

Capítulo 4

Optimización dinámica: métodos variacionales

4.1 Introducción y ejemplos

Los problemas tratados en los temas anteriores consistían en la optimización (maximización o minimización) de funciones reales de variables reales, que podían estar o no sujetas a restricciones representadas también por funciones reales de variables reales. El objetivo era buscar el punto o puntos que cumpliendo las restricciones impuestas al problema, nos proporcionaran el mejor valor (mayor o menor, dependiendo del objetivo del problema) de la función a optimizar. Para la mayoría de todos estos problemas la solución era puntual. Por ello, este tipo de optimización se suele clasificar como estática.

En este tema se enfocará la optimización que hemos denominado optimización dinámica. El objetivo de estos problemas también consiste en maximizar o minimizar cierto índice teniendo en cuenta un conjunto de restricciones que describe el sistema, pero a diferencia de la optimización estática, las soluciones de este problema están ahora formadas por trayectorias de puntos y no solamente por puntos particulares. Los objetivos utilizados como índices de rendimiento del sistema son en este caso funciones que dependen de otras funciones (funcionales) y las restricciones que describen los sistemas son en general ecuaciones diferenciales (problemas en tiempo continuo) y ecuaciones en diferencias (problemas en tiempo discreto).

Se Plantean por tanto, otro tipo de problemas, en el que la solución buscada es una curva o trayectoria, obtenida de entre todas las posibles en un sistema, que optimiza un determinado índice de coste.

A continuación se describen algunos ejemplos.

4.1.1 Problemas Geodésicos

Uno de los problemas que aparece de forma natural es el de encontrar el camino más corto entre dos puntos. Aunque en \mathbb{R}^3 este camino es la línea recta, en general no es posible tomar esta ruta debido por ejemplo a objetos naturales, entonces es necesario considerar el problema más complicado de encontrar la curva *geodésica* (es decir la que tiene menor longitud) entre todas aquellas restringidas a una “hipersuperficie” dada. En particular, por ejemplo podemos intentar caracterizar en \mathbb{R}^3 las geodésicas sobre la superficie de una esfera, de un cilindro o de un cono.

Geodésicas en el \mathbb{R}^n

Una curva que une dos puntos \mathbf{A} y \mathbf{B} en \mathbb{R}^n puede considerarse como el rango de una función vectorial $\mathbf{Y}(t) = (y_1(t), \dots, y_n(t))$, $t \in [0, 1]$ con componentes continuas en $[0, 1]$, de forma que $\mathbf{Y}(0) = \mathbf{A}$ e $\mathbf{Y}(1) = \mathbf{B}$ (figura 4.1). En particular, el segmento lineal que une ambos puntos, $\mathbf{Y}_0(t) = (1-t)\mathbf{A} + t\mathbf{B}$, $t \in [0, 1]$ es una de esas curvas. (ver figura 4.1)

Con esta generalidad, la curva puede no ser rectificable, es decir, la curva puede no tener longitud finita. Si exigimos que la función $\mathbf{Y}(t)$ tenga derivada continua en $(0, 1)$, entonces podremos contemplar la función

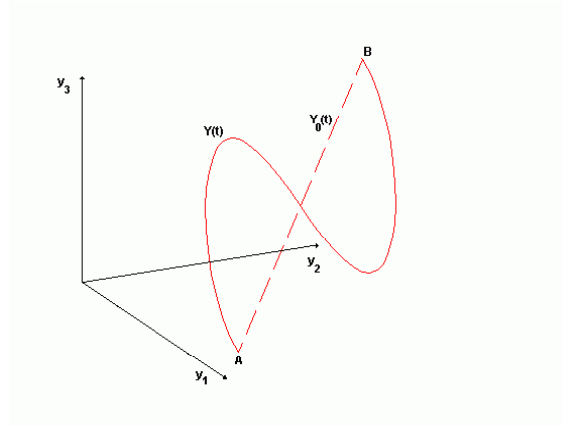


Figura 4.1: Curvas en \mathbb{R}^3 .

$\mathbf{Y}'(t) = (y'_1(t), \dots, y'_n(t))$ como la velocidad en t con módulo $\|\mathbf{Y}'(t)\|$ y por tanto la longitud de la curva debería ser la distancia recorrida durante el movimiento

$$L(\mathbf{Y}) = \int_0^1 \|\mathbf{Y}'(t)\| dt$$

Para garantizar la finitud de $L(\mathbf{Y})$, hay que exigir que cada componente de $\mathbf{Y}'(t)$ sea integrable (esta propiedad se obtiene fácilmente si cada componente es continuamente diferenciable en $[0, 1]$). De esta manera el problema consistirá en *minimizar* la longitud $\mathbf{Y}(t)$, $L(\mathbf{Y})$, sobre todas las funciones vectoriales $\mathbf{Y}(t)$ que cumplen las condiciones anteriores.

Puesto que $\mathbf{Y}_0(t)$ es una de estas curvas, con derivada $\mathbf{Y}'_0(t) = \mathbf{B} - \mathbf{A}$, sabemos que si existe la curva de longitud mínima L_{\min} , entonces

$$L_{\min} \leq \int_0^1 \|\mathbf{Y}'_0(t)\| dt = \|\mathbf{B} - \mathbf{A}\| \quad (4.1)$$

la distancia euclídea entre \mathbf{A} y \mathbf{B} .

Para comprobar la conjetura natural de que $L_{\min} = \|\mathbf{B} - \mathbf{A}\|$, solamente hay que comprobar la siguiente desigualdad para cualquier $\mathbf{Y}(t)$ admisible

$$\|\mathbf{B} - \mathbf{A}\| \leq L(\mathbf{Y}) = \int_0^1 \|\mathbf{Y}'(t)\| dt \quad (4.2)$$

que es trivial cuando $\mathbf{A} = \mathbf{B}$; cuando $\mathbf{A} \neq \mathbf{B}$ y utilizando el teorema fundamental del Cálculo

$$\mathbf{B} - \mathbf{A} = \mathbf{Y}(1) - \mathbf{Y}(0) = \int_0^1 \mathbf{Y}'(t) dt \quad (4.3)$$

de esta forma

$$\|\mathbf{B} - \mathbf{A}\|^2 = (\mathbf{B} - \mathbf{A}) \cdot (\mathbf{B} - \mathbf{A}) = (\mathbf{B} - \mathbf{A}) \cdot \int_0^1 \mathbf{Y}'(t) dt = \int_0^1 [(\mathbf{B} - \mathbf{A}) \cdot \mathbf{Y}'(t)] dt \quad (4.4)$$

utilizando la desigualdad de Cauchy

$$[(\mathbf{B} - \mathbf{A}) \cdot \mathbf{Y}'(t)] \leq \|\mathbf{B} - \mathbf{A}\| \cdot \|\mathbf{Y}'(t)\| \quad (4.5)$$

se obtiene

$$\int_0^1 [(\mathbf{B} - \mathbf{A}) \cdot \mathbf{Y}'(t)] dt \leq \int_0^1 \|\mathbf{B} - \mathbf{A}\| \cdot \|\mathbf{Y}'(t)\| dt \quad (4.6)$$

o equivalentemente

$$\|\mathbf{B} - \mathbf{A}\|^2 \leq \int_0^1 \|\mathbf{B} - \mathbf{A}\| \cdot \|\mathbf{Y}'(t)\| dt \quad (4.7)$$

y para $\mathbf{A} \neq \mathbf{B}$ se produce la desigualdad buscada dividiendo por $\|\mathbf{B} - \mathbf{A}\|$.

Geodésicas en la esfera

Para un avión es fundamental, para conseguir un ahorro de combustible, conocer cuál es la ruta más corta entre dos ciudades. La tierra es, en primera aproximación una esfera, por tanto el camino más corto entre dos puntos de su superficie no será la línea recta. Ahora, es necesario encontrar las curvas geodésicas sobre la superficie de una esfera.

Cada punto $\mathbf{Y} = (y_1, y_2, y_3) \in \mathbb{R}^3$ de la superficie de una esfera de radio R centrada en el origen (excepto los polos Norte y Sur) están determinados mediante coordenadas esféricas R , φ y θ mediante:

$$\mathbf{Y} = (R \cos \theta \sin \varphi, R \sin \theta \sin \varphi, R \cos \varphi) \quad (4.8)$$

para un único $\varphi \in (0, \pi)$ y $\theta \in [0, 2\pi)$ (figura 4.2).

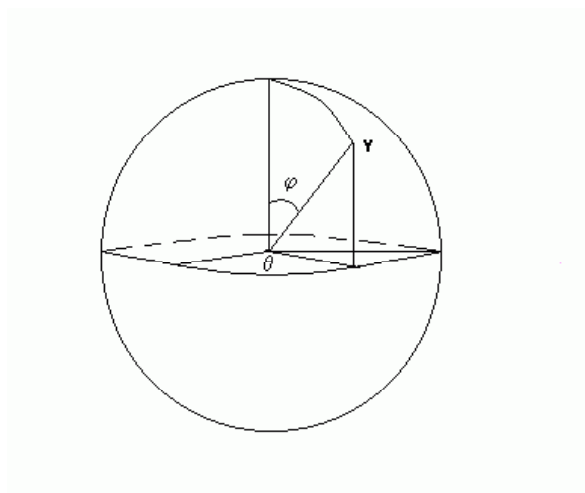


Figura 4.2: Coordenadas esféricas.

Dados dos puntos \mathbf{A} y \mathbf{B} sobre la esfera, supondremos que se eligen los ejes de forma que \mathbf{A} esté en el polo Norte ($\varphi = 0$), mientras que $\mathbf{B} \neq \mathbf{A}$ tiene las coordenadas esféricas $(R, 0, \varphi_1)$ con $\varphi_1 > 0$ (figura 4.3).

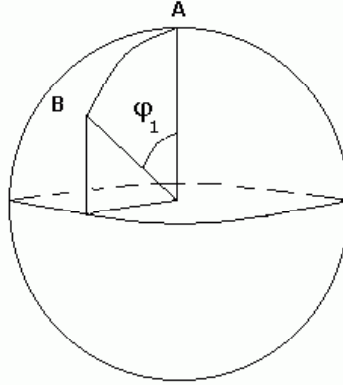


Figura 4.3: Geodésicas en la esfera.

Entonces cualquier curva que une **A** con **B** sobre la superficie de esta esfera está determinada por 4.8 mediante el par de funciones $(\theta(t), \varphi(t))$, $t \in [0, 1]$, con $\varphi(0) = 0$; $\theta(1) = 0$, $\varphi(1) = \varphi_1$. (Nota la consideración $\theta(1) = 0$ se hace por conveniencia). Como en el caso anterior queremos que las funciones utilizadas sean continuamente diferenciables en $[0, 1]$. La curva estará definida en coordenadas esféricas por

$$\mathbf{Y}(t) = R(\cos[\theta(t)] \sin[\varphi(t)], \sin[\theta(t)] \sin[\varphi(t)], \cos[\varphi(t)])$$

cuya derivada temporal es

$$\mathbf{Y}'(t) = R(-\theta' \sin \theta \sin \varphi + \varphi' \cos \theta \cos \varphi, \theta' \cos \theta \sin \varphi + \varphi' \sin \theta \cos \varphi, -\varphi' \sin \varphi)$$

(se ha prescindido de la variable temporal t por comodidad) y cuya longitud viene dada por

$$L(\mathbf{Y}) = \int_0^1 \|\mathbf{Y}'(t)\| dt = R \int_0^1 \sqrt{\theta'(t)^2 \sin^2 \varphi(t) + \varphi'(t)^2} dt$$

Si eliminamos el término $\theta'(t)^2 \sin^2 \varphi(t) \geq 0$ dentro de la raíz se obtiene

$$R \int_0^1 \sqrt{\theta'(t)^2 \sin^2 \varphi(t) + \varphi'(t)^2} dt \geq R \int_0^1 \sqrt{\varphi'(t)^2} dt = R\varphi(t)|_0^1$$

De manera que $L(\mathbf{Y}) \geq R\varphi_1$ por las propiedades de φ . La igualdad solamente se cumple si el término $\sin^2 \varphi(t) \theta'(t)^2 = 0$. Resolviendo

$$\sin^2 \varphi(t) \theta'(t)^2 = 0 \Leftrightarrow \begin{cases} \sin^2(\varphi(t)) = 0, \forall t \in [0, 1] \\ \theta'(t) = 0, \forall t \in [0, 1] \end{cases}$$

En el primer caso $\varphi(t) = 0$, con $t \in [0, 1]$, luego $\mathbf{A} = \mathbf{B}$, y en este caso no hay desplazamiento. Si, por el contrario, $\mathbf{A} \neq \mathbf{B}$, entonces $\theta'(t) = 0$, para $t \in [0, 1]$, de donde $\theta(t) = cte$, en todo el intervalo $[0, 1]$. Y por tanto $\theta(t) = \theta(1) = 0$, que corresponde con el menor círculo máximo que conecta **A** y **B**.

4.1.2 Problemas temporales

Si viajamos con velocidad constante entonces la ruta geodésica determinada en cada uno de los problemas de la sección anterior también proporciona el camino más rápido uniendo los puntos **A** y **B**. Sin embargo, si no podemos viajar con velocidad constante y en particular si la velocidad depende del camino elegido, entonces los problemas de mínima distancia y mínimo tiempo deben considerarse separadamente. A continuación vemos alguno de estos problemas.

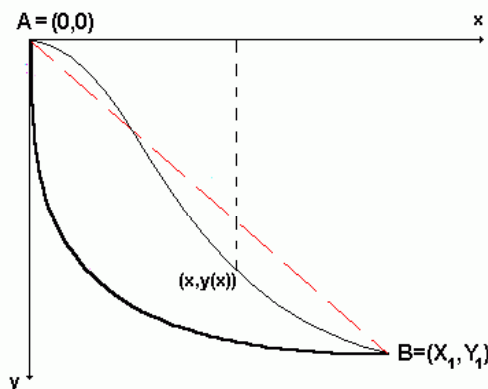


Figura 4.4: Braquistócrona.

Braquistócrona

Al caer bajo la acción de la gravedad un objeto acelera rápidamente. La curva que proporciona el menor tiempo de tránsito entre dos puntos distintos **A** y **B**, no es ni la línea recta, como puede pensarse, ni tampoco un arco circular, sino la llamada *braquistócrona* o *cicloide*. Si utilizamos el esquema indicado en la figura 4.4, en el que el punto inicial **A** está en el origen y las ordenadas y positivas se toman hacia abajo. Entonces una curva que una el punto **A** con el punto **B** = (x_1, y_1) situado más abajo ($x_1, y_1 \geq 0$), puede representarse como el *grafo* de una función continua $y = y(x)$ definida en $x \in [0, x_1]$ con $y(0) = 0$ e $y(x_1) = y_1$ (figura 4.4).

Asumiendo que $y(x)$ es suficientemente derivable, esta curva tendrá una longitud finita l y el tiempo necesario para recorrerla está dado (por consideraciones puramente cinemáticas) por

$$T = T(y) = \int_0^1 \frac{ds}{v}$$

donde $v = ds/dt$ es la velocidad del objeto.

Del cálculo infinitesimal, obtenemos que para cada $x \in [0, x_1]$

$$s = s(x) = \int_0^x \sqrt{1 + y'(\xi)^2} d\xi$$

es la longitud de arco correspondiente a la posición horizontal x , siendo $v = v(x)$ la velocidad asociada. De esta forma

$$ds(x) = \sqrt{1 + y'(x)^2} dx$$

y el tiempo de tránsito es

$$T = T(y) = \int_0^{x_1} \frac{\sqrt{1 + y'(x)^2}}{v(x)} dx$$

Podemos expresar $v(x)$ en términos de $y(x)$ utilizando las leyes de Newton de la dinámica, para ello asumimos que la aceleración gravitacional g es constante durante la caída y despreciamos los efectos de otras fuerzas (incluyendo las de fricción). La velocidad de la bola a lo largo de la trayectoria $y(x)$ en tiempo t viene dada por

$$v(x) = \sqrt{2gy(x)}$$

De esta forma

$$T = T(y) = \frac{1}{\sqrt{2g}} \int_0^{x_1} \left(\frac{1 + y'(x)^2}{y(x)} \right)^{1/2} dx$$

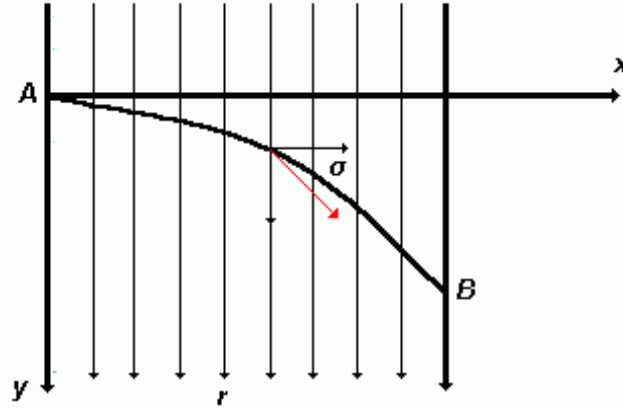


Figura 4.5: Problema de control y dirección.

y el problema consistirá en minimizar $T(y)$ sobre el conjunto de todas las funciones $y = y(x)$ para las que está definida la integral anterior.

Problemas de control y dirección

Estrechamente relacionado con el problema de la braquistócrona, está el problema de encontrar la mejor ruta de navegación al atravesar una corriente de velocidad variable. Podemos preguntarnos cuál será la ruta que proporcionará el menor tiempo de tránsito al intentar cruzar un río de velocidad cambiante desde un punto **A** hasta otro **B** en la orilla opuesta con un barco que navega a velocidad constante w respecto al agua. Si suponemos que las orillas del río son paralelas y utilizamos un determinado sistema de coordenadas, en el que el eje y representa una orilla y la recta $x = x_1$ la otra, asumimos también que $w = 1$ y que la corriente del río $r(x)$ está dirigida hacia abajo (figura 4.5).

Como estamos considerando la velocidad $w = 1$, las componentes x e y de la velocidad del barco se obtienen, de la siguiente manera. Si σ es el ángulo que forma la dirección de movimiento del barco respecto a la corriente, entonces

$$v_x = w \cos \sigma = \cos \sigma$$

$$v_y = w \sin \sigma + r(x) = r + \sin \sigma$$

puesto que a la velocidad de avance del barco hay que añadirle la velocidad de la corriente.

El tiempo de tránsito entre las orillas será

$$T = \int_0^{x_1} \frac{dx}{v_x} = \int_0^{x_1} \frac{1}{\cos \sigma} d\sigma = \int_0^{x_1} \sec \sigma d\sigma$$

El valor de la tangente en cada punto de la curva $y(x)$, es decir su derivada en cada punto $y'(x)$ se obtiene como

$$\begin{aligned} y'(x) &= \frac{v_y}{v_x} = \frac{r + \sin \sigma}{\cos \sigma} = \frac{r}{\cos \sigma} + \frac{\sin \sigma}{\cos \sigma} = r \sec \sigma + \sqrt{\frac{\sin^2 \sigma}{\cos^2 \sigma}} = r \sec \sigma + \sqrt{\frac{1 - \cos^2 \sigma}{\cos^2 \sigma}} = \\ &= r \sec \sigma + \sqrt{\frac{1}{\cos^2 \sigma} - 1} = r \sec \sigma + \sqrt{\sec^2 \sigma - 1} \end{aligned}$$

y se obtiene una relación entre $y'(x)$ y $\sec \sigma$. Despejando $\sec \sigma$ de la ecuación anterior en función de $y'(x)$, donde se tiene en cuenta que $\sigma \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \Rightarrow \cos \sigma = \frac{1}{\sec \sigma} \geq 0$

$$\sec \sigma = \left[\alpha(x) \sqrt{1 + (\alpha y')^2(x)} - (\alpha^2 r y')(x) \right]$$

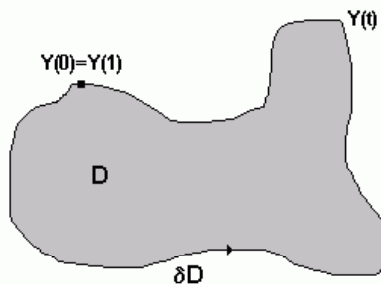


Figura 4.6: Área encerrada por una curva.

Entonces el tiempo de tránsito de un barco que cruce el río desde el origen \mathbf{A} hasta el punto $\mathbf{B} = (x_1, y_1)$ situado río abajo a lo largo de un camino suave que representa la gráfica de una función $y = y(x)$ en $[0, x_1]$ está dado por

$$T(y) = \int_0^{x_1} \left[\alpha(x) \sqrt{1 + (\alpha y')^2(x)} - (\alpha^2 r y')(x) \right] dx$$

donde $\alpha(x) = (1 - r^2(x))^{-1/2}$. (Para poder cruzar la corriente debe ocurrir $0 \leq r(x) \leq 1$). Necesitamos entonces, minimizar esta integral sobre todas las funciones y que son continuamente diferenciables en $[0, x_1]$ y cumplen $y(0) = 0$ e $y(x_1) = y_1$.

Este problema se puede generalizar al caso en el que las orillas del río están representadas mediante curvas y en el que los puntos de cruce varían. También es posible considerar que la corriente varía respecto a y igual que lo hace respecto a x . En 1931, Zermelo investigó la versión bidimensional de este problema que podría ser igualmente significativa para pilotar un submarino o un aeroplano.

4.1.3 Problemas Isoperimétricos

Uno de los problemas más antiguos de optimización de los que se tiene constancia es encontrar el área máxima que puede encerrarse por una curva de perímetro fijo. La solución de este problema es bien conocida y se trata del círculo. Este tipo de problemas se denomina isoperimétrico.

La formulación matemática de este problema sería la siguiente. Suponemos que una curva cerrada simple y suave de longitud L está representada paramétricamente por $\mathbf{Y}(t) = (x(t), y(t))$ para $t \in [0, 1]$, con $\mathbf{Y}(0) = \mathbf{Y}(1)$ para que esta sea cerrada.

De acuerdo con el teorema de Green el área del dominio D encerrado por esta curva es

$$A(\mathbf{Y}) = \iint_D dx dy = \int_{\partial D} x dy$$

donde ∂D representa la frontera de D orientada positivamente por la parametrización de \mathbf{Y} (figura 4.6).

Utilizando esta parametrización obtenemos

$$A(\mathbf{Y}) = \int_0^1 x(t) y'(t) dt$$

Y tendremos que maximizar $A(\mathbf{Y})$ sobre todas las funciones $\mathbf{Y}(t)$ que tienen componentes continuamente diferenciables en $[0, 1]$ y cumplen la condición de clausura $\mathbf{Y}(0) = \mathbf{Y}(1)$; además se cumple la *condición isoperimétrica*

$$L(\mathbf{Y}) = \int_0^1 \|\mathbf{Y}'(t)\| dt = L$$

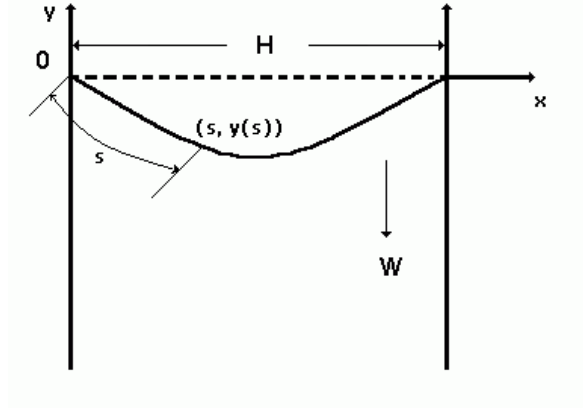


Figura 4.7: Cuerda colgante (Catenaria).

para una longitud L fija.

Una versión moderna de este problema combina los problemas isoperimétricos y de dirección. Consiste en describir el camino cerrado que debe describir un aeroplano con velocidad constante teniendo en cuenta la velocidad del viento, para encerrar el área máxima. Cuando la velocidad del viento es cero, entonces este problema es equivalente a un problema isoperimétrico clásico.

Zenodoros también consideró el problema análogo en dimensión superior que consiste en maximizar el volumen de un sólido que tiene una superficie fija.

Otro problema isoperimétrico diferente consiste en encontrar la forma que tendrá un cable o cadena delgada inextensible, con un peso por unidad de longitud de W y una longitud dada L , bajo la acción de su propio peso cuando está sujeto por sus extremos a una distancia horizontal H (figura 4.7)

Si s es la longitud de arco a lo largo del cable (que tomamos como variable independiente). Este problema requiere la minimización de la energía potencial de la cuerda dada salvo constantes por la integral centro de masas

$$F(y) = W \int_0^L y(s) ds$$

sobre todas las funciones $y \geq 0$ continuamente diferenciables en $[0, L]$, con $y(0) = y(L) = 0$ y que cumplen la condición “isoperimétrica”

$$G(x) = \int_0^L dx(s) = H$$

siendo $x(s)$ el desplazamiento en la dirección horizontal.

Teniendo en cuenta geometría elemental tenemos

$$x'(s)^2 + y'(s)^2 = 1$$

de donde la función $G(x)$ se puede escribir en términos de $y(s)$ como:

$$G(y) = \int_0^L \sqrt{1 - y'(s)^2} ds = H \quad (4.9)$$

En general, el término *isoperimétrico* se asigna a cualquier problema de optimización en el que la clase funciones implicadas están sujetas a restricciones integrales de la forma anterior.

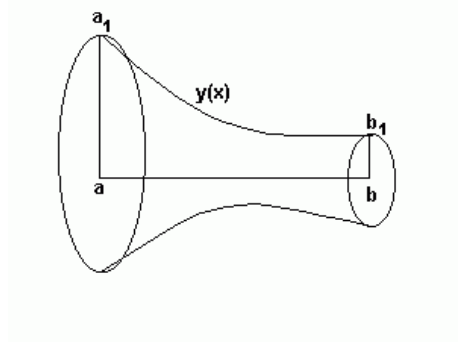


Figura 4.8: Superficie de revolución mínima.

4.1.4 Problemas de áreas de superficies

Un problema análogo al problema geodésico de la primera sección puede plantearse para una dimensión superior de la siguiente forma:

Encontrar la superficie de área mínima que abarcan curvas cerradas fija en \mathbb{R}^3

Superficie de revolución mínima

Por ejemplo, cuando las curvas consisten en una par de círculos paralelos “concéntricos”, entonces podríamos preguntarnos por la superficie de revolución que “uniendo” ambos círculos tenga área mínima, o equivalentemente, queremos encontrar la fórmula de la curva de su frontera (figura 4.8).

Utilizando un sistema de coordenadas adecuado llegamos al problema de minimizar la función área de superficie $S(y)$

$$S(y) = 2\pi \int_a^b y(x) ds(x) = 2\pi \int_a^b y(x) \sqrt{1 + y'(x)^2} dx$$

entre todas las funciones y no negativas, continuamente diferenciables en $[a, b]$ y para las cuales

$$y(a) = a_1 \quad y \quad y(b) = b_1$$

Donde a_1 y b_1 denotan los radios de cada círculo, uno de los cuales puede degenerar en un punto.

Problema del área minimal.

Supongamos que las superficies en cuestión pueden representarse como el grafo de funciones suaves $u = u(x, y)$ definidas en un dominio plano común D . Entonces el área de la superficie asociada están dadas por cálculo como

$$S(u) = \int_D \sqrt{1 + u_x^2 + u_y^2} dA$$

donde dA denota el elemento bidimensional de integración sobre D ; la frontera de D se denota como ∂D y se supone que se comporta bien de forma que la integración Riemann de funciones continuas puede definirse sobre D o \overline{D} y sobre ∂D .

Podríamos entonces buscar el mínimo de $S(u)$ sobre todas las funciones u que son continua en $\overline{D} = D \cup \partial D$, continuamente diferenciables dentro de D y tiene valores en la frontera continuos y dados, por ejemplo $u|_{\partial D} = \gamma$.

4.1.5 Resumen

En resumen todos estos problemas presenta características comunes. Cada uno de ellos implica la búsqueda de una función real definida sobre un determinado dominio \mathcal{D} de funciones \mathbf{y} , que minimiza cierta función

definida por medio de una integral de la forma

$$J(\mathbf{y}) = \int_a^b F(x, \mathbf{y}(x), \mathbf{y}'(x)) dx$$

para alguna función dada $F(x, y, z)$. Este tipo de funciones cuyo valor depende de una función real de variable real se denominan *funcionales*.

Aquí $\mathbf{y}(x)$ es, generalmente, una función real (aunque puede ser una función vectorial) continua en $[a, b]$ con derivada continua en (a, b) ; mientras que el conjunto \mathcal{D} consiste en aquellas funciones de esta clase que cumplen condiciones de frontera específicas en a o b , de la forma

$$\begin{aligned} \mathbf{y}(a) &= a_0 \\ \mathbf{y}(b) &= b_0 \end{aligned}$$

Además, como en el caso de la braquistócrona, puede ser necesario imponer restricciones posteriores tales como exigir que $\mathbf{y} \geq 0$ en $[a, b]$.

Pueden también existir restricciones de forma isoperimétrica en las que

$$H(\mathbf{y}) = \int_a^b h(x, \mathbf{y}(x), \mathbf{y}'(x)) dx = L$$

o bien desigualdades integrales de la forma

$$G(\mathbf{y}) = \int_a^b g(x, \mathbf{y}(x), \mathbf{y}'(x)) dx \leq M$$

o de forma Lagrangiana en las que

$$g(x, y(x), y'(x)) = 0 \quad x \in (a, b)$$

para funciones g y constantes L y M dadas.

Es posible ampliar el grupo de los problemas que tienen la estructura anterior, permitiendo la inclusión de derivadas de orden superior, funciones de más de una variable o condiciones de frontera variables.

La teoría general conocida como *cálculo de variaciones* ha sido desarrollada a lo largo de 300 años para manejar este tipo de problemas.

A partir de la teoría del Cálculo de Variaciones veremos que las funciones que pueden proporcionar los valores extremos deben satisfacer las ecuaciones diferenciales de Euler-Lagrange determinadas por el integrando f y las restricciones.

4.2 Espacios Lineales y Variaciones Gâteaux

Cada uno de los problemas considerado se reduce a la optimización de una función de valor real definida en un subconjunto \mathcal{D} de un espacio de funciones \mathcal{Y} .

4.2.1 Espacios Lineales Reales

Las funciones que se han utilizado en los ejemplos anteriores pertenecen a un determinado conjunto de funciones reales

$$\mathcal{Y} = \{y \mid y : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}\}$$

que en la mayoría de los casos será un espacio lineal real.

Definición 4.1 Se dice que un conjunto de funciones reales \mathcal{Y} , es un espacio lineal real, si cumple las siguientes propiedades:

1. $f, g \in \mathcal{Y} \Rightarrow f + g \in \mathcal{Y}$
2. $c \in \mathbb{R}, f \in \mathcal{Y} \Rightarrow cf \in \mathcal{Y}$

Ejemplos

Es sencillo comprobar que el conjunto de todas las funciones reales f, g definidas sobre un conjunto no vacío $A \subseteq \mathbb{R}^n$, forman un espacio lineal respecto a las operaciones de suma punto a punto:

$$(f + g)(x) = f(x) + g(x), \quad \forall x \in S$$

y multiplicaciones por un escalar:

$$(cf)(x) = cf(x), \quad \forall x \in S, c \in \mathbb{R}$$

Ejemplos de espacios vectoriales lineales de funciones:

1. El conjunto de funciones reales continuas sobre un conjunto $A \subseteq \mathbb{R}^n$

$$\mathcal{Y} = \{y \in \mathcal{C}(A) \mid A \subseteq \mathbb{R}^n\}$$

2. Los conjuntos $\mathcal{C}^m(A)$ de funciones que tienen derivadas parciales continuas hasta el orden m , donde A es un conjunto abierto de \mathbb{R}^n .

$$\mathcal{Y} = \{y \in \mathcal{C}^m(A) \mid A \subseteq \mathbb{R}^n, A \text{ abierto}\}$$

3. Si D es un conjunto acotado y ∂D es su frontera. Entonces $\overline{D} = D \cup \partial D$, y el conjunto $\mathcal{C}^m(\overline{D})$, subconjunto de $\mathcal{C}^m(D) \cap \mathcal{C}^m(\partial D)$, de aquellas funciones cuyas derivadas parciales de orden $\leq m$, admiten extensión continua a \overline{D} , es un espacio lineal real. Por ejemplo, cuando $a, b \in \mathbb{R}$, entonces $\overline{(a, b)} = [a, b]$, es un intervalo cerrado y acotado. Una función y , continua en $[a, b]$, está en $\mathcal{C}^1([a, b])$ si es continuamente diferenciable en (a, b) y su derivada y' tiene límites finito por la derecha en a y por la izquierda en b . Por ejemplo $y_0(x) = x^{3/2}$ define una función de $\mathcal{C}^1([0, 1])$, pero $y_1(x) = x^{1/2}$ no.

4. Los conjuntos $[\mathcal{C}(S)]^n$, $[\mathcal{C}^m(D)]^n$ y $[\mathcal{C}^m(\overline{D})]^n$ de funciones vectoriales n -dimensionales cuyas componentes están en $\mathcal{C}(S)$, $\mathcal{C}^m(D)$ y $\mathcal{C}^m(\overline{D})$ respectivamente, también son espacios lineales.

Sabemos que los subconjuntos \mathcal{D} de los espacios \mathcal{Y} proporcionan dominios naturales para la optimización de las funciones reales de variable real definidas en la introducción del tema. Sin embargo, estos subconjuntos no constituyen en general subespacios lineales por sí mismos. Por ejemplo,

$$\mathcal{D} = \{y \in \mathcal{C}([a, b]) : y(a) = 0, \quad y(b) = 1\}$$

no es un espacio lineal, puesto que si $y \in \mathcal{D}$, entonces $2y \notin \mathcal{D}$ ($2y(b) = 2 \neq 1$). Sin embargo,

$$\mathcal{D} = \{y \in \mathcal{C}([a, b]) : y(a) = 0, \quad y(b) = 0\}$$

sí es un espacio lineal.

En lo siguiente asumiremos la presencia de un espacio lineal real \mathcal{Y} formado por funciones y , y en el que se ha definido una operación de suma (vectorial) y una multiplicación por un escalar (real), que cumple las leyes conmutativa, asociativa y distributiva. En particular, existe un único vector \mathcal{O} tal que $c\mathcal{O} = 0y = \mathcal{O}$, $\forall y \in \mathcal{D}$, $c \in \mathbb{R}$

4.2.2 Funciones del Espacio Lineal

Aunque en principio estamos interesados en la optimización de funciones reales J sobre un subconjunto \mathcal{D} de un espacio lineal \mathcal{Y} , es natural que el dominio de J sea un espacio mayor que \mathcal{D} , incluso que sea \mathcal{Y} .

Definición 4.2 Las funciones $J(y)$ que están definidas como

$$J : \mathcal{D} \subseteq \mathcal{Y} \longrightarrow \mathbb{R}$$

con \mathcal{D} un subconjunto de un espacio lineal \mathcal{Y} , se denominan funcionales.

Ejemplo 4.3 El funcional

$$J(y) = \int_a^b \sin x + y^2(x) dx$$

está definido en todo el conjunto $\mathcal{Y} = \mathcal{C}^1([a, b])$, puesto que cualquier función continua $y \in \mathcal{Y}$, produce un integrando continuo cuya integral es finita.

Ejemplo 4.4 El funcional

$$J(y) = \int_a^b \rho(x) \sqrt{1 + y'(x)^2} dx, \quad \rho \in \mathcal{C}[a, b]$$

está definido para toda función $y \in \mathcal{Y} = \mathcal{C}^1[a, b]$, puesto que si y tiene derivada en (a, b) , con extensión continua a $[a, b]$, de nuevo tenemos un integrando continuo.

J estará definido sobre \mathcal{Y} , cuando ρ sea integrable Riemann sobre $[a, b]$.

Ejemplo 4.5 El funcional definido para la braquistócrona

$$T(y) = \frac{1}{\sqrt{2g}} \int_0^{x_1} \frac{\sqrt{1 + y'(x)^2}}{\sqrt{y(x)}} dx$$

no está definida en todo el espacio $\mathcal{Y} = \mathcal{C}^1[0, x_1]$, por la presencia del término $\sqrt{y(x)}$ en el denominador del integrando. En este caso el dominio del funcional $T(y)$ está definido por el subconjunto de \mathcal{Y}

$$\mathcal{D} = \left\{ y \in \mathcal{C}^1[0, x_1] : y(x) \geq 0, \forall x \in (0, x_1), \text{ and } \int_0^{x_1} (y(x))^{-1/2} dx < \infty \right\}$$

que no es un espacio lineal (Si $y \in \mathcal{D}$, entonces $-y \notin \mathcal{D}$).

Ejemplo 4.6 Si $f \in \mathcal{C}([a, b] \times \mathbb{R}^2)$ entonces el funcional

$$F(y) = \int_a^b f(x, y(x), y'(x)) dx$$

está definido sobre $\mathcal{Y} = \mathcal{C}^1[a, b]$, puesto que para cada $y \in \mathcal{Y}$, la función

$$f[y(x)] = f(x, y(x), y'(x)) \in \mathcal{C}[a, b]$$

Si $f \in \mathcal{C}([a, b] \times D)$, con D un subconjunto de \mathbb{R}^2 , entonces F solamente estará definida en un subconjunto de \mathcal{Y}

$$\mathcal{D} = \{y \in \mathcal{C}^1[a, b] : (y(x), y'(x)) \in D, \forall x \in [a, b]\}$$

4.2.3 Fundamentos de Optimización

Definición 4.7 Sea $J : \mathcal{D} \subseteq \mathcal{Y} \rightarrow \mathbb{R}$ un funcional real definido sobre un subconjunto \mathcal{D} de un espacio lineal \mathcal{Y} . Diremos que $y_0 \in \mathcal{D}$ es un extremal de J sobre \mathcal{D} si

$$J(y) \geq J(y_0) \quad \text{o} \quad J(y) \leq J(y_0) \quad \forall y \in \mathcal{D}$$

Puesto que de nuevo si $J(y) \leq J(y_0) \forall y \in \mathcal{D}$, entonces $-J(y) \geq -J(y_0) \forall y \in \mathcal{D}$, será suficiente caracterizar los puntos de mínimo y_0 que producen valores mínimos para J sobre \mathcal{D} .

Diremos que $y_0 \in \mathcal{D}$ minimiza a J de forma única cuando es el único punto de mínimo en \mathcal{D} , o equivalentemente cuando

$$J(y) > J(y_0), \forall y \neq y_0 \in \mathcal{D}$$

Ejemplo 4.8 Por ejemplo si queremos encontrar el mínimo de $J(y)$ sobre $\mathcal{D} = \mathcal{C}^1[0, 1]$, donde

$$J(y) = \int_0^1 y^2(x) dx$$

En este caso sabemos que

$$\int_0^1 y^2(x) dx \geq 0$$

además si $y_0(x)$ es la función nula sobre el intervalo $[0, 1]$, es decir

$$y_0(x) = 0 \quad x \in [0, 1]$$

Por una parte tendremos

$$y_0 \in \mathcal{D}$$

y por otra

$$J(y) = \int_0^1 y^2(x) dx \geq 0 = J(y_0) \quad \forall y \in \mathcal{D}$$

por tanto y_0 minimiza J sobre $\mathcal{D} = \mathcal{C}^1[0, 1]$. Además, minimiza únicamente, porque $J(y) = 0$ implica que $y^2(x) = 0$, por tanto $y = \mathcal{O} = y_0$.

Lema 4.9 Sea $J : \mathcal{D} \subseteq \mathcal{Y} \rightarrow \mathbb{R}$ un funcional real definido sobre un subconjunto \mathcal{D} de un espacio lineal \mathcal{Y} . Entonces $y_0 \in \mathcal{D}$ minimiza a J sobre \mathcal{D} [de forma única] si y sólo si

$$J(y_0 + v) - J(y_0) \geq 0 \quad \forall y_0 + v \in \mathcal{D}$$

[con igualdad si y sólo si $v = \mathcal{O}$]

Demostración: Para cada $y \in \mathcal{D}$, ponemos $v = y - y_0$, de forma que $y = y_0 + v$ y además $y = y_0$ si y sólo si $v = \mathcal{O}$.

Ejemplo 4.10 Para minimizar $J(y) = \int_a^b y'(x)^2 dx$ sobre

$$\mathcal{D} = \{y \in \mathcal{C}^1[a, b] : y(a) = 0; y(b) = 1\}$$

en primer lugar se observa que

$$J(y) \geq 0$$

y como en el ejemplo anterior, podríamos pensar en utilizar y_0 de forma que $y'_0 = 0$ y por tanto $J(y_0) = 0$. Pero en este caso

$$y'_0(x) = 0 \implies y_0(x) = cte$$

pero entonces

$$y_0(x) \notin \mathcal{D}$$

Si, por el contrario, se plantea la condición de mínimo establecida en el lema anterior, es decir, tomamos $y_0 \in \mathcal{D}$ y v , de forma que $y_0 + v \in \mathcal{D}$, entonces

$$J(y_0 + v) - J(y_0) = \int_a^b \left[(y'_0(x) + v'(x))^2 - y'_0(x)^2 \right] dx$$

$$\begin{aligned}
&= \int_a^b v'(x)^2 dx + 2 \int_a^b y'_0(x) v'(x) dx \\
&\geq 2 \int_a^b y'_0(x) v'(x) dx
\end{aligned}$$

Ahora bien, como $y_0 + v$ debe estar en \mathcal{D} , entonces

$$(y_0 + v)(a) = 0 \iff y_0(a) + v(a) = 0 \iff v(a) = 0$$

$$(y_0 + v)(b) = 1 \iff y_0(b) + v(b) = 1 \iff v(b) = 0$$

donde se ha tenido en cuenta que $y_0 \in \mathcal{D}$, y por tanto $y_0(a) = 0$ e $y_0(b) = 1$.

Si ahora tomamos $y'_0(x) = c_1$, constante, entonces

$$\int_a^b y'_0(x) v'(x) dx = c_1 \int_a^b v'(x) dx = c_1 v(x)|_a^b = 0 \quad \forall v$$

y se tendría la desigualdad buscada

$$J(y_0 + v) - J(y_0) \geq 0$$

De esta forma

$$y'_0(x) = c_1 \implies y_0(x) = c_1 x + c_2$$

y para que $y_0 \in \mathcal{D}$

$$y_0(a) = c_1 a + c_2 = 0$$

$$y_0(b) = c_1 b + c_2 = 1$$

que tiene por solución

$$\begin{aligned}
c_1 &= \frac{1}{b-a} \\
c_2 &= -\frac{a}{b-a}
\end{aligned}$$

luego

$$y_0(x) = \frac{1}{b-a}x - \frac{a}{b-a} = \frac{(x-a)}{(b-a)}$$

Por tanto por el lema, y_0 minimiza J sobre \mathcal{D} . Además lo hace de forma única, porque para que se cumpla la igualdad, tiene que cumplirse $v'(x)^2 = 0$, es decir $v(x) = \text{cte.} = v(a) = 0$, es decir $v = \mathcal{O}$

A continuación se hace una simple observación que nos permite ignorar las constantes.

Proposición 4.11 Sea $J : \mathcal{D} \subseteq \mathcal{Y} \rightarrow \mathbb{R}$ un funcional real definido sobre un subconjunto \mathcal{D} de un espacio lineal \mathcal{Y} . Entonces $y_0 \in \mathcal{D}$ minimiza a J sobre \mathcal{D} [de forma única] si y sólo si para c_0 y $c \neq 0$, y_0 minimiza $c^2 J + c_0$ sobre \mathcal{D} [de forma única]

Demostración: Si $y \in \mathcal{D}$, entonces

$$(c^2 J + c_0)(y) = c^2 J(y) + c_0 \geq c^2 J(y_0) + c_0 = (c^2 J + c_0)(y_0)$$

si y sólo si y_0 minimiza a J sobre \mathcal{D} [con igualdad si y sólo si $y = y_0$].

Utilizando la solución encontrada para el ejemplo anterior, se puede decir que la función $y_0(x) = (x-a)/(b-a)$ minimiza el funcional

$$\begin{aligned}
J_1(y) &= 3 \int_a^b \left(y'(x)^2 + \text{sen}^3(x) \right) dx = 3 \int_a^b y'(x)^2 dx + 3 \int_a^b \text{sen}^3(x) dx \\
&= 3J(y) + c_0
\end{aligned}$$

sobre el conjunto

$$\mathcal{D} = \{y \in \mathcal{C}^1[a, b] : y(a) = 0; y(b) = 1\}$$

4.2.4 Restricciones

Si buscamos la minimización de J sobre \mathcal{D} , cuando esta está restringida a una o más funciones similares G_k , entonces, puede ser suficiente minimizar una función modificada sin restricciones.

Proposición 4.12 *Si los funcionales J y G_1, \dots, G_N están definidos sobre \mathcal{D} , subconjunto de un espacio lineal \mathcal{Y} , y para determinadas constantes $\lambda_1, \dots, \lambda_N$, la función y_0 minimiza $\tilde{J} = J + \lambda_1 G_1 + \dots + \lambda_N G_N$ sobre \mathcal{D} [de forma única], entonces y_0 minimiza J sobre \mathcal{D} [de forma única] cuando nos restringimos al conjunto*

$$G_{y_0} = \{y \in \mathcal{D} : G_j(y) = G_j(y_0), \quad j = 1, \dots, N\}$$

Demostración: Para cada $y \in \mathcal{D}$ tenemos

$$\tilde{J}(y) = J(y) + \sum_{j=1}^N \lambda_j G_j(y) \geq \tilde{J}(y_0) = J(y_0) + \sum_{j=1}^N \lambda_j G_j(y_0)$$

pero cuando $y \in G_{y_0}$, entonces $J(y) \geq J(y_0)$, puesto que los términos que implican las G_j tienen el mismo valor en ambos lados de la desigualdad. Si existe, la unicidad se preserva claramente.

Corolario 4.13 *Si los funcionales J y G_1, \dots, G_N están definidos sobre \mathcal{D} , subconjunto de un espacio lineal \mathcal{Y} , y para determinadas constantes $\lambda_1, \dots, \lambda_N$, la función y_0 minimiza $\tilde{J} = J + \lambda_1 G_1 + \dots + \lambda_N G_N$ sobre \mathcal{D} [de forma única], entonces y_0 minimiza J sobre \mathcal{D} [respectivamente de forma única] cuando nos restringimos al conjunto*

$$G_{y_0}^* = \{y \in \mathcal{D} : \lambda_j G_j(y) \leq \lambda_j G_j(y_0), \quad j = 1, \dots, N\}$$

Demostración: Si y_0 minimiza \tilde{J} entonces

$$\tilde{J}(y) \geq \tilde{J}(y_0) \Leftrightarrow J(y) + \sum_{j=1}^N \lambda_j G_j(y) \geq J(y_0) + \sum_{j=1}^N \lambda_j G_j(y_0)$$

de donde

$$J(y) - J(y_0) \geq \sum_{j=1}^N [\lambda_j G_j(y_0) - \lambda_j G_j(y)]$$

Si ahora $y \in G_{y_0}^*$, entonces

$$\lambda_j G_j(y) \leq \lambda_j G_j(y_0) \Leftrightarrow [\lambda_j G_j(y_0) - \lambda_j G_j(y)] \geq 0 \quad j = 1, \dots, N$$

de forma que en el sumatorio solamente hay términos positivos y por tanto la suma total será positiva, lo que proporciona

$$J(y) - J(y_0) \geq \sum_{j=1}^N [\lambda_j G_j(y_0) - \lambda_j G_j(y)] \geq 0$$

y la función y_0 minimiza $J(y)$ en el conjunto $G_{y_0}^*$.

Si $J(y) = J(y_0)$ entonces

$$0 = J(y) - J(y_0) \geq \sum_{j=1}^N [\lambda_j G_j(y_0) - \lambda_j G_j(y)] \geq 0$$

de donde

$$\sum_{j=1}^N [\lambda_j G_j(y_0) - \lambda_j G_j(y)] = 0$$

como de nuevo todos los sumandos tienen el mismo signo, se debe cumplir

$$\lambda_j G_j(y_0) = \lambda_j G_j(y) \quad j = 1, \dots, N,$$

y por tanto $\tilde{J}(y) = \tilde{J}(y_0)$ [con unicidad, se obtiene que $y = y_0$].

Este resultado ilustra un principio importante: *La solución a un problema de minimización puede proporcionar la solución para otros problemas*

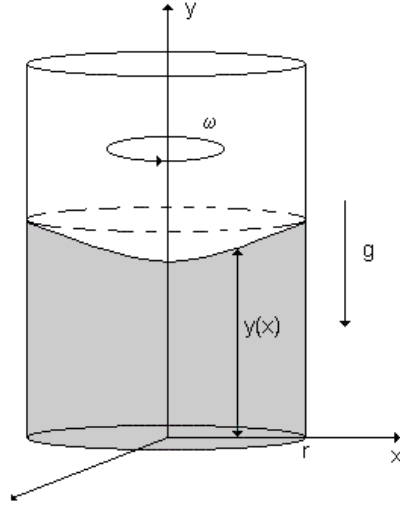


Figura 4.9: Minimización con restricciones.

Aplicación: Columna de Fluido en Rotación

Se hace rotar respecto a su eje vertical a una columna circular de agua de radio r a una velocidad constante ω , dentro de un cilindro de paredes lisas (figura 4.9). Entonces la superficie del agua adquiere una forma que preserva el volumen del fluido y minimiza la energía potencial. Esta energía potencial está dada, salvo constantes, por

$$J(y) = \rho\pi \int_0^r [gy^2(x) - \omega^2 x^2 y(x)] x dx$$

sobre

$$\mathcal{D} = \{y \in \mathcal{C}[0, l] : y(x) \geq 0\}$$

Aquí ρ es la densidad de masa del agua y g es la aceleración de la gravedad, mientras que $y(x)$ es la altura del líquido a una distancia radial x del centro.

El Volumen de la columna es fijo y está dado por

$$G(y) = 2\pi \int_0^r xy(x) dx = V$$

Por tanto y de acuerdo a la proposición 4.12, consideraremos la minimización de $\tilde{J}(y) = J(y) + \lambda G(y)$ sobre \mathcal{D} .

$$\begin{aligned} \tilde{J}(y) &= \rho\pi \int_0^r [gy^2(x) - \omega^2 x^2 y(x)] x dx + \lambda 2\pi \int_0^r xy(x) dx \\ &= \pi \int_0^r \{\rho [gy^2(x) - \omega^2 x^2 y(x)] + 2\lambda y(x)\} x dx \end{aligned}$$

Por el Lema 4.9 tenemos que examinar la diferencia $\tilde{J}(y+v) - \tilde{J}(y)$

$$\tilde{J}(y+v) - \tilde{J}(y) =$$

$$\begin{aligned}
&= \pi \int_0^r \left\{ \rho \left[g(y+v)^2 - \omega^2 x^2 (y+v) \right] + 2\lambda(y+v) \right\} x dx - \pi \int_0^r \left\{ \rho \left[gy^2 - \omega^2 x^2 y \right] + 2\lambda y \right\} x dx \\
&= \pi \int_0^r \left\{ \rho \left[g(v^2 + 2yv) - \omega^2 x^2 v \right] + 2\lambda v \right\} x dx
\end{aligned}$$

donde hemos tomado

$$y \equiv y(x)$$

$$v \equiv v(x)$$

Por claridad se hace el cambio

$$\lambda = \frac{\rho\mu}{2}$$

y al simplificar se obtiene

$$\tilde{J}(y+v) - \tilde{J}(y) = \pi\rho \int_0^r \left\{ gv^2 + [2gy + (\mu - \omega^2 x^2)] v \right\} x dx$$

si ahora se tiene en cuenta que $x \in [0, r]$ y por tanto $x \geq 0$, entonces

$$\int_0^r xgv^2 dx \geq 0$$

y por tanto se cumple

$$\tilde{J}(y+v) - \tilde{J}(y) \geq \pi\rho \int_0^r [2gy + (\mu - \omega^2 x^2)] v x dx$$

Esta última integral se anula $\forall y+v \in \mathcal{D}$ cuando el término entre corchetes es 0, es decir, cuando

$$y(x) = y_0(x) = \frac{\omega^2 x^2 - \mu}{2g}$$

Además, la igualdad solamente es posible cuando $\int_0^r v^2(x) dx = 0$, lo que implica $v = \mathcal{O}$, por tanto $y_0(x)$ minimiza \tilde{J} sobre \mathcal{D} de forma única. La superficie será el *parabolide* de revolución generado por la ecuación anterior.

En este caso

$$G(y_0) = 2\pi \int_0^r xy_0(x) dx = 2\pi \int_0^r \left(\frac{\omega^2 x^2 - \mu}{2g} \right) x dx = \frac{\pi r^2}{2g} \left[\frac{\omega^2 r^2}{2} - \mu \right] = V$$

de donde

$$\mu = \frac{\pi\omega^2 r^4 - 4gV}{2\pi r^2}$$

Se observa que el mínimo depende de ω , tal y como esperabamos, que tiene altura constante cuando $\omega = 0$ (cuando no hay rotación), pero es independiente de la densidad ρ y por tanto la forma será la misma para otro fluido perfecto.

Proposición 4.14 *Supongamos que $f = f(x, y, z)$ y $g = g(x, y, z)$ son funciones continuas sobre $[a, b] \times \mathbb{R}^2$ y que existe una función $\lambda = \lambda(x) \in \mathcal{C}[a, b]$ para la que y_0 minimiza [de forma única] la función*

$$\tilde{J}(y) = \int_a^b \tilde{f}(x, y, y') dx$$

sobre $\mathcal{D} \subseteq \mathcal{C}^1[a, b]$, donde

$$\tilde{f}(x, y, y') = f(x, y, y') + \lambda(x) g(x, y, y')$$

Entonces y_0 minimiza [respectivamente de forma única]

$$J(y) = \int_a^b f(x, y, y') dx$$

sobre \mathcal{D} , bajo la restricción de desigualdad

$$\lambda(x) g(x, y, y') \leq \lambda(x) g(x, y_0, y'_0) \quad \forall x \in [a, b] \quad (4.10)$$

Demostración: Si $y \in \mathcal{D}$ entonces como y_0 minimiza a \tilde{J} tendremos

$$\tilde{J}(y) \geq \tilde{J}(y_0) \implies J(y) + \int_a^b \lambda(x) g(x, y, y') dx \geq J(y_0) + \int_a^b \lambda(x) g(x, y_0, y'_0) dx$$

por tanto

$$J(y) - J(y_0) \geq \int_a^b \lambda(x) (g(x, y_0, y'_0) - g(x, y, y')) dx \geq 0$$

donde se ha utilizado la expresión 4.10 para establecer la última desigualdad.

Además si $J(y) = J(y_0)$ entonces

$$0 = J(y) - J(y_0) \geq \int_a^b \lambda(x) (g(x, y_0, y'_0) - g(x, y, y')) dx \geq 0$$

y por tanto

$$\lambda(x) (g(x, y_0, y'_0) - g(x, y, y')) = 0$$

lo que implica que $\tilde{J}(y) = \tilde{J}(y_0)$ y por tanto $y = y_0$ si y_0 minimiza a $\tilde{J}(y)$ de forma única.

Desafortunadamente, este resultado no puede implementarse fácilmente, puesto que no proporciona un método para determinar la expresión adecuada de $\lambda(x)$. Sin embargo, puede servir para utilizar la solución conocida de otro problema.

Ejemplo 4.15 Encuentra el mínimo del funcional

$$J(y) = \int_0^b 2xy(x) dx$$

en

$$\mathcal{D} = \mathcal{C}[0, b]$$

bajo la restricción

$$g(y(x)) = 1 - x^4$$

Solución: En primer lugar vamos a resolver el problema

$$\text{Minimizar } F(y) = \int_0^b [\rho(x) y^2(x) + \beta(x) y(x)] dx \\ \rho > 0, \beta, y \in \mathcal{D}$$

Si $y_0 \in \mathcal{D}$, minimiza a $F(y)$ entonces

$$F(y_0 + v) - F(y_0) \geq 0 \quad \forall y_0 + v \in \mathcal{D}$$

es decir

$$\begin{aligned} F(y_0 + v) - F(y_0) &= \int_0^b [\rho(y_0 + v)^2 + \beta(y_0 + v)] dx - \int_0^b [\rho y_0^2 + \beta y_0] dx \\ &= \int_0^b [\rho v^2 + 2y_0 \rho v + \beta v] dx \end{aligned}$$

donde por comodidad se ha quitado la dependencia explícita de las funciones. Como $\rho > 0$, tendremos

$$F(y_0 + v) - F(y_0) \geq \int_0^b [2y_0 \rho + \beta] v dx$$

si definimo $y_0(x)$ como

$$2y_0 \rho + \beta = 0 \Leftrightarrow y_0 = -\frac{\beta}{2\rho}$$

entonces la integral se anula y

$$F(y_0 + v) - F(y_0) \geq 0$$

La función $y_0(x) = -\frac{\beta(x)}{2\rho(x)}$, así definida está en \mathcal{D} , puesto que el denominador no se anula y es la función buscada. Es decir, y_0 minimiza $F(y)$, además lo hace de forma única, ya que es la única función que anula la integral.

Para resolver el problema inicial (minimizar $J(y)$). Tomamos $\rho(x) = 1$ y $\beta(x) = 2 - 2x^2$, y se obtiene que la función

$$y_0(x) = \frac{-\beta(x)}{2\rho(x)} = x^2 - 1$$

minimiza de forma única sobre $\mathcal{D} = \mathcal{C}^1[0, b]$ al funcional

$$\begin{aligned} \tilde{J}(y) &= \int_0^b [y^2(x) + (2 - 2x^2)y(x)] dx = \int_0^b 2xy(x) + x(y^2(x) - 2x^2y(x)) dx \\ &= \int_0^b 2xy(x) dx + \int_0^b x(y^2(x) - 2x^2y(x)) dx \end{aligned}$$

Si ahora tomamos $\lambda(x) = x$ y $g(y(x)) = y^2(x) - 2x^2y(x)$, de la proposición anterior se obtiene que y_0 también minimiza

$$J(y) = \int_0^b 2xy(x) dx$$

sobre \mathcal{D} de forma única bajo las restricciones Lagrangianas

$$g[y_0(x)] = y_0^2(x) - 2x^2y_0(x) = 1 - x^4$$

y puesto que $\lambda(x) = x \geq 0$ sobre $[0, b]$, también bajo la restricción de desigualdad $g[y(x)] \leq 1 - x^4$

4.2.5 Variaciones Gâteaux

Un papel decisivo en la optimización de funciones reales sobre un subconjunto de \mathbb{R}^n era el jugado por las derivadas parciales (o más generalmente por sus derivadas direccionales). Cuando $J(y)$ es un funcional real definido sobre un subconjunto de un espacio lineal \mathcal{Y} , no parece tan evidente la forma en la que podemos definir sus derivadas parciales (a menos que en \mathcal{Y} pueda asignarse un sistema de coordenadas). No obstante, se construirá una definición de sus “derivadas direccionales” a partir de una generalización de lo que sucede en \mathbb{R}^n .

Definición 4.16 Sea $J : \mathcal{Y} \rightarrow \mathbb{R}$ un funcional real definido sobre un espacio lineal \mathcal{Y} . Entonces para $y, v \in \mathcal{Y}$, definimos la variación Gâteaux de J en y en la dirección v , si existe, al siguiente límite

$$\delta J(y; v) = \lim_{\varepsilon \rightarrow 0} \frac{J(y + \varepsilon v) - J(y)}{\varepsilon} \quad (4.11)$$

Se observa que tanto y como la dirección v son fijos al tomar el límite cuando $\varepsilon \rightarrow 0$. La existencia de dicho límite presupone que

1. $J(y)$ está definida
2. $J(y + \varepsilon v)$ está definida para todo ε suficientemente pequeño y

$$\delta J(y; v) = \left. \frac{\partial}{\partial \varepsilon} J(y + \varepsilon v) \right|_{\varepsilon=0} \quad (4.12)$$

está derivada “ordinaria” respecto a la variable real ε , existe en $\varepsilon = 0$.

La variación Gâteaux de J en y depende solamente de la conducta local de J cerca de y , sin embargo, esta variación no necesita existir en cualquier dirección $v \neq \mathcal{O}$ o puede existir en unas direcciones y no en otras.

A partir de su definición (ecuación 4.12) se comprueba con facilidad que la variación Gâteaux es una operación lineal sobre los funcionales J ; es decir, si para $y_0 \in \mathcal{Y}$ los funcionales J y \tilde{J} tienen variaciones Gâteaux en la dirección v , entonces para $c, \tilde{c} \in \mathbb{R}$, $\delta(cJ + \tilde{c}\tilde{J})(y_0; v)$ existe y es igual a $c\delta J(y_0; v) + \tilde{c}\delta\tilde{J}(y_0; v)$. Esta propiedad se obtiene directamente a partir de 4.12 y de la linealidad de la derivada ordinaria. También podemos observar que $\delta J(y; \mathcal{O}) = 0$, $\forall y$ donde $J(y)$ esté definida y que cuando $c \in \mathbb{R}$, $\delta J(y; cv) = c\delta J(y; v)$ siempre que la variación $\delta J(y; v)$ exista. En particular si $c = -1$

$$\delta J(y; -v) = -\delta J(y; v) = \delta(-J)(y; v)$$

Ejemplo 4.17 Si $J = f \in C^1(\mathbb{R}^n)$ con $Y, V \in \mathcal{Y} = \mathbb{R}^n$ entonces

$$\delta J(Y; V) = \lim_{\varepsilon \rightarrow 0} \frac{f(Y + \varepsilon V) - f(Y)}{\varepsilon} = \nabla f(Y) \cdot V$$

Ejemplo 4.18 Si $\mathcal{Y} = C[a, b]$, entonces si

$$J(y) = \int_a^b \left[\sin^3 x + y^2(x) \right] dx$$

está definida $\forall y \in \mathcal{Y}$. Así para $y, v \in \mathcal{Y}$ fijos con $\varepsilon \neq 0$; $y + \varepsilon v \in \mathcal{Y}$, puesto que \mathcal{Y} es un espacio lineal, por tanto

$$J(y + \varepsilon v) = \int_a^b \left[\sin^3 x + (y + \varepsilon v)^2(x) \right] dx$$

Si hacemos la diferencia entre las expresiones anteriores y dividimos por ε para obtener la variación Gâteaux

$$\begin{aligned} \frac{J(y + \varepsilon v) - J(y)}{\varepsilon} &= \frac{1}{\varepsilon} \int_a^b \left\{ \left[\sin^3 x + (y + \varepsilon v)^2(x) \right] - \left[\sin^3 x + y^2(x) \right] \right\} dx \\ &= \frac{1}{\varepsilon} \int_a^b \left[(y + \varepsilon v)^2(x) - y^2(x) \right] dx \\ &= \frac{1}{\varepsilon} \int_a^b \left[y^2(x) + \varepsilon^2 v^2(x) + 2\varepsilon y(x)v(x) - y^2(x) \right] dx \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{\varepsilon} \int_a^b [\varepsilon^2 v^2(x) + 2\varepsilon y(x) v(x)] dx \\
&= \int_a^b [\varepsilon v^2(x) + 2y(x) v(x)] dx \\
&= \varepsilon \int_a^b v^2(x) dx + 2 \int_a^b y(x) v(x) dx
\end{aligned}$$

Cada una de las integrales anteriores es una constante y el límite cuando $\varepsilon \rightarrow 0$ existe. Por tanto de la definición de variación Gâteaux se tiene

$$\delta F(y; v) = 2 \int_a^b y(x) v(x) dx$$

De la misma manera utilizando 4.12, podríamos formar:

$$\begin{aligned}
J(y + \varepsilon v) &= \int_a^b [\sin^3 x + (y + \varepsilon v)^2(x)] dx \\
&= \int_a^b [\sin^3 x + y^2(x)] dt + 2\varepsilon \int_a^b y(x) v(x) dx + \varepsilon^2 \int_a^b v^2(x) dx
\end{aligned}$$

que para y y v fijos, la derivada es

$$\frac{\partial}{\partial \varepsilon} J(y + \varepsilon v) = 2 \int_a^b y(x) v(x) dx + 2\varepsilon \int_a^b v^2(x) dx$$

y evaluada en $\varepsilon = 0$

$$\left. \frac{\partial}{\partial \varepsilon} J(y + \varepsilon v) \right|_{\varepsilon=0} = 2 \int_a^b y(x) v(x) dx$$

que es el mismo resultado de antes.

Ejemplo 4.19 Cuando $\rho \in C[a, b]$, la función

$$J(y) = \int_a^b \rho(x) \sqrt{1 + y'(x)^2} dx$$

está definida $\forall y \in \mathcal{Y} = C^1[a, b]$.

Por tanto, utilizando el segundo método, construimos para $y, v \in \mathcal{Y}$ fijos

$$J(y + \varepsilon v) = \int_a^b \rho(x) \sqrt{1 + (y + \varepsilon v)'(x)^2} dx$$

Calculando su derivada, diferenciando bajo el signo de la integral obtenemos

$$\begin{aligned}
\frac{\partial}{\partial \varepsilon} J(y + \varepsilon v) &= \int_a^b \left[\rho(x) \frac{(y + \varepsilon v)'(x) v'(x)}{\sqrt{1 + [(y + \varepsilon v)'(x)]^2}} \right] dx \\
&= \int_a^b \frac{\rho(x) (y + \varepsilon v)'(x) v'(x)}{\sqrt{1 + [(y + \varepsilon v)'(x)]^2}} dx
\end{aligned}$$

que está justificado por la continuidad del último integrando en $[a, b] \times \mathbb{R}$, y evaluando en $\varepsilon = 0$

$$\delta J(y; v) = \left. \frac{\partial}{\partial \varepsilon} J(y + \varepsilon v) \right|_{\varepsilon=0} = \int_a^b \frac{\rho(x) y'(x) v'(x)}{\sqrt{1 + y'(x)^2}} dx$$

Ejemplo 4.20 Cuando $f \in \mathcal{C}([a, b] \times \mathbb{R}^2)$ la función

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

está definida para $y \in \mathcal{Y} = \mathcal{C}^1[a, b]$. Sin embargo para calcular $\delta F(y; v)$ diferenciando la expresión

$$J(y + \varepsilon v) = \int_a^b f(x, (y + \varepsilon v)(x), (y + \varepsilon v)'(x)) dx = \int_a^b f[(y + \varepsilon v)(x)] dx$$

con respecto a ε bajo el signo de la integral, necesitamos que la función $f(x, y, z)$ tenga derivadas parciales continuas en $[a, b] \times \mathbb{R}^2$

$$f_y = \frac{\partial f}{\partial y}(x, y, z)$$

$$f_z = \frac{\partial f}{\partial z}(x, y, z)$$

entonces de la regla de la cadena para x, y, z, v y w fijos

$$\frac{\partial}{\partial \varepsilon} f(x, y + \varepsilon v, z + \varepsilon w) = f_y(x, y + \varepsilon v, z + \varepsilon w) v + f_z(x, y + \varepsilon v, z + \varepsilon w) w$$

Con esta expresión y sabiendo que f, f_y y f_z son continuas

$$\begin{aligned} \frac{\partial}{\partial \varepsilon} J(y + \varepsilon v) &= \frac{\partial}{\partial \varepsilon} \int_a^b f[(y + \varepsilon v)(x)] dx \\ &= \int_a^b \frac{\partial}{\partial \varepsilon} f(x, y(x) + \varepsilon v(x), y'(x) + \varepsilon v'(x)) dx \\ &= \int_a^b f_y[(y + \varepsilon v)(x)] v(x) + f_z[(y + \varepsilon v)(x)] v'(x) dx \end{aligned}$$

(puesto que el último integrando es continuo sobre $[a, b] \times \mathbb{R}$), por tanto cuando $\varepsilon = 0$, vemos que $F(y) = \int_a^b f(x, y(x), y'(x)) dx$, tiene como variación

$$\begin{aligned} \delta F(y; v) &= \int_a^b f_y[x, y(x), y'(x)] v(x) + f_z[x, y(x), y'(x)] v'(x) dx \\ &= \int_a^b f_y[y(x)] v(x) + f_z[y(x)] v'(x) dx \quad \forall y, v \in \mathcal{C}^1[a, b] \end{aligned}$$

4.3 Lemas de Lagrange y Bois-Reymond

Los siguientes resultados son necesarios para conseguir las ecuaciones de Euler-Lagrange

Lema 4.21 (du Bois-Reymond) Si $h \in \mathcal{C}[a, b]$ con

$$\int_a^b h(x) v'(x) dx = 0 \quad \forall v \in \mathcal{D}_0 = \{v \in \mathcal{C}^1[a, b] \mid v(a) = v(b) = 0\}$$

entonces h es constante en $[a, b]$

Demostración: Para c constante, la función

$$v(x) = \int_a^x (h(t) - c) dt$$

está en $\mathcal{C}^1[a, b]$ (por el teorema fundamental del cálculo) con derivada

$$v'(x) = h(x) - c$$

que además cumple la propiedad $v(a) = 0$.

Para que v esté en el conjunto \mathcal{D}_0 , tiene que ocurrir $v(b) = 0$, es decir,

$$v(b) = \int_a^b (h(t) - c) dt = 0 \quad \text{o} \quad c = \frac{1}{b-a} \int_a^b h(t) dt$$

Si se definen c y v de la forma anterior (puede hacerse puesto que $h(x) \in \mathcal{C}[a, b]$) y se utiliza la hipótesis del Lema tenemos

$$\begin{aligned} 0 &\leq \int_a^b (h(x) - c)^2 dx = \int_a^b (h(x) - c) v'(x) dx \\ &= \int_a^b h(x) v'(x) dx - cv(x)|_a^b = 0 + 0 = 0 \end{aligned}$$

Por tanto como el integrando $(h(x) - c)^2 \geq 0$, entonces en $[a, b]$ ocurre $(h(x) - c)^2 \equiv 0$ o $h(x) = c = cte$.

Proposición 4.22 Si $g, h \in \mathcal{C}[a, b]$ y

$$\int_a^b [g(x)v(x) + h(x)v'(x)] dx = 0 \quad \forall v \in \mathcal{D}_0 = \{v \in \mathcal{C}^1[a, b] \mid v(a) = v(b) = 0\}$$

entonces $h \in \mathcal{C}^1[a, b]$ y $h' = g$

Demostración: Se define

$$G(x) = \int_a^x g(t) dt \quad x \in [a, b]$$

Como g es continua en $[a, b] \implies G \in \mathcal{C}^1[a, b]$ y $G' = g$

Integrando por partes el primer término de la integral, se obtiene

$$\int_a^b [g(x)v(x) + h(x)v'(x)] dx = \int_a^b [h(x) - G(x)] v'(x) dx + G(x)v(x)|_a^b = 0$$

pero como $v \in \mathcal{D}_0$

$$\int_a^b [h(x) - G(x)] v'(x) dx = 0 \quad \forall v \in \mathcal{D}_0$$

y aplicando el lema anterior

$$h(x) - G(x) = cte. = c \quad \text{sobre } [a, b]$$

Pero entonces

$$h = G + c \in \mathcal{C}^1[a, b]$$

y por tanto

$$h' = G' = g$$

Tomando $h \equiv 0$ se obtiene el siguiente resultado

Corolario 4.23 Si $g \in \mathcal{C}[a, b]$ y

$$\int_a^b g(x) v(x) dx = 0 \quad \forall v \in \mathcal{D}_0 = \{v \in \mathcal{C}^1[a, b] \mid v(a) = v(b) = 0\}$$

entonces $g = 0$ en $[a, b]$

Este resultado admite la siguiente generalización

Lema 4.24 (Lagrange) Si $g \in \mathcal{C}[a, b]$ y para algún $m = 0, 1, 2, \dots$

$$\int_a^b g(x) v(x) dx = 0 \quad \forall v \in \mathcal{D}_0 = \left\{v \in \mathcal{C}^m[a, b] \mid v^{(k)}(a) = v^{(k)}(b) = 0, k = 0, \dots, m\right\}$$

entonces $g \equiv 0$ en $[a, b]$

Demostración: Supongamos que $g(c) > 0$ para algún $c \in (a, b)$. Entonces por la hipótesis de continuidad de g , c está contenido en un intervalo $[\alpha, \beta] \subseteq (a, b)$ en el que $|g(x) - g(c)| \leq g(c)/2$ o $g(c)/2 > 0$. Por otra parte la función

$$v(x) = \begin{cases} [(x - \alpha)(\beta - x)]^{m+1} & x \in [\alpha, \beta] \\ 0 & x \notin [\alpha, \beta] \end{cases}$$

es de clase $\mathcal{C}^m(\mathbb{R})$, no negativa y no idénticamente igual a 0 (las m primeras derivadas se anulan en α y β). Se obtiene así que sobre $[a, b]$ el producto gv es continuo, no negativo, y no idénticamente nulo. Por tanto

$$0 < \int_a^b g(x) v(x) dx$$

que contradice la hipótesis.

Similarmente, la suposición de que $g(c) < 0$ (o $-g(c) > 0$) conduce a otra contradicción y concluimos que $g(c) = 0$, $\forall c \in (a, b)$, pero puesto que g es continua, también debe anularse en los extremos del intervalo; es decir $g \equiv 0$ sobre $[a, b]$

4.4 Extremos locales en espacios lineales normados

Como en el caso de optimización estática, la búsqueda de extremos globales de funcionales se hace a través de sus extremos locales. Para poder establecer este carácter local habría que definir dentro de los espacios lineales una distancia, esta siempre estará asociada a una norma. Se describe a continuación y de forma muy breve algunos conceptos relacionados con la topología de los espacios lineales.

4.4.1 Normas en espacios lineales

En el espacio lineal \mathcal{Y} , una norma $\|y\|$, $\forall y \in \mathcal{Y}$ es una función con las propiedades:

$$\|y\| \geq 0, \forall y \in \mathcal{Y}, \text{ con } \|y\| = 0 \Leftrightarrow y = \mathcal{O} \text{ (Definida Positiva)} \quad (4.13)$$

$$\|cy\| = |c| \|y\|, \forall c \in \mathbb{R}, y \in \mathcal{Y} \text{ (Homogénea Positiva)} \quad (4.14)$$

$$\|y + \tilde{y}\| \leq \|y\| + \|\tilde{y}\|, y, \tilde{y} \in \mathcal{Y} \text{ (Desigualdad Triangular)} \quad (4.15)$$

Como en \mathbb{R}^n , pueden definirse diferentes normas en el espacio lineal \mathcal{Y} .

Cualquier norma también cumple la desigualdad triangular inversa

$$|\|y\| - \|\tilde{y}\|| \leq \|y - \tilde{y}\|, \forall y, \tilde{y} \in \mathcal{Y} \quad (4.16)$$

Ejemplo 4.25 Para $\mathcal{Y} = \mathcal{C}[a, b]$, podemos pensar en los valores $y(x)$ como las “componentes” del “vector” $y \in \mathcal{Y}$. En este caso la expresión

$$\|y\|_M = \max |y(x)| = \max_{x \in [a, b]} |y(x)|$$

determina una norma, denominada norma del máximo.

Ejemplo 4.26 Para $\mathcal{Y} = \mathcal{C}[a, b]$ otra norma viene dada por

$$\|y\|_1 = \int_a^b |y(x)| dx$$

Ejemplo 4.27 Para $\mathcal{Y} = \mathcal{C}^1[a, b]$ se pueden definir normas equivalentes de la siguiente forma

$$\|y\|_M = \max_{x \in [a, b]} (|y(x)| + |y'(x)|)$$

y

$$\|y\|_1 = \int_a^b (|y(x)| + |y'(x)|) dx$$

4.4.2 Convergencia

Cuando asociamos una norma a un espacio lineal \mathcal{Y} , inmediatamente podemos definir sobre \mathcal{Y} una estructura topológica asociada que permite un análisis del espacio lineal normado $(\mathcal{Y}, \|\cdot\|)$.

Es posible definir la “distancia” entre los vectores y e \tilde{y} , mediante $\|y - \tilde{y}\|$. La desigualdad triangular implica que para tres vectores x, y , y z de \mathcal{Y} se cumple:

$$\|x - z\| \leq \|x - y\| + \|y - z\|$$

Es posible también, introducir el concepto de convergencia declarando que si $y_n \in \mathcal{Y}$, $n = 1, 2, \dots$, entonces la sucesión $\{y_n\}_n$ tiene límite $y_0 \in \mathcal{Y}$, si y sólo si $\|y_n - y_0\| \rightarrow 0$, cuando $n \rightarrow \infty$. Si el límite existe entonces será único.

Con la definición de distancia podemos definir una bola abierta de radio δ y centro y_0 con $\delta > 0$, al conjunto

$$B_\delta(y_0) = \{y \in \mathcal{Y} : \|y - y_0\| \leq \delta\}$$

Por ejemplo, para $\mathcal{Y} = \mathcal{C}[a, b]$ y la norma del máximo, la bola de centro y_0 y radio δ , es el conjunto de funciones y cuyo grafo no difieren en más de δ del grafo de y_0

4.4.3 Continuidad

Definición 4.28 En un espacio lineal normado $(\mathcal{Y}, \|\cdot\|)$, si $\mathcal{D} \subseteq \mathcal{Y}$, un funcional $J : \mathcal{D} \rightarrow \mathbb{R}$ es una función continua en $y_0 \in \mathcal{D} \Leftrightarrow$

$$\forall \varepsilon > 0, \exists \delta > 0 : \forall y \in \mathcal{D}, \|y - y_0\| < \delta \Rightarrow |J(y) - J(y_0)| < \varepsilon$$

Equivalentemente podemos decir que J es continua en $y_0 \in \mathcal{D} \Leftrightarrow \forall \{y_n\}_n$ de \mathcal{D} , con $\lim_{n \rightarrow \infty} y_n = y_0 \Rightarrow \lim_{n \rightarrow \infty} J(y_n) = J(y_0)$

También diremos que J es continua sobre $\mathcal{D} \Leftrightarrow J$ es continua para cada $y_0 \in \mathcal{D}$.

Ejemplo 4.29 En cualquier espacio lineal normado $(\mathcal{Y}, \|\cdot\|)$, el funcional norma $J(y) = \|y\|$ es una función continua.

Ejemplo 4.30 Cuando $f \in \mathcal{C}([a, b] \times \mathbb{R}^2)$ la función

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

está definida $\forall y \in \mathcal{Y} = \mathcal{C}^1[a, b]$. Para demostrar su continuidad respecto a la norma del máximo $\|\cdot\|_M$ procedemos como sigue:

f es continua sobre $[a, b] \times \mathbb{R}^2$, por tanto f es uniformemente continua sobre cada “caja” compacta de la forma $[a, b] \times [-c, c]^2$, cuando $c > 0$. De esta forma para $y_0 \in \mathcal{Y}$ fijo, cuando $y \in B_1(y_0)$ tendremos

$$\|y\|_M = \|y - y_0 + y_0\|_M \leq \|y - y_0\|_M + \|y_0\|_M \leq 1 + \|y_0\|_M$$

donde en la última desigualdad se ha utilizado el hecho de que $y \in S_1(y_0)$ y de la desigualdad triangular.

Entonces $\forall x \in [a, b]$

$$|y(x)| \leq |y(x)| + |y'(x)| \leq \|y\|_M < 1 + \|y_0\|_M$$

$$|y'(x)| \leq |y(x)| + |y'(x)| \leq \|y\|_M < 1 + \|y_0\|_M$$

Si tomamos $c = c_0 = 1 + \|y_0\|_M$, por la continuidad uniforme de f sobre $[a, b] \times [-c_0, c_0]^2$, obtenemos que para todo $\varepsilon > 0$, $\exists \delta \in (0, 1)$ tal que $\|y - y_0\|_M < \delta (< 1) \Rightarrow$

$$|f[y(x)] - f[y_0(x)]| = |f(x, y(x), y'(x)) - f(x, y_0(x), y'_0(x))| < \varepsilon \quad \forall x \in [a, b]$$

Por tanto

$$|J(y) - J(y_0)| \leq \int_a^b |f[y(x)] - f[y_0(x)]| dx \leq \varepsilon(b - a)$$

cuando

$$\|y - y_0\| \leq \delta < 1$$

y se prueba la continuidad de J en un punto arbitrario $y_0 \in \mathcal{Y}$. Se observa que δ depende de c_0 (y por tanto de y_0), y en general decrecerá cuando c_0 crece.

4.4.4 Puntos Extremos Locales

La definición de extremal que se ha dado es de tipo global por naturaleza y puede hacerse sin consideraciones de norma. Sin embargo, la presencia de una norma permite una descripción análoga a la conducta *local* de J en y_0 .

Definición 4.31 En un espacio lineal normado $(\mathcal{Y}, \|\cdot\|)$ se dice que un punto $y_0 \in \mathcal{D} \subseteq \mathcal{Y}$ es un punto extremal local para J en \mathcal{D} , si para algún $r > 0$, y_0 es un punto extremal para J sobre $\mathcal{D}_r(y_0) = \{y \in \mathcal{D} : \|y - y_0\| < r\}$, es decir ocurre alguna de las dos opciones siguientes

$$J(y) \leq J(y_0), \forall y \in \mathcal{D}_r(y_0) \quad (y_0 \text{ es un punto de máximo local para } J \text{ en } \mathcal{D})$$

$$J(y) \geq J(y_0), \forall y \in \mathcal{D}_r(y_0) \quad (y_0 \text{ es un punto de mínimo local para } J \text{ en } \mathcal{D})$$

Por supuesto, cada punto extremal de $J(y)$ es automáticamente un punto extremal local cualquiera que sea la norma utilizada. Sin embargo, y_0 puede ser un punto extremal local con respecto a una norma y sin embargo *no* serlo respecto a otra.

Recordemos que la construcción de las variaciones Gâteaux de un funcional se ha realizado sin consideraciones de norma. Cuando esta variación no se anula, entonces es indicativo de que no hay conducta de extremal local respecto a ninguna norma.

Supongamos, por ejemplo, que en el punto y_0 , la función J tiene una variación positiva $\delta J(y_0; v)$ en la dirección $v \in \mathcal{Y}$. A partir de la definición de variación Gâteaux, se obtiene que $\forall \varepsilon$ suficientemente pequeño, la razón $[J(y_0 + \varepsilon v) - J(y_0)]/\varepsilon$ es también positiva, por tanto $J(y_0 + \varepsilon v) - J(y_0)$ tiene el signo de ε . Por tanto

$$\delta J(y_0; v) > 0 \Rightarrow J(y_0 - \varepsilon v) < J(y_0) < J(y_0 + \varepsilon v) \quad \forall \varepsilon > 0, \text{ pequeño} \quad (4.17)$$

y podemos decir que en y_0 , J es estrictamente creciente en la dirección v (y decrece estrictamente en la dirección opuesta $-v$).

Cuando $\delta J(y_0; v) < 0$, entonces $\delta J(y_0; -v) > 0$, y las desigualdades anteriores cambian de sentido. En cualquier caso, puesto que cuando $\varepsilon \rightarrow 0$, $\|(y_0 \pm \varepsilon v) - y_0\| = \varepsilon \|v\| \rightarrow 0$, los puntos del tipo $y_0 \pm \varepsilon v$ de 4.17 están en cada entorno de y_0 y la conducta de extremo local de J en y_0 no es posible en la dirección v .

Por ejemplo, si $\mathcal{Y} = \mathcal{C}[a, b]$ y la función $J(y)$ está definida como

$$J(y) = \int_a^b \sin^3 x + y^2(x) dx$$

su variación en la dirección v viene dada por

$$\delta J(y; v) = 2 \int_a^b y(x) v(x) dx$$

Supongamos que $y_0(x) = x^2$, la variación es

$$\delta J(y_0; v) = 2 \int_a^b x^2 v(x) dx$$

que claramente no se anula, por ejemplo para $v(x) = e^x$, y por tanto, en esta dirección, J no presenta conducta de extremo local en y_0 , y como consecuencia de ello, y_0 no puede ser un punto extremal local para J sobre \mathcal{Y} , independientemente de la norma utilizada.

4.4.5 Condiciones Necesarias: Dirección Admisibles

Al minimizar una función real J sobre $\mathcal{D} \subseteq \mathcal{Y}$, donde $(\mathcal{Y}, \|\cdot\|)$, es un espacio lineal normado, es natural considerar para cada $y \in \mathcal{D}$, aquellas direcciones $v \in \mathcal{Y}$, para las cuales la función restringida $J|_{\mathcal{D}}$ admita variaciones en y ; es decir, queremos distinguir aquellas direcciones $v \in \mathcal{Y}$ para las cuales:

1. $\exists \lambda > 0$, tal que $y + \varepsilon v \in \mathcal{D}$, $\forall 0 < \varepsilon < \lambda$
2. Existe la variación de y en la dirección v : $\delta J(y; v)$.

Este tipo de direcciones se denomina *admisibles en y* para \mathcal{D} , o \mathcal{D} -admisibles en y (para J). Podemos observar que si v es \mathcal{D} -admisible en y , entonces también es admisible cada múltiplo escalar de cv para $c \in \mathbb{R}$. La función nula \mathcal{O} es siempre admisible.

Para apreciar su significado, supongamos que y_0 es un punto mínimo local para J sobre \mathcal{D} . Entonces J no puede decrecer estrictamente en ninguna dirección admisible v en y_0 para \mathcal{D} , y por tanto a partir de 4.17 $\delta J(y; v) = 0$ en esas direcciones. Se obtiene el mismo resultado cuando y_0 es un punto de máximo local sobre \mathcal{D} . De esta manera se establece la siguiente:

Proposición 4.32 *En un espacio lineal normado $(\mathcal{Y}, \|\cdot\|)$, si $y_0 \in \mathcal{D} \subseteq \mathcal{Y}$, es un punto extremal local para una función real J sobre \mathcal{D} , entonces*

$$\delta J(y_0; v) = 0, \forall v \text{ direcciones } \mathcal{D}\text{-admisibles en } y_0$$

Como en el caso de la optimización estática, no es posible distinguir entre máximos y mínimos locales (o entre puntos de mínimo local y mínimo global), si solamente se utiliza esta condición. Además, también como en el caso de \mathbb{R}^n , tenemos que admitir la posibilidad de *puntos de silla*, que cumplen la condición pero que no son ni máximo ni mínimo local. Por eso esta condición es *necesaria* pero *no suficiente* para que una función sea un mínimo local.

Ejemplo 4.33 *Para caracterizar los extremos locales de la función*

$$J(y) = \int_a^b \left[\sin^3 x + y^2(x) \right] dx$$

sobre el dominio

$$\mathcal{D} = \{y \in \mathcal{C}[a, b] : y(a) = a_1, y(b) = b_1\}$$

con $a_1, b_1 \in \mathbb{R}$ conocidos. Sabemos que $\delta J(y; v)$ está definida $\forall y, v \in \mathcal{C}[a, b]$. Sin embargo, las únicas direcciones \mathcal{D} -admisibles en $y \in \mathcal{Y}$ son aquellas para las que $y + \varepsilon v \in \mathcal{D}$, con $\varepsilon \neq 0$, suficientemente pequeño. Para que $y + \varepsilon v \in \mathcal{D}$ se deben cumplir las siguientes ecuaciones

$$a_1 = (y + \varepsilon v)(a) = y(a) + \varepsilon v(a) = a_1 + \varepsilon v(a)$$

$$b_1 = (y + \varepsilon v)(a) = y(b) + \varepsilon v(b) = b_1 + \varepsilon v(b)$$

de donde se obtiene directamente $v(a) = v(b) = 0$. Por tanto v es \mathcal{D} -admisible en $y \in \mathcal{D}$ si y sólo si

$$v \in \mathcal{D}_0 = \{v \in \mathcal{C}[a, b] : v(a) = v(b) = 0\}$$

Se observa que el conjunto \mathcal{D}_0 no depende de y y por la proposición anterior la condición necesaria para que $y_0 \in \mathcal{D}$ sea un punto extremal local es que su variación Gâteaux sea nula

$$\delta J(y_0; v) = 2 \int_a^b y_0(x) v(x) dx = 0 \quad \forall v \in \mathcal{D}_0$$

Esta condición se cumple completamente cuando $y_0 \equiv 0$, pero esta función no está en \mathcal{D} , a menos que $a_1 = b_1 = 0$. Sin embargo, por el Lema de Lagrange no existe otra posibilidad para $y_0(x)$; de esta forma no existe extremo local.

4.5 Ecuaciones de Euler-Lagrange

La solución que Jakob Bernouilli (1696) dió a su hermano Johann acerca del problema de la braquistócrona marcaba el comienzo de consideraciones variacionales. Sin embargo fue con los trabajos de Euler (1742) y Lagrange (1755) cuando emergió el cálculo de variaciones tal y como se conoce actualmente. Inicialmente este análisis estaba restringido a la búsqueda de condiciones que fueran *necesarias* para que una función integral del tipo

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

tuviera un extremo (local) en el conjunto

$$\mathcal{D} \subseteq \{y \in \mathcal{C}^1[a, b] : y(a) = a_1; \quad y(b) = b_1\}$$

Para a_1 y b_1 dados, este es un problema de *punto final fijo* (Ver figura 4.4).

Sin embargo para Jakob Bernouilli el interés se centraba en un conjunto mayor

$$\mathcal{D}^b \subseteq \{y \in \mathcal{C}^1[a, b] : y(a) = a_1\}$$

al describir una braquistócrona para la que se pide que descienda sobre una distancia horizontal $(b - a)$ en el mínimo tiempo posible, sin especificar la distancia vertical que hay que cubrir. Este tipo de problemas es llamado de *un punto extremo libre* (ver figura 4.10).

Notar que también pueden considerarse problemas sobre el conjunto

$$\mathcal{D}^a \subseteq \{y \in \mathcal{C}^1[a, b] : y(b) = b_1\}$$

es decir el extremo final es ahora fijo, mientras que el extremo inicial es libre.

También podemos pensar en problemas con los *dos puntos extremos libres* o también en aquellos donde los extremos locales están en subconjuntos arbitrarios de $\mathcal{C}^1[a, b]$, por ejemplo, un problema relacionado con las condiciones sobre los puntos extremos es el que caracteriza la braquistócrona que une dos curvas fijas llamadas *transversales* lo que requiere la minimización de una integral con límites variables

$$J(y) = \int_{x_1}^{x_2} f(x, y(x), y'(x)) dx$$

sobre un conjunto

$$\mathcal{D}_\tau \subseteq \{y \in \mathcal{C}^1[x_1, x_2] : \tau_j(x_j, y(x_j)) = 0; \quad j = 1, 2\}$$

donde $[x_1, x_2] \subseteq \mathbb{R}$ y $\{\tau_j\}_{j=1}^2$ son funciones dadas (figura 4.11)

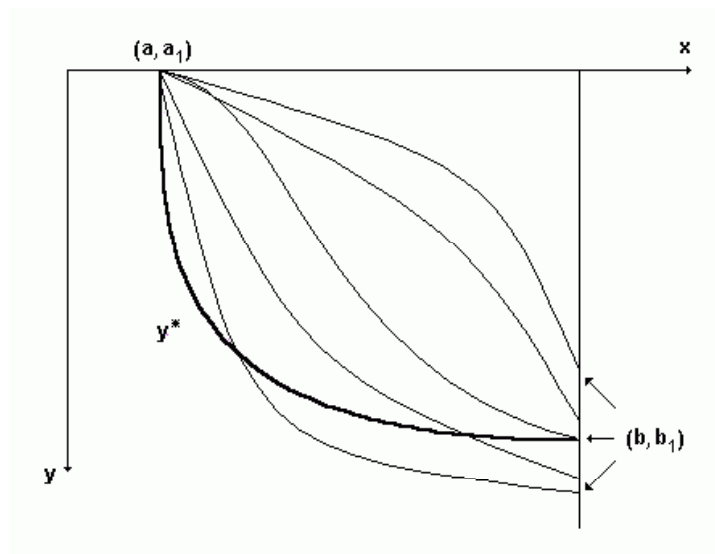


Figura 4.10:

Todos estos problemas admiten una resolución variacional común: Si $y_0 \in \mathcal{C}^1([a, b])$, es por ejemplo, una función que minimiza localmente alguno de estos problemas, entonces con una apropiada selección de a_1 , b_1 y \mathcal{D} podemos suponer que $y_0 \in \mathcal{D}$ y como condición que $\mathcal{D} \subseteq \mathcal{D}^b$ o $\mathcal{D} \subseteq \mathcal{D}_\tau$. En cada caso, y_0 es una función que minimiza (localmente) a J sobre \mathcal{D} , el problema de punto fijo considerado inicialmente.

Consecuentemente de la proposición 4.32 obtendremos

$$\delta J(y_0; v) = 0 \quad \forall v \in \mathcal{D}_0 = \{v \in \mathcal{C}^1[a, b] : v(a) = v(b) = 0\}$$

4.5.1 La primera Ecuación: Funciones estacionarias

Por simplicidad, supongamos inicialmente que la función $f = f(x, y, z)$ y sus dos derivadas f_y y f_z son continuas en $[a, b] \times \mathbb{R}^2$, es decir, $f(x, y, z) \in \mathcal{C}^1([a, b] \times \mathbb{R}^2)$.

Para $y \in \mathcal{Y} = \mathcal{C}^1[a, b]$ se define el siguiente funcional

$$J(y) = \int_a^b f(x, y(x), y'(x)) dx \quad (4.18)$$

En las secciones anteriores se ha comprobado que $J(y)$ tiene en cada dirección $v(x)$ una variación Gâteaux dada por la expresión

$$\delta J(y; v) = \int_a^b [f_y(x) v(x) + f_{y'}(x) v'(x)] dx \quad (4.19)$$

donde se ha utilizado la notación más compacta

$$f_y(x) = f_y(x, y(x), y'(x))$$

$$f_{y'}(x) = f_{y'}(x, y(x), y'(x))$$

Con estas condiciones tenemos la siguiente proposición:

Proposición 4.34 Si $y \in \mathcal{Y}$ hace $\delta J(y; v) = 0$, $\forall v \in \mathcal{D}_0 = \{v \in \mathcal{Y} : v(a) = v(b) = 0\}$ entonces $f_{y'}$ es de clase $\mathcal{C}^1([a, b] \times \mathbb{R}^2)$ y

$$\frac{d}{dx} f_{y'}(x) = f_y(x) \quad \forall x \in (a, b) \quad (4.20)$$

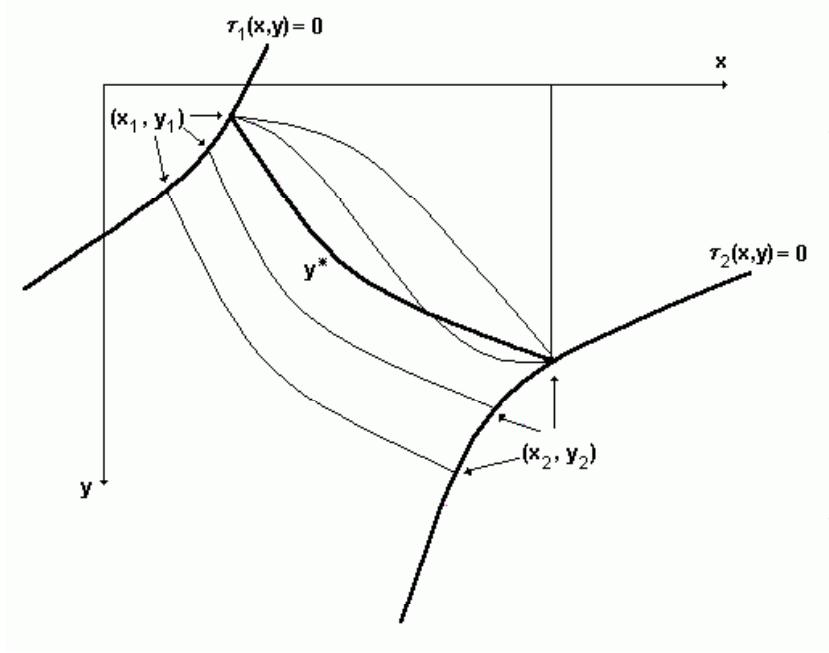


Figura 4.11:

por tanto

$$\delta J(y; v) = f_{y'}(x) v(x) \Big|_a^b \quad \forall v \in \mathcal{Y} \quad (4.21)$$

y si $v \in \mathcal{D}_0$, entonces

$$\delta J(y; v) = 0$$

Demostración: El primer resultado es una repetición de la proposición 4.22 (proposición siguiente al lema de du Bois-Reymond) para las funciones continuas $g(x) = f_y(x)$ y $h(x) = f_{y'}(x)$. En aquel lema se obtenía

$$h'(x) = g(x) \Leftrightarrow \frac{d}{dx} [f_{y'}(x)] = f_y(x)$$

que es la expresión buscada.

Ahora bien, utilizando la expresión 4.20 se permite reconocer en el integrando de 4.19 como

$$\delta J(y; v) = \int_a^b [f_y(x) v(x) + f_{y'}(x) v'(x)] dx = \int_a^b \left[\frac{d}{dx} [f_{y'}(x)] v(x) + f_{y'}(x) v'(x) \right] dx = \int_a^b \left[\frac{d}{dx} f_{y'}(x) v(x) \right] dx$$

de donde, al integrar, se obtiene de forma directa la ecuación 4.21.

La ecuación 4.20 es la llamada *primera ecuación de Euler-Lagrange* (que la obtuvo de forma heurística en 1736 variando los vértices de una curva solución poligonal imaginada) y Lagrange (que la obtuvo en 1755 de forma incorrecta) integrando el segundo término de 4.19 por partes. La demostración anterior que es correcta fue dada por P. du Bois-Reymond en 1879. La ecuación 4.20 debe cumplirse para cada curva que proporcione un extremal local de J en \mathcal{D} (también en \mathcal{D}^b o en \mathcal{D}_τ).

Definición 4.35 *Cualquier función $y \in \mathcal{C}^1$ que cumple la primera ecuación diferencial de Euler-Lagrange, es decir, para la que*

$$\frac{d}{dx} f_{y'}(x, y(x), y'(x)) = f_y(x, y(x), y'(x)) \quad (4.22)$$

en algún intervalo se llama función estacionaria de $J(y)$.

Tradicionalmente se denomina a estas funciones *funciones extremales* o simplemente *extremales*, aunque no proporcionen ni máximo ni mínimo local para el problema. No es necesario que las funciones estacionarias cumplan ninguna condición inicial o de contorno particular.

4.5.2 Casos especiales de la primera ecuación

En general resolver la ecuación 4.20 es bastante difícil. Sin embargo cuando una o más variables de f no están presentes explícitamente, entonces es posible encontrar una primera integral de la ecuación diferencial. Veamos tres de estos casos. Notar en primer lugar que si $y(x) \in \mathcal{C}^2[a, b]$, al aplicar la regla de la cadena al primer miembro de la ecuación 4.20 se obtiene

$$\frac{d}{dx} f_{y'}(x) = f_{y'x} + f_{y'y}y' + f_{y'y'}y'' = f_y(x) \quad (4.23)$$

Cuando $f = f(z)$

En este caso $f_y = 0$ y la ecuación 4.20 se transforma en

$$\frac{d}{dx} f_{y'}(x) = 0$$

o equivalentemente

$$f_{y'}(x) = cte.$$

En el caso particular de que $y \in \mathcal{C}^2[a, b]$, entonces la ecuación 4.23 se transforma en

$$\frac{d}{dx} f_{y'}(x) = f_{y'x} + f_{y'y}y' + f_{y'y'}y'' = 0$$

como en este caso también $f_{y'x} = 0$ y $f_{y'y} = 0$

$$f_{y'y'}y'' = 0$$

De aquí se obtiene

$$\left. \begin{array}{c} y'' = 0 \\ \text{ó} \\ f_{y'y'} = 0 \end{array} \right\}$$

Si $y'' = 0$ entonces

$$y = c_1x + c_2$$

y obtenemos una familia biparamétrica de rectas.

Si, en caso contrario, ocurre $f_{y'y'} = 0$ y esta ecuación tiene una o varias raíces reales $y' = k_i$, entonces integrando

$$y = k_ix + c$$

que es una familia monoparamétrica de rectas, que podemos suponer, está contenida en la familia biparamétrica descrita con anterioridad.. De esta forma, en el caso $f = f(z)$ todas las líneas rectas posibles de la forma

$$y = c_1x + c_2$$

son *estacionarias*.

Ejemplo 4.36 La longitud del arco de una curva es de la forma:

$$L(y) = \int_a^b \sqrt{1 + (y')^2} dx$$

se trata por tanto del caso anterior, donde $f(z) = \sqrt{1 + z^2}$. La ecuación de Euler-Lagrange para este caso particular es

$$\frac{d}{dx} f_{y'} = 0 \Rightarrow \frac{d}{dx} \left(\frac{y'}{\sqrt{1 + (y')^2}} \right) = 0$$

y derivando respecto a x

$$\frac{1 \cdot \sqrt{1 + (y')^2} - \frac{y'}{\sqrt{1 + (y')^2}} \cdot y'}{1 + (y')^2} \cdot y'' = 0 \Leftrightarrow \frac{1}{[1 + (y')^2]^{3/2}} \cdot y'' = 0$$

cuya única solución es

$$y'' = 0 \Leftrightarrow y(x) = c_1 x + c_2$$

Cuando $f = f(x, z)$

En este caso, también ocurre $f_y = 0$ y la ecuación 4.20 toma la forma

$$\frac{d}{dx} f_{y'}(x, y') = 0$$

que tiene como primera integral

$$f_{y'} = C_1$$

Además, como la ecuación de primer orden obtenida no contiene a y , ésta puede integrarse o bien resolviéndola directamente respecto a y' e integrando o bien introduciendo un parámetro escogido en forma adecuada.

Ejemplo 4.37 Como se ha visto al principio de la sección para caracterizar las geodésicas suaves sobre una esfera de radio R había que minimizar el funcional

$$L(Y) = R \int_0^1 \sqrt{\theta'(t)^2 \operatorname{sen}^2 \varphi(t) + \varphi'(t)^2} dt$$

con

$$\varphi(0) = 0$$

$$\theta(1) = 0$$

$$\varphi(1) = \varphi_1$$

A partir de las ecuaciones paramétricas $(\theta(t), \varphi(t))$ podemos obtener una expresión de θ en términos de φ : $\theta = y(\varphi)$, de donde por una parte

$$\theta'(t) = \frac{dy}{d\varphi} \frac{d\varphi}{dt} = y'(\varphi) \varphi'(t) \Rightarrow \frac{\theta'(t)}{\varphi'(t)} = y'(\varphi) \quad (4.24)$$

y por otra

$$\varphi'(t) = \frac{d\varphi(t)}{dt} \Rightarrow \varphi'(t) dt = d\varphi \quad (4.25)$$

por lo que el funcional se puede expresar como

$$L(Y) = R \int_0^1 \sqrt{\theta'(t)^2 \operatorname{sen}^2 \varphi(t) + \varphi'(t)^2} dt = R \int_0^1 \varphi'(t) \sqrt{\left(\frac{\theta'(t)}{\varphi'(t)}\right)^2 \operatorname{sen}^2 \varphi(t) + 1} dt$$

y utilizando 4.24 y 4.25 obtenemos

$$L(y) = R \int_0^{\varphi_1} \sqrt{1 + (y'(\varphi) \operatorname{sen} \varphi)^2} d\varphi$$

Si tomamos ahora $\varphi = x$ el problema es buscar el mínimo de

$$L(y) = R \int_0^{x_1} \sqrt{1 + (y'(x) \operatorname{sen} x)^2} dx$$

sobre el conjunto

$$\mathcal{D}^1 = \{y \in \mathcal{C}^1[0, x_1] : y(x_1) = 0\}$$

En este caso

$$f = f(x, z) = R \sqrt{1 + z^2 \operatorname{sen}^2 x}$$

por tanto

$$f_z(x, z) = \frac{Rz \operatorname{sen}^2 x}{\sqrt{1 + z^2 \operatorname{sen}^2 x}}$$

de manera que las funciones estacionarias y son aquellas para las que sucede

$$\frac{Ry' \operatorname{sen}^2 x}{\sqrt{1 + y'(\varphi)^2 \operatorname{sen}^2 x}} = c$$

siendo c una constante, cuyo valor podemos determinar teniendo en cuenta que esta expresión es válida para todos los valores de x en el intervalo $[0, x_1]$ y tomando en particular $x = 0$

$$\left. \frac{Ry' \operatorname{sen}^2 x}{\sqrt{1 + y'(\varphi)^2 \operatorname{sen}^2 x}} \right|_{x=0} = 0$$

luego

$$c = 0$$

De manera que tiene que cumplirse $y'(x) = 0$ y de ahí

$$y(x) = A$$

con A constante. Esta expresión de $y(x)$ corresponde a los círculos máximos de la esfera.

Cuando $f = f(y, z)$

Si f no depende explícitamente de la variable x y la función $y(x) \in \mathcal{C}^2([a, b])$, la ecuación de Euler-Lagrange 4.20 se reduce a una de primer orden, ya que llamando

$$H(x) = f(y(x), y'(x))$$

al derivar respecto de x , se obtiene

$$H' = \frac{d}{dx} [H(x)] = f_x + f_y \cdot y' + f_{y'} \cdot y''$$

y teniendo en cuenta que $f_x = 0$ y que también se cumple la ecuación de Euler-Lagrange 4.20 entonces

$$H' = \frac{d}{dx} (f_{y'} \cdot y' + f_y \cdot y'') = \frac{d}{dx} (f_{y'} \cdot y')$$

y el segundo miembro es la derivada de un producto, de donde

$$H = f_y' \cdot y' + A$$

con A una constante. Como $H = f$ se tiene la ecuación diferencial de primer orden

$$f - f_{y'} \cdot y' = A \quad (4.26)$$

Ejemplo 4.38 Se aplica el resultado al problema de la braquistócrona formulado al principio. El problema consistía en minimizar la expresión

$$T(y) = \frac{1}{\sqrt{2g}} \int_0^{x_1} \frac{\sqrt{1 + y'(x)^2}}{\sqrt{y(x)}} dx$$

sobre el conjunto de funciones

$$\mathcal{D} = \left\{ 0 \leq y \in \mathcal{C}^1[0, x_1] : y(0) = 0, y(x_1) = y_1; \int_0^{x_1} (y(x))^{-3/2} dx < +\infty \right\}$$

En este caso

$$f(y, z) = \frac{\sqrt{1 + z^2}}{\sqrt{y}}$$

salvo el factor constante $\frac{1}{\sqrt{2g}}$. De aquí obtenemos

$$f_z = \frac{z}{\sqrt{y}\sqrt{1 + z^2}} \Rightarrow f_{y'} = \frac{y'}{\sqrt{y}\sqrt{1 + (y')^2}}$$

Y utilizando la expresión 4.26

$$\frac{\sqrt{1 + (y')^2}}{\sqrt{y}} - \frac{y'}{\sqrt{y}\sqrt{1 + (y')^2}} \cdot y' = k$$

o de forma equivalente

$$\frac{1}{\sqrt{y}\sqrt{1 + (y')^2}} = k = \frac{1}{c}$$

Si se eleva ambos miembros al cuadrado teniendo en cuenta además que $y > 0$

$$y(1 + (y')^2) = c^2 \Leftrightarrow \sqrt{\frac{y}{c^2 - y}} y' = 1 \quad (4.27)$$

Si ahora introducimos el cambio de variable $\theta = \theta(x)$ de forma que

$$y = c^2 \sin^2 \frac{\theta}{2} = \frac{c^2}{2} (1 - \cos \theta) \quad 0 \leq \theta < 2\pi$$

entonces

$$c^2 - y = c^2 \cos^2 \frac{\theta}{2} \quad e \quad y' = c^2 \theta' \sin \frac{\theta}{2} \cos \frac{\theta}{2}$$

Y sustituyendo ambas expresiones en la ecuación diferencial 4.27

$$c^2 \left(\sin \frac{\theta}{2} \right) \theta' = 1 \quad o \quad \frac{c^2}{2} (1 - \cos \theta) \theta' = 1$$

e integrando

$$\left(\frac{c^2}{2} \right) (\theta - \sin \theta) = x - c_1$$

siendo c_1 una constante.

Cambiando $c^2/2$ por c^2 se obtienen las ecuaciones paramétricas de la braquistócrona

$$x = c^2 (\theta - \sin \theta) + c_1$$

$$y = c^2 (1 - \cos \theta)$$

4.5.3 La Segunda Ecuación

Cuando $f = f(x, y, z) \in \mathcal{C}^1([a, b] \times \mathbb{R}^2)$ e y es una solución de clase $\mathcal{C}^1([a, b])$ de la primera ecuación Euler-Lagrange (ecuación 4.20), la integración directa produce

$$\frac{d}{dx} [f_{y'}] = f_y \Rightarrow \int_a^x \frac{d}{dt} [f_{y'}] dt = \int_a^x f_y dt \Rightarrow f_{y'} = \int_a^x f_y dt + A \quad (4.28)$$

con A una constante.

Por otra parte, si $y \in \mathcal{C}^2([a, b])$ y derivamos f respecto a x , se obtiene

$$\frac{d}{dx} [f] = f_x + f_y y' + f_{y'} y'' = f_x + \frac{d}{dx} [f_{y'}] y' + f_{y'} y'' = f_x + \frac{d}{dx} [y' f_{y'}]$$

y por tanto

$$\frac{d}{dx} [f - y' f_{y'}] = f_x$$

o integrando

$$f - y' f_{y'} = \int_a^x f_x(t) dt + c_0$$

donde c_0 es una constante.

Proposición 4.39 Sea

$$J(y) = \int_a^b f(x, y(x), y'(x)) dx$$

con $f \in \mathcal{C}^1([a, b] \times \mathbb{R}^2)$ y

$$\mathcal{D} = \{y \in \mathcal{C}^1[a, b] : y(a) = a_1; y(b) = b_1\}$$

Si $y_0 \in \mathcal{D}$ es una función extremal para J en \mathcal{D} , entonces en $[a, b]$, la función y_0 cumple la segunda ecuación de Euler-Lagrange

$$f - y' f_{y'} = \int_a^x f_x(t) dt + c_0 \quad (4.29)$$

para alguna constante c_0 .

4.6 Condiciones de transversalidad

Utilizaremos ahora el mismo análisis que el descrito anteriormente para encontrar las ecuaciones que deben cumplir las funciones extremales, cuando alguno de los extremos es libre (o ambos), y también cuando existen restricciones isoperimétricas (condiciones integrales de igualdad o desigualdad) en el problema.

4.6.1 Extremos libres

Supongamos de nuevo que la función $f = f(x, y, z) \in \mathcal{C}^1([a, b] \times \mathbb{R}^2)$. Y para cada $y \in \mathcal{Y} = \mathcal{C}^1[a, b]$ definimos el funcional

$$J(y) = \int_a^b f(x, y(x), y'(x)) dx$$

Queremos minimizar $J(y)$ sobre el conjunto

$$\mathcal{D}^b \subseteq \{y \in \mathcal{C}^1[a, b] : y(a) = a_1\}$$

donde ahora el valor de las funciones y no está fijo para $x = b$. Recordemos que la variación Gâteaux para este tipo de funcionales, viene dada por la expresión

$$\delta J(y; v) = \left. \frac{dJ(y + \varepsilon v)}{d\varepsilon} \right|_{\varepsilon=0} = \int_a^b f_y v + f_{y'} v' dx$$

donde se ha utilizado la notación simplificada.

En este caso no podemos aplicar los lemas anteriores, puesto que ahora si v es una dirección admisible, debe ocurrir

$$y + \varepsilon v \in \mathcal{D}^b$$

lo que implica que

$$v(a) = 0$$

pero no condiciona el valor de $v(b)$. Sin embargo, si integramos por partes el segundo término del integrando

$$\int_a^b f_y v + f_{y'} v' dx = \int_a^b f_y v dx + \int_a^b f_{y'} v' dx = \int_a^b f_y v dx + f_{y'} v|_a^b - \int_a^b \frac{d}{dx} (f_{y'}) v dx$$

Después de simplificar obtenemos

$$\delta J(y; v) = \int_a^b \left(f_y - \frac{d}{dx} (f_{y'}) \right) v dx + f_{y'}(b) v(b)$$

Si ahora y_0 es un extremal de $J(y)$ debería ocurrir

$$\delta J(y_0; v) = 0$$

es decir

$$\int_a^b \left(f_y - \frac{d}{dx} (f_{y'}) \right) v dx + f_{y'}(b) v(b) = 0$$

Puesto que $v(b)$ no tiene porqué ser 0, esta variación será cero si y cumple

$$f_y - \frac{d}{dx} (f_{y'}) = 0 \quad (\text{Ecuación de Euler-Lagrange})$$

$$f_{y'}(b) = 0 \quad (\text{Condición de transversalidad})$$

Del mismo modo, se obtiene la condición de transversalidad que debería cumplir un extremal si y es libre en el extremo inferior del intervalo a

$$f_{y'}(a) = 0$$

Para el caso más general en el que la función y que define el funcional no esté condicionada por los extremos, es decir, si $y(a)$ e $y(b)$ son libres, entonces, las condiciones que debe cumplir un extremal serán

$$\begin{aligned} f_y - \frac{d}{dx} (f_{y'}) &= 0 \\ f_{y'}(b) &= 0 \\ f_{y'}(a) &= 0 \end{aligned}$$

4.6.2 Restricciones Integrales

En el caso de que las funciones y que minimizan el funcional

$$J(y) = \int_a^b f(x, y(x), y'(x)) dx$$

deban cumplir restricciones integrales del tipo

$$G_k(y) = \int_a^b g_k(x, y(x), y'(x)) dx = c_k \in \mathbb{R} \quad k = 1, \dots, N \quad (4.30)$$

se ha visto en la primera parte que en este caso si y_0 minimiza $\tilde{J} = J + \sum_{k=1}^N \lambda_k G_k$, entonces y_0 también minimiza J bajo las condiciones $\lambda_k G_k(y) = \lambda_k G_k(y_0)$. Por tanto, si y_0 minimiza a \tilde{J} tendrán que cumplirse las ecuaciones de Euler-Lagrange para su integrando

$$\tilde{J} = \int_a^b \left\{ f(x, y(x), y'(x)) + \sum_{k=1}^N \lambda_k g_k(x, y(x), y'(x)) \right\} dx = \int_a^b \tilde{f}(x, y(x), y'(x)) dx$$

siendo

$$\tilde{f} = f + \sum_{k=1}^N \lambda_k g_k$$

luego

$$\tilde{f}_y - \frac{d}{dx} (\tilde{f}_{y'}) = 0$$

Los valores de λ_k se determinan utilizando las ecuaciones 4.30.

4.7 Bibliografía Básica del Tema

1. *Variational Calculus and Optimal Control. Optimization with Elementary Convexity.* John L. Troutman. **Ed. Springer Verlag (UTM Series)**
2. *Ecuaciones diferenciales y cálculo variacional.* L. Elsgoltz **Ed. MIR (Rubiños-1860, S.A.)**
3. *Introducción al Control Óptimo.* Agustín Jiménez Avello & Eduardo Jiménez Moreno. **Ed. Universidad Politécnica de Madrid (ETSII, Servicio de Publicaciones)**