Teorema 3.116 son satisfechas, y vamos a obtener la negación de la condición (iii) del Teorema 3.116. Primero que todo, obsérvese que como $I(x) = I_0(x)$ para toda x admisible, entonces (3.78) implica que para toda $\mu, \nu > 0$, existe x admisible con $\|x - x_0\|_0 < \nu$ e

$$I_0(x) < I_0(x_0) + \mu D(x - x_0). \tag{3.79}$$

Ahora, notemos que para toda x admisible y para toda $i = 0, 1, \dots, r$,

$$I_i(x) = I_i(x_0) + I_i'(x_0, x - x_0) + \mathcal{K}_i(x) + \mathcal{E}_i(x) \tag{3.80}$$

donde

$$\mathcal{E}_i(x) := \int_{t_0}^{t_1} E_i(t, x(t), \dot{x}_0(t), \dot{x}(t))dt,$$

$$\mathcal{K}_i(x) := \int_{t_0}^{t_1} \{M_i(t, x(t)) + \langle \dot{x}(t) - \dot{x}_0(t), N_i(t, x(t))\rangle\}dt,$$

y las funciones M_i y N_i están dadas por

$$M_i(t, x) := L_i(t, x, \dot{x}_0(t)) - L_i(\tilde{x}_0(t)) - L_{ix}(\tilde{x}_0(t))(x - x_0(t)),$$

$$N_i(t, x) := L^*_{i\dot{x}}(t, x, \dot{x}_0(t)) - L^*_{i\dot{x}}(\tilde{x}_0(t)).$$

Tenemos que para toda $i = 0, 1, \ldots, r$,

$$M_i(t, x) = \tfrac{1}{2}\langle x - x_0(t), P_i(t, x)(x - x_0(t))\rangle, \quad N_i(t, x) = Q_i(t, x)(x - x_0(t)),$$

donde

$$P_i(t, x) := 2\int_0^1 (1 - \lambda)L_{ixx}(t, x_0(t) + \lambda(x - x_0(t)), \dot{x}_0(t))d\lambda,$$

$$Q_i(t, x) := \int_0^1 L_{i\dot{x}x}(t, x_0(t) + \lambda(x - x_0(t)), \dot{x}_0(t))d\lambda.$$

Como se hizo en la Sección 22, escojamos $c_1 > 0$ tal que para toda x admisible con $\|x - x_0\|_0 < 1$,

$$|\mathcal{K}_0(x)| \leq c_1\|x - x_0\|_0[1 + D(x - x_0)]. \tag{3.81}$$

Ahora, por (3.79), para toda $q \in \mathbf{N}$ existe x_q admisible tal que

$$\|x_q - x_0\|_0 < \text{mín}\{\epsilon, 1/q\}, \quad I_0(x_q) - I_0(x_0) < \frac{1}{q}D(x_q - x_0). \tag{3.82}$$

Consecuentemente, por la segunda desigualdad en (3.82),

$$d_q := [2D(x_q - x_0)]^{1/2} > 0 \quad (q \in \mathbf{N}).$$

Por el Teorema 3.116(i) se sigue que $I_0'(x_0, y) = 0$ para toda $y \in \mathcal{X}$ con $y(t_0) = y(t_1) = 0$. Con esto en mente, por (3.80), el Teorema 3.116(v), (3.81) y la primera desigualdad de (3.82),

$$I_0(x_q) - I_0(x_0) = \mathcal{K}_0(x_q) + \mathcal{E}_0(x_q) \geq -c_1\|x_q - x_0\|_0 + D(x_q - x_0)(h - c_1\|x_q - x_0\|_0).$$

Por (3.82), obtenemos que

$$D(x_q - x_0)\left(h - \frac{1}{q} - \frac{c_1}{q}\right) < \frac{c_1}{q}$$

y en consecuencia,

$$\lim_{q\to\infty} D(x_q - x_0) = 0.$$

Para toda $t \in T$ y $q \in \mathbf{N}$, definamos

$$y_q(t) := \frac{x_q(t) - x_0(t)}{d_q}.$$

Por el Lema 3.106, existen una subsucesión de $\{x_q\}$, nuevamente denotada por $\{x_q\}$, y $y_0 \in \mathcal{X}$ con $\dot{y}_0 \in L^2(T; \mathbf{R}^n)$, tales que $y_q(t) \to y_0(t)$ uniformemente en T, y la sucesión $\{\dot{y}_q\}$ converge débilmente en $L^1(T; \mathbf{R}^n)$ a $\dot{y}_0$.

Probemos que $y_0 \in \mathcal{Y}(x_0)$, $y_0 \neq 0$ e $I_0''(x_0, y_0) \leq 0$ con lo cual concluiríamos la demostración del teorema.

El hecho de que $y_0(t_0) = y_0(t_1) = 0$ se sigue de la definición de y_q, la admisibilidad de x_q y puesto que $y_q(t) \to y_0(t)$ uniformemente en T.

Para toda $q \in \mathbf{N}$ e $i = 0, 1, \ldots, r$, tenemos que

$$\frac{\mathcal{K}_i(x_q)}{d_q^2} = \int_{t_0}^{t_1} \left\{ \frac{M_i(t, x_q(t))}{d_q^2} + \left\langle \dot{y}_q(t), \frac{N_i(t, x_q(t))}{d_q} \right\rangle \right\} dt.$$

Nuevamente, por el Lema 3.106, para toda $i = 0, 1, \ldots, r$,

$$\lim_{q \to \infty} \frac{M_i(t, x_q(t))}{d_q^2} = \tfrac{1}{2} \langle y_0(t), L_{ixx}(\tilde{x}_0(t)) y_0(t) \rangle,$$

$$\lim_{q \to \infty} \frac{N_i(t, x_q(t))}{d_q} = L_{i\dot{x}x}(\tilde{x}_0(t)) y_0(t)$$

todos uniformemente en T y, puesto que $\{\dot{y}_q\}$ converge débil a $\dot{y}_0$ en $L^1(T; \mathbf{R}^n)$,

$$\tfrac{1}{2} I_0''(x_0, y_0) = \lim_{q \to \infty} \frac{\mathcal{K}_0(x_q)}{d_q^2} + \tfrac{1}{2} \int_{t_0}^{t_1} \langle \dot{y}_0(t), L_{0\dot{x}\dot{x}}(\tilde{x}_0(t)) \dot{y}_0(t) \rangle dt. \tag{3.83}$$

Tenemos que,

$$\liminf_{q \to \infty} \frac{\mathcal{E}_0(x_q)}{d_q^2} \geq \tfrac{1}{2} \int_{t_0}^{t_1} \langle \dot{y}_0(t), L_{0\dot{x}\dot{x}}(\tilde{x}_0(t)) \dot{y}_0(t) \rangle dt. \tag{3.84}$$

En efecto, por el Lema 3.106, podemos escoger $S \subset T$ medible tal que $\dot{x}_q(t) \to \dot{x}_0(t)$ uniformemente en S. Además, para toda $t \in S$ y $q \in \mathbf{N}$,

$$\frac{1}{d_q^2} E_0(t, x_q(t), \dot{x}_0(t), \dot{x}_q(t)) = \tfrac{1}{2} \langle \dot{y}_q(t), R_q(t) \dot{y}_q(t) \rangle$$

donde

$$R_q(t) := 2 \int_0^1 (1 - \lambda) L_{0\dot{x}\dot{x}}(t, x_q(t), \dot{x}_0(t) + \lambda[\dot{x}_q(t) - \dot{x}_0(t)]) d\lambda.$$

Claramente,

$$\lim_{q \to \infty} R_q(t) = R_0(t) := L_{0\dot{x}\dot{x}}(\tilde{x}_0(t)) \quad \text{uniformemente en } S.$$

Por el Teorema 3.116(ii), $R_0(t) \geq 0$ $(t \in S)$. Además, por el Teorema 3.116(iv), y el Lema 3.107, hay una subsucesión de $\{x_q\}$, nuevamente denotada por $\{x_q\}$, tal que

$$\liminf_{q \to \infty} \frac{\mathcal{E}_0(x_q)}{d_q^2} \geq \tfrac{1}{2} \int_S \langle \dot{y}_0(t), L_{0\dot{x}\dot{x}}(\dot{x}_0(t)) \dot{y}_0(t) \rangle dt.$$

Puesto que S se puede escoger de tal forma que difiera de T por un conjunto de medida arbitrariamente pequeña, y la función

$$t \mapsto \langle \dot{y}_0(t), L_{0\dot{x}\dot{x}}(\tilde{x}_0(t)) \dot{y}_0(t) \rangle$$

pertenece a $L^1(T; \mathbf{R})$, esta desigualdad se cumple cuando $S = T$ y esto establece (3.84). Entonces, por (3.80), (3.82), (3.83) y (3.84),

$$\tfrac{1}{2} I_0''(x_0, y_0) \leq \lim_{q \to \infty} \frac{\mathcal{K}_0(x_q)}{d_q^2} + \liminf_{q \to \infty} \frac{\mathcal{E}_0(x_q)}{d_q^2} = \liminf_{q \to \infty} \frac{I_0(x_q) - I_0(x_0)}{d_q^2} \leq 0.$$

Adicionalmente, si $y_0 = 0$, entonces

$$\lim_{q \to \infty} \frac{\mathcal{K}_0(x_q)}{d_q^2} = 0$$

y en consecuencia, por el Teorema 3.116(v),

$$\tfrac{1}{2}h \leq \liminf_{q\to\infty} \frac{\mathcal{E}_0(x_q)}{d_q^2} \leq 0,$$

contradiciendo la positividad de h.

El teorema estará demostrado si establecemos que

$$I_i'(x_0, y_0) = 0 \quad (i = 1, \ldots, r).$$

Para obtener este resultado, observemos primero que por el Lema 3.106,

$$\lim_{q\to\infty} \frac{\mathcal{K}_i(x_q)}{d_q} = \lim_{q\to\infty} \frac{\mathcal{K}_0(x_q)}{d_q} = 0 \quad (i = 1, \ldots, r). \tag{3.85}$$

Con esto en mente, por el Teorema 3.116(i), (3.80) y (3.82),

$$0 \geq \limsup_{q\to\infty} \frac{I_0(x_q) - I_0(x_0)}{d_q} = \limsup_{q\to\infty} \frac{\mathcal{E}_0(x_q)}{d_q}.$$

Por la primera desigualdad en (3.82) y el Teorema 3.116(iv), para toda $q \in \mathbf{N}$, $\mathcal{E}_0(x_q) \geq 0$, y por lo tanto

$$\lim_{q\to\infty} \frac{\mathcal{E}_0(x_q)}{d_q} = 0,$$

y en consecuencia, por el Teorema 3.116(vi),

$$\lim_{q\to\infty} \frac{\mathcal{E}_i(x_q)}{d_q} = 0 \quad (i = 1, \ldots, r). \tag{3.86}$$

Puesto que $I_i(x_q) = I_i(x_0) = 0$ $(i = 1, \ldots, r)$, tenemos por (3.80), (3.85) y (3.86),

$$0 = \lim_{q\to\infty} I_i'(x_0, y_q) + \lim_{q\to\infty} \frac{\mathcal{K}_i(x_q)}{d_q} + \lim_{q\to\infty} \frac{\mathcal{E}_i(x_q)}{d_q} = \lim_{q\to\infty} I_i'(x_0, y_q),$$

para toda $i = 1, \ldots, r$. En virtud del Lema 3.106,

$$0 = \lim_{q\to\infty} I_i'(x_0, y_q) = I_i'(x_0, y_0) \quad (i = 1, \ldots, r). \ \blacksquare$$

3.117 Corolario: *Sea x_0 un arco admisible C^1 a pedazos en T. Supongamos que existen dos números positivos h, ϵ, una constante $c \in \mathbf{R}^n$, y multiplicadores $\lambda_1, \ldots, \lambda_r$ tales que si*

$$L_0(t, x, \dot{x}) := L(t, x, \dot{x}) + \sum_1^r \lambda_i L_i(t, x, \dot{x}),$$

lo siguiente se satisface:

(i) $L_{0\dot{x}}(t, x_0(t), \dot{x}_0(t)) = \int_{t_0}^t L_{0x}(s, x_0(s), \dot{x}_0(s))ds + c^*$ $(t \in T)$.

(ii) $L_{0\dot{x}\dot{x}}(t, x_0(t), \dot{x}_0(t)) > 0$ $(t \in T)$.

(iii) $I_0''(x_0, y) > 0$ *para toda* $y \in \mathcal{Y}(x_0)$, $y \neq 0$.

(iv) $E_0(t, x, \dot{x}, u) \geq 0$ *para* $(t, x, \dot{x}, u)$ *con* $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$.

(v) $\int_{t_0}^{t_1} E_0(t, x(t), \dot{x}_0(t), \dot{x}(t))dt \geq h|\int_{t_0}^{t_1} E_i(t, x(t), \dot{x}_0(t), \dot{x}(t))dt|$ $(i = 1, \ldots, r)$ *siempre que x sea admisible y $\|x - x_0\|_0 < \epsilon$.*

Entonces, para alguna $\mu, \nu > 0$ y para todo arco admisible x con $\|x - x_0\|_0 < \nu$, se tiene que

$$I(x) \geq I(x_0) + \mu D(x - x_0).$$

En particular, x_0 es un mínimo fuerte estricto de (P).

Demostración: Por el Lema 3.105, el Corolario 3.117(ii) y (iv), disminuyendo $\epsilon > 0$ si es necesario, existe $h_0 > 0$ tal que

$$E_0(t, x, \dot{x}, u) \geq h_0 V(u - \dot{x})$$

para $(t, x, \dot{x}, u)$ con $(t, x, \dot{x}) \in \mathcal{T}_1(x_0; \epsilon)$. Si es necesario, disminuimos al número positivo h dado en la condición (v) del Corolario 3.117 de tal forma que h sea igual que h_0 y obtenemos que $E_0(t, x(t), \dot{x}_0(t), \dot{x}(t)) \geq 0$ (c.s. en T) y

$$\int_{t_0}^{t_1} E_0(t, x(t), \dot{x}_0(t), \dot{x}(t)) dt \geq hD(x - x_0)$$

siempre que x sea admisible y $\|x - x_0\|_0 < \epsilon$. La conclusión del corolario se sigue del Teorema 3.116. ∎

Bibliografía

[1] *Berlanga, R., & Rosenblueth, J. F.*, *Jacobi's Condition for Singular Extremals: An extended Notion of Conjugate Points*, Applied Mathematics Letters **15**, pp. 453-458, 2002.

[2] *Berlanga, R., & Rosenblueth, J. F.* (2004) *Extended conjugate points in the calculus of variations*, IMA Journal of Mathematical Control and Information **21**, pp. 159-173, 2004.

[3] *F. H. Clarke* and *V. M. Zeidan*, *Sufficiency and the Jacobi condition in the calculus of variations*, Canadian Journal of Mathematics, 38 (1986), pp. 1199–1209.

[4] *F. H. Clarke*, *Functional Analysis, Calculus of Variations and Optimal Control*, Springer-Verlag, London, 2013.

[5] *M. R. De Pinho* and *J. F. Rosenblueth*, *Mixed constraints in optimal control: an implicit function theorem approach*, IMA Journal of Mathematical Control and Information, 24 (2007), pp. 197–218

[6] *G. M. Ewing*, *Calculus of Variations with Applications*, Dover, New York, 1985.

[7] *M. R. Hestenes*, *Calculus of Variations and Optimal Control Theory*, John Wiley, New York, 1966.

[8] *A. D. Ioffe* and *V. M. Tihomirov*, *Theory of Extremal Problems*, Translated from the Russian by K. Makowski, Studies in Mathematics and its Applications **6**, North-Holland Publishing, Amsterdam, 1979.

[9] *P. D. Loewen*, *Second-order sufficiency criteria and local convexity for equivalent problems in the calculus of variations*, Journal of Mathematical Analysis and Applications, 146 (1990), pp. 512–522.

[10] *K. Malanowski*, *Sufficient optimality conditions for optimal control subject to state constraints*, SIAM Journal on Control and Optimization, 35 (1997), pp. 205–227.

[11] *K. Malanowski, H. Maurer* and *S. Pickenhain*, *Second order sufficient conditions for state- constrained optimal control problems*, Journal of Optimization Theory and Applications, 123 (2004), pp. 595–617.

[12] *H. Maurer*, *First and second order sufficient optimality conditions in mathematical programming and optimal control*, Mathematical Programming Study, 14 (1981), pp. 163–177.

[13] *H. Maurer* and *S. Pickenhain*, *Second order sufficient conditions for control problems with mixed control-state constraints*, Journal of Optimization Theory and Applications, 86 (1995), pp. 649–667.

[14] *H. Maurer* and *H. J. Oberle*, *Second order sufficient conditions for optimal control problems with free final time: the Riccati approach*, SIAM Journal on Control and Optimization, 41 (2002), pp. 380–403.

[15] *D. Q. Mayne*, *Sufficient conditions for a control to be a strong minimum*, Journal of Optimization Theory and Applications, 21 (1977), pp. 339–351.

[16] *E. J. McShane*, *Sufficient conditions for a weak relative minimum in the problem of Bolza*, Transactions of the American Mathematical Society, 52 (1942), pp. 344–379.

[17] *A. A. Milyutin* and *N. P. Osmolovskiĭ*, *Calculus of Variations and Optimal Control*, Translations of Mathematical Monographs **180**, American Mathematical Society, Providence, Rhode Island, 1998.

[18] *G. Stefani* and *P. L. Zezza*, *Optimality conditions for a constrained optimal control problem*, SIAM Journal on Control and Optimization, 34 (1996), pp. 635–659.

[19] *V. M. Zeidan*, *First and second order sufficient conditions for optimal control and the calculus of variations*, Applied Mathematics and Optimization 11 (1984), pp. 209–226.

[20] *V. M. Zeidan*, *Sufficiency conditions with minimal regularity assumptions*, Applied Mathematics and Optimization 20 (1989), pp. 19–31.

[21] *V. M. Zeidan*, *The Riccati equation for optimal control problems with mixed state-control constraints: necessity and sufficiency*, SIAM Journal on Control and Optimization 32 (1994), pp. 1297–1321.