Apuntes de Análisis Convexo y Optimización

Antonio Cañada Villar

Departamento de Análisis Matemático

Universidad de Granada

MÉTODOS VARIACIONALES. OPTATIVA DE SEGUNDO CICLO DE LA LICENCIATURA EN MATEMÁTICAS, UNIVERSIDAD DE GRANADA

MÉTODOS VARIACIONALES, A. CAÑADA, CURSO 2010/11

En las notas que siguen se recogen de manera esquemática algunas de las ideas fundamentales que se explican en la asignatura "Métodos Variacionales", optativa de segundo ciclo de la Licenciatura en Matemáticas en la Universidad de Granada. Estas notas están en constante proceso de elaboración. Espero que cada alumno encuentre algo de utilidad en ellas.

1. Introducción, motivación y notas históricas: del problema de la braquistocrona a la teoría de puntos críticos.

1.1. Recomendaciones bibliográficas.

Los problemas de máximos y mínimos constituyeron la motivación fundamental para la creación del Cálculo Diferencial e Integral a finales del siglo XVII. El origen del Cálculo Diferencial e Integral se atribuye fundamentalmente a Newton y Leibniz.

El significativo trabajo de Leibniz titulado "Nova methodus pro maximis et minimis itemque tangentibus" que puede considerarse como el nacimiento formal del Cálculo, se publicó en 1684 y las palabras con las que Leibniz describe el título de esta obra nos dan una idea muy precisa del contenido: nuevo método para el estudio de máximos y mínimos de funciones usando la noción de tangente. En este tratado se proporciona un método muy preciso para el cálculo de máximos y mínimos de funciones concretas que surgen en las aplicaciones más diversas. Dicho método usa la noción de derivada, punto crítico, derivadas de orden superior, etc.

Un buen resumen de la evolución histórica del cálculo de variaciones se puede ver en los siguientes artículos:

Erwin Kreyszig: On the Calculus of Variations and Its Major Influences on the Mathematics of the First Half of Our Century. Part I. American Mathematical Monthly, volumen 101, no. 7, 674-678, (1994).

Erwin Kreyszig: On the Calculus of Variations and Its Major Influences on the Mathematics of the First Half of Our Century. Part II. American Mathematical Monthly, volumen 101, no. 9, 902-908, (1994).

La introducción del libro de texto recomendado [3] ofrece información histórica de interés.

Por cierto, la revista mencionada, American Mathematical Monthly, es muy recomendable para los alumnos de los últimos cursos de la licenciatura en Matemáticas. Tiene secciones muy diversas: artículos, problemas, reseñas de libros, etc. El nivel es, en general, asequible. Os animo a que la consultéis habitualmente.

Como veremos a continuación, los problemas que trata el cálculo de variaciones se formularon inemdiatamente después del nacimiento del Cálculo.

1.2. **Algunos problemas significativos del Cálculo de Variaciones.** .

Según M. Kline ([10]) el primer problema significativo del Cálculo de Variaciones fue propuesto y resuelto por Newton en el libro II de sus "Principia". Estudiando la forma que debía tener una superficie de revolución, moviéndose en el agua (o cualquier otro fluido) a velocidad constante a lo largo de su eje para ofrecer una resistencia mínima al movimiento, se planteó el estudio de mínimos de funcionales de la forma

$$\Phi(y) = \int_a^b \frac{y(x)(y'(x))^3}{1+(y'(x))^2} \tag{1.1}$$

donde la función $y(x)$ pertenece a un espacio apropiado de funciones de clase $C^1[a,b]$. Por ejemplo, podemos pensar que $y(\cdot) \in Y$, donde

$$Y = \{y : [a,b] \to \mathbb{R} : \ y \in C^1[a,b], \ y(a) = y_1, y(b) = y_2 \ \}$$

donde $(a, y_1), (b, y_2)$ son dados.

Ya apreciamos la primera diferencia fundamental respecto de los problemas de máximos y mínimos estudiados por el alumno en la licenciatura: en (1.1) la "variable independiente" $y(x)$ pertenece a un conjunto adecuado de funciones y no a un subconjunto del espacio euclídeo finito dimensional. Tengamos en cuenta que los espacios de funciones son, en general, espacios vectoriales de dimensión infinita. Os recomiendo la lectura y consulta del artículo: G. Buttazzo y B. Kawohl: On Newton's problem of minimal resistance, Mathematical Intelligencer, volumen 15, 7-12, (1992). Por cierto, esta revista, Mathematical intelligencer, es también muy recomendable para los alumnos.

De todas formas, suele considerarse que lo que se entiende hoy en día por Cálculo de Variaciones, nació con la proposición de Johann Bernoulli,en 1.696, del problema de la braquistocrona (según algunos autores, este problema fue formulado con anterioridad por Galileo en 1638). Puede describirse así: una masa puntual se desliza, por acción de la gravedad, desde un punto A del plano , hasta otro punto B, a través de alguna curva. El tiempo empleado para ir desde A hasta B depende de la curva elegida. ¿Para qué curva se tendrá que el tiempo es mínimo? Puede verse ([17]) que esto exige el estudio de funcionales de la forma

$$\int_a^b \left(\frac{1 + y'(x)^2}{y(x)} \right)^{1/2} dx \tag{1.2}$$

donde, como en (1.1), la variable independiente $y(\cdot)$, pertenece a un conjunto adecuado de funciones.

Contrariamente a lo que puede sugerir nuestra intuición, la solución de este problema no es ni el segmento rectilíneo que une A con B, ni ningún arco de circunferencia que pase por A y B. La solución, un arco de cicloide, fue encontrada por diferentes matemáticos: Jakob Bernoulli, Newton, Leibniz, L'Hopital, etc.

En los ejemplos citados con anterioridad, la variable independiente es una función $y(x)$. Hay otros problemas, donde de manera natural surge el estudio de funcionales cuya variable independiente pueden ser varias funciones. Incluso, se pueden plantear problemas que recuerdan a los máximos y mínimos condicionados (multiplicadores de Lagrange). Por ejemplo, esto ocurre en el llamado problema isoperimétrico que describimos a continuación. Dado un número real positivo L, se trata de estudiar la cuestión siguiente: de todas las curvas cerradas del plano de longitud dada $L > 0$, ¿cuál es la que encierra mayor área?

Si escribimos la curva mediante sus ecuaciones paramétricas

$$x = x(s) \tag{1.3}$$

$$y = y(s) \tag{1.4}$$

(donde s es, por ejemplo, el parámetro arco), entonces es conocido que el área encerrada por la misma viene dada por:

$$A = \Phi(x, y) = \frac{1}{2} \int_0^L \big(x(s)y'(s) - x'(s)y(s) \big)\, ds \tag{1.5}$$

Por tanto, se trata de hacer máxima una expresión donde la variable es un par de funciones $x(s), y(s)$. Más precisamente, el par de funciones $x(\cdot), y(\cdot)$ pertenecen al espacio de funciones de clase $C^1[0, L]$ y, además, $x(0) = x(L), y(0) = y(L)$ (la curva es cerrada). Por otra parte, al ser s el parámetro arco se ha de cumplir la restricción

$$(x'(s))^2 + (y'(s))^2 = 1, \ \forall \ s \in [0, L].$$

Esto es pues un problema de máximos y mínimos condicionados.

Puede probarse ([7]) la relación

$$L^2 - 4\pi A \geq 0 \tag{1.6}$$

y que la igualdad se da sólo para cualquier circunferencia de longitud L. Un método elegante y sencillo para estudiar este problema usa algunos hechos básicos de la teoría de Series de Fourier ([5]).

El problema puede plantearse también en dimensiones superiores $n \geq 3$, probándose que si $B \subset \mathbb{R}^n$, es un conjunto con algunas restricciones adicionales, entonces ([7])

$$[L(\partial B)]^n - n^n \omega_n [M(B)]^{n-1} \geq 0 \tag{1.7}$$

donde ω_n es el volumen de la bola unidad $n-$dimensional, es decir, $\omega_n = \int_{B_{\mathbb{R}^n}(0;1)} 1$ y $[M(B)]$ es la medida de Lebesgue $n-$dimensional de B. Por su parte, $L(\partial B)$ es la medida de Lebesgue $n-1$-dimensional de la frontera de B. Además la igualdad se da en la expresión anterior si y sólo si B es una bola.

Un ejemplo de especial significación, desde el punto de vista de la física, lo constituye el principio de Hamilton. Vamos a ver que proporciona una formulación variacional (en términos de máximos y mínimos, o al menos, puntos estacionarios o críticos), de la segunda ley de Newton.

Supongamos que una partícula de masa 1 se desplaza en el espacio euclídeo tridimensional desde un punto dado p_1 a otro p_2 en el intervalo de tiempo $[t_1, t_2]$, bajo la acción de una fuerza $F : \mathbb{R}^3 \longrightarrow \mathbb{R}^3$ que depende sólo de la posición de la partícula. En este caso, $F = F(x),\ x = (x_1, x_2, x_3)$.

Si $x : [t_1, t_2] \longrightarrow \mathbb{R}^3, t \to x(t)$, es la trayectoria de la partícula, entonces la segunda ley de Newton ($masa \times aceleracion = fuerza$) nos proporciona la relación

$$x''(t) = F(x(t)),\ x(t_1) = p_1, x(t_2) = p_2 \tag{1.8}$$

(1.8) es una ecuación diferencial ordinaria de segundo orden.

El Cálculo de Variaciones nos proporciona otra interpretación de (1.8), cercana a problemas de máximos y mínimos cuando la fuerza F es conservativa, esto es, deriva de un potencial $U : \mathbb{R}^3 \to \mathbb{R}$. Esto quiere decir que $\nabla U(y) = F(y), \quad \forall y \in \mathbb{R}^3$, donde $\nabla U(y) = \left(\frac{\partial U(y)}{\partial y_1}, \frac{\partial U(y)}{\partial y_2}, \frac{\partial U(y)}{\partial y_3}\right) = (F_1(y), F_2(y), F_3(y))$.

Definamos el conjunto $\mathcal{A}$ de "trayectorias posibles entre los puntos p_1 y p_2 " de la forma siguiente:

$$\mathcal{A} = \left\{ y : [t_1, t_2] \longrightarrow \mathbb{R}^3 \quad / \quad y(t_1) = p_1, y(t_2) = p_2, y \in \mathcal{C}^1[t_1, t_2] \right\} \tag{1.9}$$

Dado un elemento $y \in \mathcal{A}$, su energía cinética como función del tiempo t viene dada por

$$(M\dot{y})(t) = \frac{1}{2}\left(\dot{y}_1(t)^2 + \dot{y}_2(t)^2 + \dot{y}_3(t)^2\right) \tag{1.10}$$

y su energía potencial viene expresada por $-U(y(t))$ donde U es el potencial mencionado.

La lagrangiana de la trayectoria y se define como

$$L(y(t)) = M(\dot{y}(t)) + U(y(t)) \tag{1.11}$$

El funcional, llamado acción integral de Hamilton, se define como

$$\begin{gathered} \Phi : \mathcal{A} \longrightarrow \mathbb{R} \\ \Phi(y) = \int_{t_1}^{t_2} L(y(t))\ dt \end{gathered} \tag{1.12}$$

que hace corresponder a cada trayectoria posible, y, un número real,$\Phi(y)$.

¿Qué relación existe entre el funcional dado en (1.12) y la ecuación diferencial (1.8)? Por una parte, el funcional dado hace corresponder, a cada trayectoria posible y, un número real $\Phi(y)$. Por otra, (1.8) es una ecuación diferencial ordinaria. No parece sencillo, ni intuitivo, qué relación puede haber entre ambos conceptos.

El principio de Hamilton afirma: "la trayectoria $x \in \mathcal{A}$ que sigue la partícula bajo la acción de la fuerza F para ir de p_1 a p_2 es aquella que hace estacionaria la acción integral Φ".

¿Qué es eso de hacer estacionaria la acción integral?. Recordemos que dada una función $H : \mathbb{R}^n \to \mathbb{R}$, un punto estacionario de la misma es cualquier punto $x \in \mathbb{R}^n$ para el que $H'(x) = 0$ (equivalentemente, todas las derivadas parciales de H, en el punto x son cero; es decir $\frac{\partial H(x)}{\partial x_i} = 0,\ 1 \leq i \leq n.$) Ahora quizás no tenga sentido, o quizás sea un concepto nada trivial, la noción de "derivada parcial de la acción integral de Hamilton". Esto puede ser subsanado con el concepto de derivada direccional, más amplio que el de derivada parcial.

Un elemento $x \in \mathcal{A}$ se dice estacionario cuando la primera variación de Φ en x es cero. La primera variación del funcional Φ, en el punto x y en la dirección h se define como

$$\delta\Phi(x; h) = \frac{d}{d\tau}\Phi(x + \tau h)_{|\tau=0} \tag{1.13}$$

donde

$$h \in \left\{ g : [t_1, t_2] \longrightarrow \mathbb{R}^3 \quad / \quad g(t_1) = 0, g(t_2) = 0, g \in \mathcal{C}^1[t_1, t_2] \right\} = B \tag{1.14}$$

Que la primera variación sea cero significa que

$$\delta\Phi(x;h) = 0 \quad \forall h \in B. \tag{1.15}$$

Veremos en el curso que (1.15) no es sino la la segunda ley de Newton (1.8) expresada en forma variacional. En realidad veremos que, para trayectorias $x \in C^2[t_1, t_2]$, tales que $x(t_1) = p_1, x(t_2) = p_2$, (1.8) es equivalente a (1.15).

Otras muchas leyes de la Física se expresan también en forma variacional. Todas ellas responden a la filosofía marcada por las frases siguientes:

(1) "*En la Naturaleza no se realiza ningún proceso superfluo o innecesario (Olympiodorus, siglo VI)*"
(2) "*En un medio no homogéneo, formado por dos homogéneos y separados por una recta, un rayo luminoso que va desde un punto del primer medio a otro del segundo, lo hace de tal forma que minimiza el tiempo empleado en su recorrido (Fermat, siglo XVII)*"
(3) "*Puesto que el Universo es perfecto y fue creado por el Creador más sabio, nada ocurre en él, sin que esté presente alguna ley de máximo o mínimo (Euler, siglo XVIII)*"
(4) "*Una partícula que se mueve en el espacio bajo la acción de una fuerza conservativa, lo hace de tal forma que su trayectoria hace mínima, o al menos estacionaria, la acción integral (Hamilton, siglo XIX)*"

Detengámonos en un ejemplo concreto, la ley de la refracción que afirma lo siguiente:
Supongamos un medio no homogéneo, formado por dos medios homogéneos I y II separados por una recta r. Sean A y B dos puntos dados de I y II respectivamente. Entonces un rayo luminoso que va de A hasta B lo hace a través de una trayectoria que es rectilínea en cada medio considerado y tal que si O es el punto de la recta r por el que pasa, se tiene que el cociente entre el seno del ángulo de incidencia φ_1 y el seno del ángulo de refracción φ_2, es igual a $\frac{v_1}{v_2}$, siendo v_1 y v_2, respectivamente, las velocidades de propagación de la luz en I y II. Es decir

$$\frac{\text{sen}\varphi_1}{\text{sen}\varphi_2} = \frac{v_1}{v_2} \tag{1.16}$$

Una demostración variacional de la ley anterior, puede obtenerse usando el principio de Fermat (siglo XVII) que afirma:
en la situación anteriormente descrita, un rayo luminoso que va de A hasta B, lo hace a través de una trayectoria tal que el tiempo empleado en su recorrido es mínimo.

Usando este principio y eligiendo convenientemente los ejes de coordenadas tal que $A = (0, a), B = (d, -b)$, entonces si $O = (x, 0)$, tenemos que el tiempo empleado en la trayectoria es

$$t(x) = \frac{(a^2 + x^2)^{1/2}}{v_1} + \frac{(b^2 + (d - x)^2)^{1/2}}{v_2}$$

Es trivial comprobar que existe un único punto $x_0 \in [0, d]$ tal que $t'(x_0) = 0$. Además, $t''(x) > 0, \ \forall \ x \in [0, d]$. Esto prueba que existe un único punto $x_0 \in [0, d]$ tal que $t(x_0) = \min_{[0,d]} \ t$. Calculando $t'(x_0)$, es trivial que la relación $t'(x_0) = 0$ es equivalente a (1.16).

Terminamos la descripción de problemas concretos, con el denominado Principio de Dirichlet, donde como variables independientes aparecen de manera natural, funciones de varias variables reales (hasta ahora sólo han aparecido funciones de una variable).

Sea Ω un dominio (subconjunto abierto y conexo) acotado de $\mathbb{R}^n$. Consideremos el problema de contorno

$$(1.17) \qquad \left. \begin{array}{rl} \Delta u(x) = 0 & x \in \Omega \\ u(x) = f(x) & x \in \partial\Omega \end{array} \right\}$$

donde Δ es el operador Laplaciano definido como

$$(1.18) \qquad \Delta u(x) = \frac{\partial^2 u(x)}{\partial x_1^2} + ... + \frac{\partial^2 u(x)}{\partial x_n^2}$$

Problemas del tipo anterior (estudio de la existencia de funciones armónicas en Ω que tomen valores prefijados, dados por la función f, en la frontera de Ω) surgen, por ejemplo, en electrostática.

Como en el caso de la segunda ley de Newton, la ecuación que aparece en (1.17) es una ecuación diferencial de segundo orden, pero en este caso se trata de una ecuación en derivadas parciales. Al tratar de conectar (1.17) con problemas de máximos y mínimos, surge de manera natural el llamado funcional de energía:

$$(1.19) \qquad \Phi(u) = \int_\Omega |\nabla u(x)|^2 \, dx = \int_\Omega \left[\left(\frac{\partial u(x)}{\partial x_1} \right)^2 + ... + \left(\frac{\partial u(x)}{\partial x_n} \right)^2 \right] dx$$

que está definido sobre el conjunto de funciones:

$$(1.20) \qquad \mathcal{A} = \left\{ u : \bar{\Omega} \longrightarrow \mathbb{R} \quad / \quad u \in \mathcal{C}^1(\overline{\Omega}), u_{|\partial\Omega} = f \right\}$$

En electrostática, u es el potencial eléctrico y $\Phi(u)$ la energía. Obsérvese la analogía entre este funcional y el funcional Φ dado por la acción integral de Hamilton en (1.12) (ahora la fuerza F y por tanto, el potencial U son las funciones idénticamente cero).

Los puntos estacionarios del funcional Φ son, en un sentido que se precisará en el curso, soluciones del problema (1.17) Este es el principio de Dirichlet.

1.3. **Las condiciones necesarias de Euler-Lagrange y Legendre.**

Como resumen de lo dicho hasta ahora, nos encontraremos con el problema de minimizar (o maximizar), o al menos encontrar puntos estacionarios (por cierto, un concepto que habrá que definir rigurosamente), de funcionales Φ de la forma

$$\Phi(y) = \int_a^b L(t, y(t), ..., y^{n)}(t))\ dt$$

donde $y : [a, b] \to \mathbb{R}$ de clase $C^n[a, b]$ (sistemas con un grado de libertad). Este es el caso del problema de Newton (1.1), descrito con anterioridad, o el problema de la braquistocrona (1.2). En estos dos ejemplos $n = 1$.

De manera más general, $y : [a, b] \to \mathbb{R}^m$ de clase $C^n[a, b]$ (sistemas con m grados de libertad). Por ejemplo, en el caso de problema isoperimétrico (1.5) tenemos $n = 1, m = 2$ y en el caso del principio de Hamilton para el funcional Φ dado en (1.12) tenemos $n = 1, m = 3$. Además, la función y y sus derivadas hasta el orden n, satisfacen en general algunas condiciones adicionales en la frontera del intervalo $[a, b]$.

Trataremos también el caso de funciones de varias variables (sistemas con infinitos grados de libertad) como en el caso del funcional (1.19), relacionado con el principio de Hamilton. Concretamente funcionales de la forma

$$\Phi(u) = \int_\Omega L(x, u(x), ..., D^\alpha u(x))\ dx,$$

siendo Ω un dominio acotado de $\mathbb{R}^n$. Aquí α es un multiíndice $\alpha = (\alpha_1, \alpha_2, \ldots, \alpha_n)$ tal que $\alpha_i \in \mathbb{N} \cup \{0\},\ 1 \leq i \leq n$ y

$$D^\alpha u(x) = \frac{\partial^{\alpha_1 + \ldots + \alpha_n} u(x)}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}$$

Por ejemplo, en el caso del funcional (1.19) los multiíndices α vienen dados por los vectores canónicos de $\mathbb{R}^n$. Además, es usual que la función u verifique algunas condiciones en la frontera de Ω. Por ejemplo, en el principio de Hamilton $u(x) = f(x), \forall\ x \in \partial\Omega$.

Como en la teoría de máximos y mínimos para funcionales $\Phi : D \to \mathbb{R},\ y \to \Phi(y)$, donde D es un subconjunto del espacio euclídeo $\mathbb{R}^k$, las principales cuestiones que hemos de resolver se centran fundamentalmente en el establecimiento de condiciones necesarias y condiciones suficientes que nos permitan reconocer un punto de mínimo o al menos un punto estacionario (si es que tiene sentido) del funcional Φ. Sabemos que esto depende no sólo del funcional Φ, sino en gran medida del subconjunto D.

En el caso finito dimensional, donde $D \subset \mathbb{R}^k$, si D es abierto, dichas condiciones se expresan en términos de $\Phi'(x)$ y de $\Phi''(x)$ (o derivadas de orden superior a dos). No obstante, en este caso el problema suele ser, con frecuencia, probar que Φ tiene mínimo. Si D es compacto, no suele ser problema probar que Φ tiene mínimo. Si, además, su frontera se expresa en términos "buenos", es útil el teorema de los multiplicadores de Lagrange.

Todas estas son las ideas que flotan en el ambiente del Cálculo de Variaciones. No obstante, la situación ahora es bastante más complicada, puesto que en general tratamos con problemas donde la variable independiente pertenece a espacios de dimensión infinita. ¿Tiene sentido ahora la definición de derivada parcial, o derivada direccional?, ¿cómo podemos definir la noción de derivada, y la noción de derivadas de orden superior a uno?...

En lo que se refiere al establecimiento de condiciones necesarias, un avance fundamental fue dado por Euler (¿cómo no?) en 1744, año en que se publicó su famoso libro sobre curvas geodésicas. Para funcionales de la forma

$$\Phi(y) = \int_a^b f(t, y(t), y'(t)) \, dt$$

la condición necesaria para que un elemento (en este caso una función) y_0 sea un mínimo es

$$f_y(t, y_0(t), y_0'(t)) - \frac{d}{dt} f_{y'}(t, y_0(t), y_0'(t)) = 0, \ \forall \ t \in [a, b] \tag{1.21}$$

que es una ecuación diferencial ordinaria de segundo orden (como la de la segunda ley de Newton). Esta es la condición similar a que la primera derivada sea cero en el caso finito-dimensional. Dependiendo del tipo de problema que se considere la anterior ecuación ha de completarse con alguna condición sobre la función y_0 en la frontera $t = a$ y $t = b$ del intervalo $[a, b]$. Como veremos en el curso, esta condición en la frontera implica las cantidades $y_0(a), y_0(b), y_0'(a), y_0'(b)$, etc, dependiendo de que en el problema original considerado los extremos de la función se consideren fijos, variables, etc.

Algunas ideas de Euler fueron posteriormente desarrolladas en profundidad por Lagrange. En particular, en su libro de 1760-61, sobre "un nuevo método para determinar los máximos y los mínimos de fórmulas integrales" llegó a la misma conclusión que Euler, sobre la ecuación (1.21) usando un camino diferente: Lagrange consideró las variaciones

$$J(y_0, \eta, \varepsilon) = F(y_0 + \varepsilon\eta) \tag{1.22}$$

para una función η fija y ε variable. Así llegó a la misma ecuación (1.21), que por este motivo es conocida con el nombre de ecuación de Euler-Lagrange. Parece ser que el nombre "Cálculo de Variaciones" también viene de esta idea de Lagrange.

La ecuación de Euler-Lagrange es una condición necesaria de mínimo que implica la derivada de orden uno. Partiendo de (1.22) y considerando la "segunda variación" $J_{\varepsilon\varepsilon}$, introducida por Legendre en 1786 ([10], [11]), éste demostró que en un punto de mínimo y_0 ha de cumplirse la condición necesaria

$$fy'y'(t, y_0(t), y_0'(t)) \geq 0, \ \forall \ t \in [a, b]. \tag{1.23}$$

Esta es una condición similar al hecho de que la derivada segunda ha de ser semidefinida positiva en el caso finito-dimensional.

Respecto de las posibles condiciones suficientes, hubo que esperar 50 años hasta que Jacobi introdujo la llamada hoy en día condición de Jacobi, que junto con la ecuación de Euler-Lagrange y (1.23) permiten demostrar que el punto y_0 es un mínimo. No obstante, dicha condición no es fácil de comprobar, como veremos más adelante.

1.4. **Teoría de puntos críticos.**

Como hemos tenido ocasión de comentar, la segunda ley de Newton (1.8), que viene dada por una ecuación diferencial ordinaria de segundo orden, puede expresarse en términos variacionales de la forma (1.15), que no refleja otra realidad sino que todas las derivadas direccionales del funcional Φ (definido en (1.12)) en el punto x son cero. En el caso finito dimensional, si hablamos de funcionales $\Phi : \mathbb{R}^n \to \mathbb{R}$ que sean derivables, la manera más fácil de conseguir que todas las derivadas direccionales de Φ en x sean cero es que la derivada de Φ en x, $\Phi'(x)$ sea cero. Ahora estamos hablando no de funcionales definidos en espacios normados de dimensión finita, $\mathbb{R}^n$, sino de funcionales Φ definidos en espacios normados de dimensión infinita E. La teoría de puntos críticos trata sobre la ecuación

$$\Phi'(u) = 0$$

para funcionales derivables $\Phi : E \to \mathbb{R}$, derivables (en un sentido que se ha de precisar), donde E es cualquier espacios normado de dimensión infinita.

En este curso se proporcionará una introducción a la teoría de puntos críticos, y se expondrán algunos resultados importantes referentes tanto al llamado método directo (en el que se trata de probar la existencia de mínimo global del funcional Φ dado), como a los métodos min-max, de desarrollo relativamente reciente. En particular, dedicaremos una atención especial a dos teoremas relativamente modernos, pero ya clásicos: el Teorema del paso de montaña de Ambrosetti y Rabinowitz (1971) y el Teorema del punto de silla de Rabinowitz (1978).

Según mi opinión, los resultados seleccionados son ejemplos muy representativos de las diferentes posibilidades del método variacional, pudiéndose, con la ayuda de la bibliografía recomendada, estudiar otros de los muchos que se conocen en la actualidad.

Un comentario más.

En un curso de esta naturaleza, de carácter introductorio sobre métodos variacionales, pueden obtenerse los resultados fundamentales sin recurrir a nociones de diferenciabilidad en espacios de dimensión infinita (véase, por ejemplo [7], [18]). No obstante, en mi opinión, es muy instructivo hacerlo desde este punto de vista, puesto que el desarrollo del cálculo diferencial en espacios de dimensión infinita puede hacerse muy similarmente al caso de dimensión finita; eso sí hay que hacer un hincapié especial en las diferencias básicas entre ambos casos, lo que es muy instructivo para el alumno. Además, este punto de vista hace necesaria la introducción de diversos conceptos (espacios normado, de Hilbert, aplicaciones lineales, bases hilbertianas, etc.) que permite al alumno tener una formación no sólo en este tema sino también en otras disciplinas relacionadas con Ecuaciones Diferenciales y Análisis Funcional. Es indudable que estos conocimientos pueden serle de gran utilidad en otras disciplinas que curse, e incluso para su posible iniciación en la investigación actual.

2. Cálculo diferencial en espacios normados

2.1. **Derivabilidad en espacios normados de dimensión arbitraria.** A continuación introducimos brevemente algunas nociones sobre derivabilidad en espacios de dimensión arbitraria. Puede consultarse [3] y [8] para los detalles y para profundizar en el tema.

La idea básica para extender la noción de derivada a espacios de dimensión infinita está basada en las siguientes reflexiones escalonadas:

(1) Si $f : (a,b) \to \mathbb{R}$ y $x_0 \in (a,b)$, entonces f es derivable en x_0 si y solamente si existe el límite

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0) \tag{2.1}$$

en cuyo caso, a tal límite se le llama derivada de f en x_0, $f'(x_0) \in \mathbb{R}$. (2.1) es equivalente a

$$\lim_{h\to 0} \left| \frac{f(x_0+h)-f(x_0)}{h} - f'(x_0) \right| = 0. \tag{2.2}$$

(2) Si $f : (a,b) \to \mathbb{R}^m$ y $x_0 \in (a,b)$, entonces todo es lo mismo que en el caso anterior, salvo que, ahora, $f'(x_0) \in \mathbb{R}^m$. Más precisamente, f es derivable en x_0 si y solamente si existe el límite

$$\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0) \tag{2.3}$$

en cuyo caso, a tal límite se le llama derivada de f en x_0, $f'(x_0) \in \mathbb{R}^m$. (2.3) es equivalente a

$$\lim_{h\to 0} \left\| \frac{f(x_0+h)-f(x_0)}{h} - f'(x_0) \right\| = 0. \tag{2.4}$$

donde $\|\cdot\|$ indica cualquier norma en el espacio normado finito dimensional $\mathbb{R}^m$ (recordemos que al ser $\mathbb{R}^m$ de dimensión finita, todas las normas son equivalentes).

(3) Si $f : \Omega \to \mathbb{R}$, donde Ω es un subconjunto abierto de $\mathbb{R}^n$, y $n > 1$, la anterior definición no es válida, puesto que no podemos dividir por $h \in \mathbb{R}^n$. La extensión adecuada de derivada no la constituye la mera existencia de derivadas parciales, puesto que es conocido que una función puede tener derivadas parciales y no ser continua.

Para proceder a la definición en este caso, es necesario obtener una caracterización de las situaciones presentadas en los dos primeros apartados, que pueda extenderse a situaciones generales. En este sentido, el límite (2.1) es un número real o más generalmente, en el caso de (2.3), un vector de $\mathbb{R}^m$. En ambos casos, el vector $f'(x_0)$ define una aplicación lineal (automáticamente continua) $L : \mathbb{R} \to \mathbb{R}^m$

dada por $L(h) = f'(x_0)h, \ \forall\, h \in \mathbb{R}$ y, en ambos casos, podemos decir que f es derivable en x_0 si

$$\lim_{h\to 0} \frac{|f(x_0+h) - f(x_0) - L(h)|}{\|h\|_{\mathbb{R}^n}} = 0 \tag{2.5}$$

donde $\|\cdot\|_{\mathbb{R}^m}$ es cualquier norma en $\mathbb{R}^m$. Estos reflexiones se ven avaladas porque, como puede comprobarse fácilmente, sólo puede existir una aplicación lineal $L : \mathbb{R} \to \mathbb{R}^m$ que satisfaga (2.5).

De acuerdo con las anteriores ideas, si $f : \Omega \to \mathbb{R}$, donde Ω es un subconjunto abierto de $\mathbb{R}^n$, $n > 1$, y $x_0 \in \Omega$, diremos que f es derivable en x_0 si y solamente si existe alguna aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}$ (automáticamente continua) tal que

$$\lim_{h\to 0} \frac{|f(x_0+h) - f(x_0) - L(h)|}{\|h\|_{\mathbb{R}^n}} = 0 \tag{2.6}$$

(4) En la situación general finito dimensional, donde $f : \Omega \to \mathbb{R}^m$, y Ω es un abierto de $\mathbb{R}^n$, diremos que f es derivable en x_0 si y solamente si existe alguna aplicación lineal $L : \mathbb{R}^n \to \mathbb{R}^m$ (automáticamente continua) tal que

$$\lim_{h\to 0} \frac{\|f(x_0+h) - f(x_0) - L(h)\|_{\mathbb{R}^m}}{\|h\|_{\mathbb{R}^n}} = 0 \tag{2.7}$$

Cuando intentemos dar el salto a dimensión infinita, la única novedad es que, ahora, una aplicación lineal no es necesariamente continua y por tanto, hemos de exigirlo en la definición (véase [4]).

Definición 2.1. (Derivada de Fréchet).

Sean $(X, \|\cdot\|_X)$, $(Y, \|\cdot\|_Y)$ espacios normados reales, $\Phi : \Omega \longrightarrow Y$ donde Ω es un abierto de X y $x_0 \in \Omega$.

Se dice que Φ es derivable (según Fréchet) en el punto $x_0 \in \Omega$, si existe alguna aplicación lineal y continua $L : X \longrightarrow Y$ tal que

$$\lim_{h\to 0} \frac{\|\Phi(x_0+h) - \Phi(x_0) - L(h)\|_Y}{\|h\|_X} = 0 \tag{2.8}$$

Como primeras propiedades de esta definición, podemos indicar las siguientes:

Proposición 2.2. (1) *L, si existe, es única. Usaremos la notación $L = \Phi'(x_0)$.*

(2) *La derivabilidad es un concepto que no cambia, si se toman otras normas equivalente. Más precisamente, si Φ es derivable en x_0 y tomamos normas equivalentes en X e Y respectivamente, la derivada es la misma.*

(3) *Si* $\exists\ \Phi'(x_0)$, *entonces* Φ *es continua en* x_0.

(4) *Sea* $\Phi : X \longrightarrow Y$ *lineal y continua (recordemos que si* Φ *es lineal, esto es equivalente a que exista alguna constante* $k \geq 0$ *tal que* $\|\Phi(x)\|_Y \leq k\|x\|_X,\ \forall\ x \in X$*). Entonces se tiene* $\Phi'(x_0) = \Phi, \quad \forall x_0 \in X$.

(5) *Sea* $\Phi : X \times Z \to Y$ *bilineal y continua (recordemos que si* Φ *es bilineal, esto es equivalente a que exista alguna constante* $k \geq 0$ *tal que* $\|\Phi(x,z))\|_Y \leq k\|x\|_X\|z\|_Z,\ \forall\ (x,z) \in X \times Z$*). Entonces existe* $\Phi'(x_0,z_0), \quad \forall (x_0,z_0) \in X \times Z$ *y* $\Phi'(x_0,z_0)(h_1,h_2) = \Phi(x_0,h_2) + \Phi(h_1,z_0)$. *En particular, el producto escalar en cualquier espacio de Hilbert es un ejemplo de ello.*

(6) *Si* H *es un espacio de Hilbert real con producto escalar* $\langle \cdot \rangle$, $A : H \longrightarrow H$ *es lineal y continua y* $\Phi : H \longrightarrow \mathbb{R}$, *viene definida por* $\Phi(x) = \langle Ax, x\rangle$, *entonces existe* $\Phi'(x_0),\ \forall x_0 \in H$ *y* $\Phi'(x_0)(h) = \langle Ax_0, h\rangle + \langle x_0, Ah\rangle$.

En particular, si $A = I$, *entonces* $\Phi(x) = \|x\|^2$ *y* $\Phi'(x_0)(h) = 2\langle x_0, h\rangle,\ \forall\ x_0, h \in H$.

(7) *Si* $\Phi : \Omega \to \mathbb{R}$, *donde* Ω *es un subconjunto abierto de algún espacio de Hilbert* H, *y* Φ *es derivable en* $x_0 \in \Omega$, *entonces el teorema de representación de Riesz garantiza la existencia de un único elemento* $\nabla\Phi(x_0) \in H$, *tal que* $\Phi'(x_0)(h) =< \nabla\Phi(x_0), h >,\ \forall\ h \in H$. *A* $\nabla\Phi(x_0)$ *se le denomina gradiente de* Φ *en* x_0.

Un operador $\Psi : \Omega \to H$ *se dice que es un operador gradiente si existe algún operador* $\Phi : \Omega \to \mathbb{R}$ *tal que* $\Psi(x) = \nabla\Phi(x),\ \forall\ x \in \Omega$. *En este caso, a* Φ *se le llama potencial de* Ψ.

Como en el caso finito dimensional (las demostraciones son idénticas), puede comprobarse inmediatamente que la derivada es un operador lineal en el siguiente sentido: la derivada de la suma es la suma de derivadas, la derivada del producto de un escalar por una función es el escalar por la derivada de la función, etc. y que es válida la regla de la cadena ([3]).

Si $\Phi : \Omega \longrightarrow Y$ es tal que existe su derivada de Fréchet en todos los puntos de Ω, entonces se puede considerar la aplicación derivada

$$\begin{aligned} \Phi' : \Omega &\longrightarrow \mathcal{L}(X,Y) \\ x &\longrightarrow \Phi'(x) \end{aligned} \tag{2.9}$$

Aquí, $\mathcal{L}(X,Y)$ indica el espacio normado real formado por las aplicaciones lineales y continuas de X en Y, dotado de la norma usual. Recordemos que si $L \in \mathcal{L}(X,Y)$, entonces existe alguna constante $k \geq 0$ tal que $\|L(x)\|_Y \leq k\|x\|_X,\ \forall\ x \in X$. En este caso, $\|L\|_{\mathcal{L}}$ es el ínfimo (mínimo) de los números k que verifican la relación anterior. También $\|L\|_{\mathcal{L}(X,Y)} = \sup_{x\in X, \|x\|_X \leq 1} \ \|L(x)\|_Y$.

Tiene así sentido hablar de la derivabilidad de Φ'. Si existe $(\Phi')'(x_0)$ se dirá que Φ posee derivada segunda en el punto $x_0 : \Phi''(x_0) \in \mathcal{L}(X, \mathcal{L}(X,Y)$. Por ejemplo, es fácil comprobar que si $L \in \mathcal{L}(X,Y)$, entonces $L'' = 0$ en todo punto. Si $X = C([0,1],\mathbb{R})$, y $\Phi : X \to X$ está definida como $\Phi(x)(t) = x^2(t)$, entonces $(\Phi''(x))(h)(k) = 2hk$, $\forall\ h,k \in X$.

Para poder trabajar cómodamente con derivadas de orden superior, conviene hacer las identificaciones que siguen. Comencemos por un lema, de demostración sencilla, que identifica el espacio $\mathcal{L}(X,\mathcal{L}(X,Y))$ con $\mathcal{L}_2(X,Y)$, donde $\mathcal{L}_2(X,Y)$ es el espacio de las aplicaciones bilineales y continuas de $X \times X$ en Y. Recordemos que si $L \in \mathcal{L}_2(X,Y)$, $\exists m \geq 0, \quad / \left\| L(h,k) \right\|_Y \leq m \left\| h \right\|_X \left\| k \right\|_X, \ \forall (h,k) \in X \times X$. Entonces $\|L\|$ es el ínfimo (mínimo) de los números m que verifican la desigualdad anterior. También, $\|L\|_{\mathcal{L}_2(X,Y)} = \sup_{\|h\|\leq 1, \|k\| \leq 1} \ \|L(h,k)\|_Y$.

Lema 2.3.
Sea

$$\begin{array}{rccc} \Gamma: & \mathcal{L}(X,\mathcal{L}(X,Y)) & \longrightarrow & \mathcal{L}_2(X,Y) \\ & A & \longrightarrow & \Gamma_A \end{array}$$

donde $\Gamma_A : X \times X \longrightarrow Y$, $\Gamma_A(h,k) = (A(h))(k)$, $\forall (h,k) \in X \times X$.

Entonces Γ *es lineal, biyectiva y* $\|\Gamma_A\|_{\mathcal{L}_2(X,Y)} = \|A\|_{\mathcal{L}(X,\mathcal{L}(X,Y))}, \quad \forall A \in \mathcal{L}(X,\mathcal{L}(X,Y))$.

Una propiedad destacable es que la derivada segunda de una función, si existe, es siempre una aplicación bilineal simétrica (véanse [2], [8]).

En general, si $\mathcal{L}_n(X,Y)$ denota el conjunto de aplicaciones multilineales y continuas L de $X \times \ldots^{n)} \times X$ en Y, con la norma usual, es decir,

$$\|L\| = \sup_{\|x_1\|\leq 1, \cdots, \|x_n\|\leq 1} \|L(x_1, \cdots, x_n)\|_Y,$$

entonces la derivada $n-$ésima puede definirse por inducción, sin más que usar la identificación

$$\mathcal{L}(X, \mathcal{L}_n(X,Y)) \simeq \mathcal{L}_{n+1}(X,Y) \tag{2.10}$$

Como en el caso de dimensión dos, una propiedad destacable es que la derivada $n-$ésima de una función, si existe, es siempre una aplicación multilineal simétrica (véanse [2], [8]).

La siguiente fórmula, conocida como fórmula de Taylor, es semejante al caso de dimensión finita, y de gran utilidad en el estudio de máximos y mínimos de funcionales. Para la demostración véanse [2], [3] o bien [8]. Como hasta ahora, Ω denota cualquier abierto de un espacio normado real X, e Y denota también cualquier espacio normado real.

Teorema 2.4. *Sea* $\Phi : \Omega \to Y$, *de clase* $C^n(\Omega)$ *(es decir, tanto* Φ *como sus derivadas hasta el orden* n *son continuas en* Ω*), y dos puntos* $u, u+h \in \Omega$

tales que el segmento $[u, u+h] \subset \Omega$. *Entonces*

$$\Phi(u+h) = \Phi(u) + \Phi'(u)(h) + \ldots + \frac{1}{(n)!}\Phi^{n)}(u)(h)^n + \omega(u,h)(h)^n, \tag{2.11}$$

donde $\omega(u,h)$ *es un aplicación multilineal de* $X \times \ldots^{n)} \times X$ *en* Y *tal que* $\lim_{h\to 0}\ \omega(u,h) = 0$. *En particular* $\lim_{h\to 0}\ \frac{\omega(u,h)(h)^n}{\|h\|_X^n} = 0$.

Aquí, la notación $\Phi^{k)}(u)(h)^k$ significa la aplicación multilineal $\Phi^{k)}(u)$ aplicada al elemento $(h, \ldots, h) \in X^k$.

Los resultados anteriores permiten demostrar los teoremas siguientes, sobre condiciones necesarias y condiciones suficientes de máximos y mínimos.

Teorema 2.5.
Sea $\Phi : \Omega \longrightarrow \mathbb{R}$, $\Omega \subset X$ *abierto, y* $x_0 \in \Omega$ *un mínimo local de* Φ.

- *Si* Φ *es derivable en* x_0, *entonces* $\Phi'(x_0) = 0$.
- *Si existe la derivada segunda de* Φ *en* x_0, *entonces* $\Phi''(x_0)(h,h) \geq 0, \quad \forall h \in X$

Teorema 2.6.
Sea $\Phi : \Omega \longrightarrow \mathbb{R}$, $\Omega \subset X$ *abierto,* $\Phi \in \mathcal{C}^2(\Omega, \mathbb{R})$ *y* $x_0 \in \Omega$ *tal que:*

(1) $\Phi'(x_0) = 0$
(2) $\exists K > 0 \quad / \quad \Phi''(x_0)(h,h) \geq K\|h\|^2, \quad \forall h \in X$

Entonces Φ *tiene un mínimo local estricto en* x_0.

Nota 1. Notemos que la condición

$$\exists K > 0 \quad / \quad \Phi''(x_0)(h,h) \geq K\|h\|^2, \quad \forall h \in X$$

es equivalente a

$$\exists K > 0 \quad / \quad \Phi''(x_0)(h,h) \geq K,\ \forall h \in X \text{ tal que } \|h\| = 1.$$

Si el espacio X tiene dimensión finita, esto último es claramente equivalente a que

$$\Phi''(x_0)(h,h) > 0,\ \forall h \in X \text{ tal que } \|h\| = 1.$$

En dimensión infinita no es necesariamente así, debido a la no compacidad de la bola unidad del espacio X.

Veamos un ejemplo de lo anterior, de especial significación por las aplicaciones al cálculo de variaciones. Sea $f : [a,b] \times \mathbb{R} \longrightarrow \mathbb{R}$, $(t,x) \to f(t,x)$, continua tal que $\exists\ \frac{\partial f}{\partial x} : [a,b] \times \mathbb{R} \longrightarrow \mathbb{R}$ y es continua.

Definimos

$$\Phi : \mathcal{C}([a,b],\mathbb{R}) \longrightarrow \mathbb{R}$$

$$\Phi(x) = \int_a^b f(t,x(t))dt \tag{2.12}$$

- $\mathcal{C}([a,b],\mathbb{R})$ lo consideramos como espacio normado real, con la norma $\|x\|_0 = \max_{t\in[a,b]} |x(t)|$.

Proposición 2.7.
$\forall x_0 \in \mathcal{C}([a,b],\mathbb{R})$, $\exists\ \Phi'(x_0)$ y

$$\Phi'(x_0) : \mathcal{C}([a,b],\mathbb{R}) \longrightarrow \mathbb{R} \textit{ viene definida por}$$

$$\Phi'(x_0)(h) = \int_a^b \frac{\partial f(t,x_0(t))}{\partial x} h(t)dt, \ \forall\ h \in \mathcal{C}([a,b]. \tag{2.13}$$

Principales ideas de la demostración.
Tendremos que comprobar en primer lugar que $L(h) \equiv \Phi'(x_0)(h)$ tal y como se ha definido anteriormente, es lineal y continua respecto de h.

$$L : (\mathcal{C}([a,b]),\mathbb{R}) \longrightarrow \mathbb{R}$$

$$h \longrightarrow \int_a^b \frac{\partial f(t,x_0(t))}{\partial x} h(t)dt \tag{2.14}$$

es lineal de forma trivial. Para ver que es continua hemos de demostrar que existe $M \in \mathbb{R}^+$ tal que $|L(h)| \le M\,\|h\|_0$, $\forall h \in \mathcal{C}([a,b],\mathbb{R})$. Ahora bien, como M puede tomarse $M_1(b-a)$, donde M_1 es el máximo de la aplicación de $[a,b]$ en $\mathbb{R}$ definida como

$$t \longrightarrow \left|\frac{\partial f(t,x_0(t))}{\partial x}\right| \tag{2.15}$$

El siguiente paso es comprobar que se verifica la condición (2.8). Ahora bien,

$$\Phi(x_0+h) - \Phi(x_0) - L(h) =$$

$$\int_a^b \left[f(t,x_0(t)+h(t)) - f(t,x_0(t)) - \tfrac{\partial f(t,x_0(t))}{\partial x} h(t)\right] dt = (*) \tag{2.16}$$

Ahora podemos usar el Teorema del Valor Medio, en versión integral. Este Teorema, muy útil y sencillo de usar, afirma lo siguiente:

$$g \in \mathcal{C}^1([x,y],\mathbb{R}) \Rightarrow g(y) - g(x) = \int_0^1 g'(ty + (1-t)x)(y-x)dt \tag{2.17}$$

Pensemos que la demostración es trivial, puesto que

$$g'(ty + (1-t)x)(y-x) = \frac{d}{dt}\ g(ty + (1-t)x)$$

Usando este Teorema, la expresión (2.16), se transforma en

$$(*) = \int_a^b \left[\int_0^1 \frac{\partial f(t, rx_0(t) + rh(t) + (1-r)x_0(t))}{\partial x} h(t)dr - \frac{\partial f(t,x_0(t))}{\partial x} h(t) \right] dt = \tag{2.18}$$

$$= \int_a^b \left[\int_0^1 \left(\frac{\partial f(t, x_0(t) + rh(t))}{\partial x} - \frac{\partial f(t,x_0(t))}{\partial x} \right) dr \right] h(t)dt = \tag{2.19}$$

Tomamos valores absolutos, y obtenemos la expresión

$$\begin{gathered} |\Phi(x_0 + h) - \Phi(x_0) - L(h)| \leq \\ \|h\|_0 \int_a^b \int_0^1 \left| \frac{\partial f(t,x_0(t)+rh(t))}{\partial x} - \frac{\partial f(t,x_0(t))}{\partial x} \right| drdt \end{gathered} \tag{2.20}$$

Para concluir la demostración, bastaría comprobar que

$$\lim_{\|h\|_0 \to 0} \int_a^b \int_0^1 \left| \frac{\partial f(t, x_0(t) + rh(t))}{\partial x} - \frac{\partial f(t,x_0(t))}{\partial x} \right| drdt = 0 \tag{2.21}$$

Ahora bien, tenemos que $x_0(t) + rh(t)$ converge (uniformemente respecto de r) a la función $x_0(t)$ con la norma $\| \cdot \|_0$. Además, como $\dfrac{\partial f}{\partial x}$ es continua, esta función es uniformemente continua en compactos, y por ello se tiene que

$$\begin{gathered} \frac{\partial f(t,x_0(t)+rh(t))}{\partial x} \longrightarrow \frac{\partial f(t,x_0(t))}{\partial x} \\ \text{con la norma } \| \cdot \|_0, \text{ uniformemente respecto de } r \in [0,1]. \end{gathered} \tag{2.22}$$

Ideas análogas pueden usarse para el estudio del siguiente funcional: Sea $X = \mathcal{C}^1([a,b],\mathbb{R})$, con la norma dada por $\|x\| = \|x\|_0 + \|x'\|_0$.

Sea

$$\begin{gathered} f : [a,b] \times \mathbb{R} \times \mathbb{R} \longrightarrow \mathbb{R} \\ (t,x,y) \longrightarrow f(t,x,y) \end{gathered} \tag{2.23}$$

continua tal que para cualquier $t \in [a,b]$ fijo, la aplicación $f(t,x,y)$ es C^1 con respecto a $(x,y) \in \mathbb{R}^2$.

Si definimos el funcional

$$\Phi : X \longrightarrow \mathbb{R}$$

(2.24)
$$\Phi(x) = \int_a^b f(t,x(t),\dot{x}(t))dt$$

Entonces Φ es derivable y

(2.25)
$$\Phi'(x_0)(h) = \int_a^b \left[\frac{\partial f(t,x_0(t),\dot{x}_0(t))}{\partial x} h(t) + \frac{\partial f(t,x_0(t),\dot{x}_0(t))}{\partial \dot{x}} \dot{h}(t) \right] dt, \ \forall\, h \in X.$$

Otros ejemplos son los siguientes. Sea $X = C^k([a,b],\mathbb{R})$ con la norma $\|x\|_k = \sum_{p=0}^{k} \|x^{p)}\|_0$ y $f : [a,b]\times\mathbb{R}^{k+1} \to \mathbb{R}$, $(t,x_0,x_1,...,x_k) \to f(t,x_0,x_1,...,x_k)$ continua y tal que para cada $t \in [a,b]$ fijo, la aplicación $f(t,x_0,x_1,...,x_k)$ es C^1 con respecto a $(x_0,x_1,...,x_k) \in \mathbb{R}^{k+1}$. Si $\Phi : X \to \mathbb{R}$ se define como

$$\Phi(x) = \int_a^b f(t,x(t),...,x^{p)}(t),...,x^{k)}(t))\ dt,$$

entonces

$$\Phi'(x_0)(h) = \sum_{p=0}^{k} \int_a^b \frac{\partial f(t,x_0(t),...,x_0^{k)}(t))}{\partial x_p} h^{p)}(t)\ dt$$

Respecto de las derivadas de orden superior, las mismas ideas pueden usarse para probar el resultado siguiente:

Sea $f : [a,b] \times \mathbb{R} \longrightarrow \mathbb{R}$, $(t,x) \to f(t,x)$ continua tal que para cada $t \in [a,b]$ fijo, la aplicación $f(t,x)$ es C^2 respecto de $x \in \mathbb{R}$. Si Φ es el funcional definido en (2.12) entonces

$$\Phi''(x_0)(h,k) = \int_a^b \frac{\partial^2 f(t,x_0(t))}{\partial x^2} h(t)k(t)dt.$$

Para el funcional definido en (2.24) se tendría

(2.26)
$$\Phi''(x_0)(h,k) = \int_a^b [\tfrac{\partial^2 f(t,x_0(t),\dot{x}_0(t))}{\partial x^2} h(t)k(t) + \tfrac{\partial^2 f(t,x_0(t),\dot{x}_0(t))}{\partial x \partial y} h(t)\dot{k}(t) +$$

$$\tfrac{\partial^2 f(t,x_0(t),\dot{x}_0(t))}{\partial y \partial x} \dot{h}(t)k(t) + \tfrac{\partial^2 f(t,x_0(t),\dot{x}_0(t))}{\partial y^2} \dot{h}(t)\dot{k}(t)]dt$$

A veces es necesario trabajar con conceptos más débiles que la derivabilidad según Fréchet. Vamos a hablar brevemente de esto para funcionales que están definidos en algún abierto de un espacio normado y con valores reales.

Definición 2.8. Sea $(E, \|\cdot\|)$ un espacio normado real, Ω un subconjunto abierto de E, $\Phi : \Omega \longrightarrow \mathbb{R}$ y $x_0 \in E$.

- Recordemos que Φ es derivable según Fréchet en x_0 si $\exists\, L : E \longrightarrow \mathbb{R}$ lineal y continua tal que

$$\lim_{h \to 0} \frac{|\Phi(x_0 + h) - \Phi(x_0) - L(h)|}{\|h\|_X} = 0 \tag{2.27}$$

- Se dice que Φ es derivable según Gateaux en x_0 si

$$\forall h \in E \quad \exists \lim_{\tau \to 0} \frac{\Phi(x_0 + \tau h) - \Phi(x_0)}{\tau} \equiv \Phi'_G(x_0)(h) \tag{2.28}$$

  y la aplicación $\Phi'_G(x_0) : E \longrightarrow \mathbb{R}$, $h \to \Phi'_G(x_0)(h)$, es lineal y continua en $h \in E$. En este caso, a $\Phi'_G(x_0)$ se le llama derivada de Gateaux de Φ en x_0.

- Se define la primera variación de Φ en x_0, en la dirección $h \in E$ como

$$\delta\Phi(x_0, h) = \lim_{\tau \to 0} \frac{\Phi(x_0 + \tau h) - \Phi(x_0)}{\tau} \tag{2.29}$$

- En general, se define la n-ésima variación de la función Φ en x_0 y en la dirección $h \in E$ como

$$\delta^n\Phi(x_0, h) = \left. \frac{d^n\Phi(x_0 + th)}{dt^n} \right|_{t=0} \tag{2.30}$$

La derivabilidad según Fréchet implica la derivabilidad según Gateaux, y no se verifica la implicación contraria. La derivabilidad según Gateaux implica la existencia de la primera variación en todas las direcciones, pero no al contrario. También, las reglas algebraicas usuales, regla de la cadena, etc. son válidas para estos conceptos relacionados con la derivabilidad. Véase [17] para la relación entre estos conceptos.

Algunos ejemplos de interés son los siguientes:

Ejemplo 1.
Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$.

$$f(x,y) = \begin{cases} \dfrac{x^6}{(y - x^2)^2 + x^8} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

Esta función es derivable según Gateaux en el punto $(0,0)$ (de hecho, su derivada es la función nula), pero f no es continua en este punto. En particular f es derivable según Gateaux en $(0,0)$ pero no es Fréchet derivable en $(0,0)$.

Ejemplo 2.
Sea $f : \mathbb{R}^2 \longrightarrow \mathbb{R}$.

$$f(x,y) = \begin{cases} \dfrac{x^2y\sqrt{x^2+y^2}}{x^4+y^2} & \text{si } (x,y) \neq (0,0) \\ 0 & \text{si } (x,y) = (0,0) \end{cases}$$

Esta función es derivable según Gateaux en el punto $(0,0)$ (de hecho, su derivada es la función nula), es continua en este punto, pero no es derivable en el sentido de Fréchet.

¡Anímese el alumno a comprobar las afirmaciones anteriores!

El Teorema del valor medio es cierto, usando la noción de derivada de Gateaux. De hecho, tenemos el Teorema siguiente:

Teorema 2.9. *Sean X, Y espacios normados, Ω un abierto de X y $\Phi : \Omega \to Y$ derivable Gateaux en Ω. Si $x_1, x_2 \in \Omega$ son tales que el segmento $[x_1, x_2] \subset \Omega$, entonces*

$$\|\Phi(x_1) - \Phi(x_2)\|_Y \leq \left(\sup_{0 \leq t \leq 1} \|\Phi'_G(tx_1 + (1-t)x_2)\|_{\mathcal{L}(X,Y)} \right) \|x_1 - x_2\|_X$$

Usando este Teorema, puede probarse el siguiente, de interés en las aplicaciones, para determinar la derivada de Fréchet de los funcionales que estemos tratando. De hecho, la función dada Φ puede usarse de manera directa para intentar calcular la derivada de Gateaux, pero no así para calcular la derivada de Fréchet.

Teorema 2.10. *Sea $\Phi : \Omega \to Y$, derivable según Gateaux en Ω (es decir, en todos los puntos de Ω) y tal que la aplicación $\Phi'_G : \Omega \to \mathcal{L}(X,Y)$ es continua en $x_0 \in \Omega$. Entonces existe la derivada de Fréchet $\Phi'(x_0)$ y $\Phi'(x_0) = \Phi'_G(x_0)$.*

La versión tradicional del Teorema del valor medio es válida también. Esto exige introducir el concepto de integral definida,

$$\int_a^b f(t)\ dt \tag{2.31}$$

para funciones $f : [a,b] \to E$, continuas, donde E es un espacio normado cualquiera. No es difícil: en primer lugar se introduce el concepto anterior para funciones escalonadas. En segundo lugar, usando el hecho de que si $f : [a,b] \to E$, es continua, entonces es uniformemente continua, puede usarse un método límite para definir (2.31). Véase, por ejemplo ([12]).

Se obtiene así el Teorema fundamental del Cálculo que se enuncia a continuación.

Teorema 2.11. *Sea E un espacio normado cualquiera y $f : [a,b] \to E$ continua. Si $F : [a,b] \to E$ se define como $F(t) = \int_a^t f(s)\ ds$, entonces $F'(t) = f(t),\ \forall\ t \in [a,b]$.*

Lo que sigue es el Teorema del valor medio para derivada de Fréchet.

Teorema 2.12. *Sea $f : \Omega \to Y$ de clase $C^1(\Omega)$. Entonces, para todo par de puntos $x_1, x_2 \in \Omega$ tales que $[x_1, x_2] \subset \Omega$, se tiene:*

$$f(x_1) - f(x_2) = \int_0^1 f'(x_2 + t(x_1 - x_2))(x_1 - x_2)\ dt$$

¡Puede ser muy instructivo comparar los Teoremas 2.9 y 2.12!

2.2. **Convexidad de un operador y monotonía de su derivada.** Veamos a continuación algunas de las relaciones existentes entre la convexidad de una función y la monotonía de su derivada. Para ello, sea $(E, \|\cdot\|)$ un espacio normado, Ω un abierto de E y $f : \Omega \longrightarrow \mathbb{R}$ derivable. De acuerdo con la definición de derivada, se tiene

$$\begin{aligned} f' : \Omega &\longrightarrow \mathcal{L}(E, \mathbb{R}) \\ x_0 &\longrightarrow f'(x_0) \end{aligned}$$

donde $\mathcal{L}(E, \mathbb{R})$ es el espacio de Banach de las aplicaciones lineales y continuas de E en $\mathbb{R}$, con la norma usual, es decir:

$$\forall\ L \in \mathcal{L}(E, \mathbb{R}) \quad \|L\| = \sup_{\|x\| \le 1,\ x \in E} |L(x)|$$

Definición 2.13.

- f' se dice que es monótona si $(f'(x) - f'(y))(x - y) \ge 0 \quad \forall x, y \in \Omega$.
- f' se dice que es estrictamente monótona si $(f'(x) - f'(y))(x - y) > 0, \quad \forall x, y \in \Omega$ con $x \neq y$.

Teorema 2.14. *Sea $\Omega \subset E$, abierto y convexo y $f : \Omega \longrightarrow \mathbb{R} \quad f \in \mathcal{C}^1(\Omega)$.*

Son equivalentes:

(1) *f es convexa en Ω; es decir, $\forall\ x, y \in \Omega$ se tiene $f(\lambda x + (1-\lambda)y) \le \lambda f(x) + (1-\lambda) f(y),\ \forall\ \lambda \in [0,1]$.*
(2) *$f' : \Omega \longrightarrow \mathcal{L}(E, \mathbb{R})$ es monótona.*

También son equivalentes:

(1) *f es estrictamente convexa en Ω; es decir, $\forall\ x, y \in \Omega\ x \neq y$, se tiene $f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda) f(y),\ \forall\ \lambda \in (0,1)$.*
(2) *$f' : \Omega \longrightarrow \mathcal{L}(E, \mathbb{R})$ es estrictamente monótona.*

Demostremos la primera parte. *¡Inténtelo el alumno con la segunda. Es posible que se lleve alguna sorpresa!*

1. $\Rightarrow$ 2.

Si $x, y \in \Omega$ entonces $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y), \ \forall \ \lambda \in [0,1]$. De aquí obtenemos que $f(y + \lambda(x-y)) - f(y) \leq \lambda(f(x) - f(y)), \forall \lambda \in (0,1)$. Así,

$$\frac{f(y+\lambda(x-y)) - f(y)}{\lambda} \leq f(x) - f(y) \tag{2.32}$$

Si $\lambda \longrightarrow 0^+$ obtenemos

$$f'(y)(x-y) \equiv \delta f(y; x-y) \leq f(x) - f(y) \tag{2.33}$$

Análogamente, cambiando x por y, se obtiene la expresión $f'(x)(y-x) \leq f(y) - f(x)$. Sumando ambas expresiones, obtenemos finalmente que $(f'(y) - f'(x))(x-y) \leq 0$, por lo que $(f'(y) - f'(x))(y-x) \geq 0$.

2. $\Rightarrow$ 1.

Veamos que $\forall x, y \in \Omega$ y $\forall \lambda \in [0,1]$ se tiene que $f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$. En efecto, para ello definimos la función

$$\begin{gathered} p : [0,1] \longrightarrow \mathbb{R} \\ p(\lambda) = f(\lambda x + (1-\lambda)y) - \lambda f(x) - (1-\lambda) f(y) \end{gathered} \tag{2.34}$$

Hemos de comprobar que $p(\lambda) \leq 0 \quad \forall \lambda \in [0,1]$.

En principio $p(0) = p(1) = 0$. Supongamos que $\exists \lambda_0 \in (0,1)$ tal que $p(\lambda_0) > 0$. Entonces $\max_{[0,1]} p > 0$. Sea λ_1 un punto donde se alcance este máximo. Entonces $p'(\lambda_1) = 0$. Además, usando que f' es monótona, se tiene

$$\begin{gathered} p'(\lambda) - p'(\lambda_1) = f'(\lambda x + (1-\lambda)y)(x-y) - f(x) + f(y) \\ -f'(\lambda_1 x + (1-\lambda_1)y)(x-y) + f(x) - f(y) = \\ \left[f'(\lambda x + (1-\lambda)y) - f'(\lambda_1 x + (1-\lambda_1 y)\right](x-y) \end{gathered}$$

.

Si $u = \lambda x + (1-\lambda)y$, $v = \lambda_1 x + (1-\lambda_1)y$, la expresión anterior queda $(f'(u) - f'(v))(\frac{u-v}{\lambda-\lambda_1})$, que es una expresión no negativa para $\lambda > \lambda_1$, por la monotonía de f'. Esto implica que $p'(\lambda) \geq 0 \quad \forall \lambda > \lambda_1$, lo que no es posible si tenemos en cuenta los valores de p en $\lambda = \lambda_1$ y $\lambda = 1$.

Observando detenidamente la demostración anterior, obtenemos el siguiente resultado que será muy útil en la práctica.

Corolario 2.15.

Sea Ω convexo y $f : \Omega \longrightarrow \mathbb{R}$ convexa. Supongamos que para algún $x_0 \in \Omega$ se cumple

$$\delta f(x_0; x - x_0) = 0, \quad \forall x \in \Omega. \tag{2.35}$$

Entonces x_0 es un punto de mínimo global de f en Ω. Además, si f es estrictamente convexa, el punto donde se alcanza el mínimo es único.

Un ejemplo de lo anterior es el siguiente.
Sea $C = \{y \in \mathcal{C}^1[0,1] : \ y(0) = 0, \ y(1) = 1\}$, y $f : C \longrightarrow \mathbb{R}$ definida por

$$f(y) = \int_0^1 (y'^2(x) + 4y(x))dx$$

. Puede probarse que:

- Existe mínimo de f en C.
- Este mínimo se alcanza en la función $y_0(x) = x^2$.
- El punto donde se alcanza el mínimo es único.

De hecho, para este ejemplo se tiene

$$(f'(x) - f'(y), x - y) = 2 \int_0^1 (x' - y')^2$$

lo que prueba que el funcional f es estrictamente convexo. Además, una condición suficiente para que se cumpla (2.35) es que $x_0 \in C^2[0,1], \ x_0'' = 2, \ x_0(0) = 0, \ x_0(1) = 1$ (¿por qué?) Esto lo satisface la función $x_0(x) = x^2$.

Otro ejemplo puede ser el siguiente.
Sea $C = \{y \in \mathcal{C}^1[0,1] : \ y(0) = 0, \ y(1) = 0\}$, y $f : C \longrightarrow \mathbb{R}$ definida por

$$f(y) = \int_0^1 (y'^2(x) - y^2(x) + 2xy(x))dx$$

. Entonces puede probarse:

- Existe mínimo de f en C.
- Este mínimo se alcanza en la función $y_0(x) = x - \dfrac{\text{sen} x}{\text{sen} 1}$.
- El punto donde se alcanza el mínimo es único.

De hecho, para este ejemplo se tiene

$$(f'(y) - f'(z), y - z) = 2 \int_0^1 (y' - z')^2 - 2 \int_0^1 (y - z)^2$$

Usando series de Fourier ([5]) puede probarse fácilmente que

$$(f'(y) - f'(z), y - z) \geq 2(\pi^2 - 1) \int_0^1 (y - z)^2$$

lo que prueba que f es estrictamente convexo. Además, una condición suficiente para que se cumpla (2.35) es que $y_0 \in C^2[0,1], \ -y_0'' - y_0 + x = 0, \ y_0(0) = 0, \ y_0(1) = 1$ (¿por qué?) Esto lo satisface la función $y_0(x) = x - \dfrac{\text{sen } x}{\text{sen } 1}$.

2.3. **Condiciones necesarias (Euler-Lagrange y Legendre) en problemas clásicos del cálculo de variaciones.**

Comenzamos estudiando problemas clásicos del cálculo de variaciones donde ambos extremos de la función están fijos.

Teorema 2.16. *Sea* $X = \left\{ q \in \mathcal{C}^1([a,b],\mathbb{R}) \quad / \quad q(a) = q_1, \; q(b) = q_2 \right\}$ *con* (a, q_1) *y* (b, q_2) *dados.*

Sea $L \in \mathcal{C}^2([a,b] \times \mathbb{R} \times \mathbb{R}, \mathbb{R})$ *y*

$$\Phi : X \longrightarrow \mathbb{R}$$

$$\Phi(q) = \int_a^b L(t, q(t), \dot{q}(t))dt \tag{2.36}$$

En X *se considera la norma* $\| \cdot \|_1$, *esto es,* $\|q\|_1 = \max_{t\in[a,b]} \; |q(t)| + \max_{t\in[a,b]} \; |q'(t)|$.

Entonces:

(1) *(Condición necesaria de Euler-Lagrange) Si* $q_0 \in X$ *es tal que* q_0 *es un mínimo local de* Φ *en* X *(por tanto, mínimo local con la norma* $\| \cdot \|_1$*), entonces la función* $L_{\dot{q}}(t, q_0(t), \dot{q}_0(t))$ *es* $\mathcal{C}^1[a,b]$ *y*

$$\frac{d}{dt}(L_{\dot{q}}(t, q_0(t), \dot{q}_0(t)) = L_q(t, q_0(t), \dot{q}_0(t)) \tag{2.37}$$

(Ecuación de Euler-Lagrange). En particular, la ecuación de Euler-Lagrange (2.37) se cumple si $q_0 \in X$ *es un mínimo global de* Φ *en* X.

(2) *(Condición necesaria de Legendre) Si* $q_0 \in X$ *es tal que* q_0 *es un mínimo local de* Φ *en* X, *entonces*

$$L_{\dot{q}\dot{q}}(t, q_0(t), \dot{q}_0(t)) \geq 0, \quad \forall t \in [a,b]. \tag{2.38}$$

En particular, la condición necesaria de Legendre (2.38) se cumple si $q_0 \in X$ *es un mínimo global de* Φ *en* X.

(3) *Si* $q_0 \in X$ *satisface la condición necesaria de Euler-Lagrange y para cada* $t \in [a,b]$ *fijo, la aplicación* $\mathbb{R}^2 \to \mathbb{R}$, $(q, \dot{q}) \to L(t, q, \dot{q})$ *es convexa, entonces* q_0 *es mínimo global de* Φ *en* X.

(4) *Si* $q_0 \in X$ *satisface la condición necesaria de Euler-Lagrange y para cada* $t \in [a,b]$ *fijo la aplicación* $\mathbb{R}^2 \to \mathbb{R}$, $(q, \dot{q}) \to L(t, q, \dot{q})$ *es estrictamente convexa, entonces* q_0 *es el único mínimo global de* Φ *en* X. *Además,* q_0 *es mínimo global estricto de* Φ *en* X.

En lo que sigue,

$$X_0 = \left\{ h \in \mathcal{C}^1([a,b],\mathbb{R}) \quad / \quad h(a) = 0, \; h(b) = 0 \right\} \tag{2.39}$$

En la demostración de este teorema se usan, además de las nociones de convexidad del apartado anterior, los siguientes lemas (véanse [3], [17] y [18]).

¡Animamos a los alumnos a que proporcionen una demostración de los lemas que siguen, demostración que está perfectamente a su alcance!

Lema 2.17. *(Lema fundamental del cálculo de variaciones)*

Sea $q \in \mathcal{C}([a,b],\mathbb{R})$ *tal que* $\displaystyle\int_a^b q(t)h(t)dt = 0 \quad \forall h \in X_0$. *Entonces* $q \equiv 0$.

Lema 2.18. *(Lema de Du Bois Reymond)*

Sea $q \in \mathcal{C}([a,b],\mathbb{R})$ *tal que* $\displaystyle\int_a^b q(t)\dot{h}(t)dt = 0 \quad \forall h \in X_0$. *Entonces* q *es constante.*

Lema 2.19.

Sea la forma cuadrática

$$Q : \mathcal{C}^1([a,b],\mathbb{R}) \longrightarrow \mathbb{R}$$

$$Q(h) = \int_a^b \left[F_1(t)(h(t))^2 + F_2(t)h(t)\dot{h}(t) + F_3(t)(\dot{h}(t))^2\right] dt \tag{2.40}$$

donde $F_1, F_2, F_3 \in \mathcal{C}([a,b],\mathbb{R})$ *son dados y tal que* $Q(h) \geq 0 \quad \forall h \in X_0$.

Entonces $F_3(t) \geq 0 \quad \forall t \in [a,b]$.

Una vez que los anteriores lemas se dan por válidos, la demostración del teorema sigue las líneas siguientes:

1. Si q_0 es un punto de mínimo local de Φ entonces podemos definir

$$W : X_0 \longrightarrow \mathbb{R}$$

$$W(h) = \int_a^b L(t, q_0(t) + h(t), \dot{q}_0(t) + \dot{h}(t))dt \tag{2.41}$$

donde en X_0 consideramos la norma $\|\cdot\|_1$. Es claro que W tiene un mínimo local en 0. Luego $W'(0) = 0$ de donde obtenemos (téngase en cuenta (2.24) y (2.25))

$$\int_a^b (L_q(t, q_0(t), \dot{q}_0(t))h(t) + L_{\dot{q}}(t, q_0(t), \dot{q}_0(t))\dot{h}(t))dt = 0, \ \forall\ h \in X_0. \tag{2.42}$$

Integramos por partes el primer miembro de esta ecuación y obtenemos:

$$
\begin{aligned}
&h(t) \int_a^t (L_q(s, q_0(s), \dot{q}_0(s))ds \Big]_a^b \\
(2.43) \quad &- \int_a^b \left[\int_a^t L_q(s, q_0(s), \dot{q}_0(s))ds\dot{h}(t) + \int_a^b L_{\dot{q}}(t, q_0(t), \dot{q}_0(t))\dot{h}(t) \right] dt = 0
\end{aligned}
$$

con h anulándose en los extremos. Concluimos por el lema de Du Bois Reymond que la función

$$
(2.44) \qquad - \int_a^t L_q(s, q_0(s), \dot{q}_0(s))ds + L_{\dot{q}}(t, q_0(t), \dot{q}_0(t))
$$

es una constante.

Por tanto, $L_{\dot{q}}(t, q_0(t), \dot{q}_0(t))$ es de clase $\mathcal{C}^1$ en $[a, b]$.

Ahora hacemos de nuevo una integración por partes en (2.42) pero usando el segundo sumando. Así obtenemos
(2.45)

$$
\int_a^b (L_q(t, q_0(t), \dot{q}_0(t))h(t) - \frac{d}{dt}[L_{\dot{q}}(t, q_0(t), \dot{q}_0(t))]h(t))dt = 0, \ \forall \ h \in X_0.
$$

Usando el lema fundamental del cálculo de variaciones obtenemos la ecuación de Euler-Lagrange.

2. Si q_0 es un punto de mínimo local de Φ, entonces

$$
(2.46) \qquad W''(0)(h, h) \geq 0, \ \forall \ h \in X_0
$$

Como
(2.47)

$$
W''(0)(h, h) =
$$

$$
\int_a^b L_{qq}(t, q_0(t), \dot{q}_0(t))h^2(t) + 2L_{q\dot{q}}(t, q_0(t), \dot{q}_0(t))h(t)\dot{h}(t) + L_{\dot{q}\dot{q}}(t, q_0(t), \dot{q}_0(t))\dot{h}^2(t)
$$

usando el tercer lema tendríamos la condición necesaria de Legendre

$$
(2.48) \qquad L_{\dot{q}\dot{q}}(t, q_0(t), \dot{q}_0(t)) \geq 0 \quad \forall t \in [a, b]
$$

Las dos últimas partes del teorema son sencillas usando las ideas de convexidad del apartado anterior; en particular el Corolario 2.15. Tengamos en cuenta que, bajo las hipótesis del apartado (3) del Teorema, Φ es convexa y que si $q_0 \in X$ satisface la ecuación de Euler-Lagrange (2.37), entonces, $\delta\Phi(q_0; q - q_0) = 0, \ \forall \ q \in X$.

Nota 2. La condición de Euler-Lagrange (2.37) es una condición necesaria de mínimo local. En general no es suficiente. Para poder garantizar la existencia de mínimo global, necesitamos probar que la función L satisface, por ejemplo, propiedades adicionales de convexidad. Esto no es sencillo, pero algunas nociones de cálculo elemental pueden ayudarnos.

Pensemos que, en realidad, tenemos que probar que una función dada $f : \mathbb{R}^2 \to \mathbb{R}$ es convexa. La propia definición de convexidad no suele ser útil en estos casos, puesto que es complicada de comprobar.

Si $f \in C^1(\mathbb{R}^n)$, el Teorema 2.14 puede ayudarnos, pero tampoco suele ser fácil su aplicación. Si $f \in C^2(\mathbb{R}^n)$, entonces tenemos criterios más útiles usando la matriz Hessiana $H(f)(x)$. Si $H(f)(x)$ es semidefinida positiva, $\forall\ x \in \mathbb{R}^n$, entonces f es convexa en $\mathbb{R}^n$. Si $H(f)(x)$ es definida positiva, $\forall\ x \in \mathbb{R}^n$, entonces f es estrictamente convexa en $\mathbb{R}^n$. Lógicamente puede sustituirse $\mathbb{R}^n$ por cualquier subconjunto $C \subset \mathbb{R}^n$ que sea abierto y convexo. Por último, para probar que $H(f)(x)$ es semidefinida positiva $\forall\ x \in \mathbb{R}^n$ puede usarse el criterio de los menores principales: Si los $n-1$ primeros menores principales son positivos y el último es no negativo, entonces $H(f)(x)$ es semidefinida positiva. Si todos los menores principales son positivos entonces $H(f)(x)$ es definida positiva. También pueden ayudar mucho las propiedades siguientes:

(1) La suma finita de funciones convexas es una función convexa.
(2) El producto de un número real positivo por una función convexa es una función convexa.
(3) Si f es convexa y g creciente y convexa, entonces la composición $g(f(x))$ es convexa.

Un libro muy claro y útil para familiarizarse con los anteriores criterios es [14]. Muchos ejemplos concretos, de aplicación del Teorema 2.16, se pueden ver en [17].

Discutamos a continuación, de manera detallada, el problema de la braquistocrona, comentado en la introducción histórica. Este problema tiene sus peculiaridades. Por ejemplo, el funcional (1.2) no es del tipo contemplado en el Teorema 2.16, ya que el denominador se anula cuando $y(x) = 0$. Para solucionar éste y otros problemas relacionados con la convexidad, tomamos como eje de abscisas el eje vertical (con dirección positiva hacia abajo), como eje de ordenadas el eje horizontal. Entonces, si $A = (0,0), B = (b, b')$, con b y b' positivos, el funcional a minimizar adopta la forma (véase [17])

$$\Phi(y) = \int_0^b \left(\frac{1 + y'(x)^2}{2gx} \right)^{1/2} dx \tag{2.49}$$

donde $y \in X = \{y \in C^1([0,b], \mathbb{R}) : y(0) = 0, y(b) = b'\}$. A pesar de la presencia del término $(2gx)^{-1/2}$, $x \in [0,b]$ en la expresión anterior, se puede comprobar ¿fácilmente? que $\Phi(y) \in \mathbb{R},\ \forall\ y \in X$. En realidad, esto es tan sencillo como probar que $\int_0^b (x)^{-1/2}\ dx$ existe. Así pues, podemos hablar de posibles mínimos de Φ en X, sin usar necesariamente el Teorema 2.16. Además,

(1) Para cada $x \in (0,b]$ fijo, la aplicación $\mathbb{R} \to \mathbb{R}$, $y' \to \left(\frac{1+y'^2}{2gx}\right)^{1/2}$ es estrictamente convexa. De hecho, la derivada segunda de esta función, respecto de y', es estrictamente positiva.

(2) Si $L : (0,b] \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}$, está definida como $L(x,y,y') = \left(\frac{1+y'^2}{2gx}\right)^{1/2}$ y una función $y_0 \in X$ satisface la ecuación de Euler-Lagrange, en este caso,

$$\frac{d}{dx} L_{y'}(x, y_0(x), y_0'(x)) = 0, \ \forall \ x \in (0,b], \tag{2.50}$$

entonces la primera variación $\delta\Phi(y_0; h) = 0, \ \forall \ h \in C^1([0,b], \mathbb{R}) : y(0) = y(b) = 0$.

De las anteriores consideraciones se deduce trivialmente que X es convexo, que Φ es estrictamente convexo en X y que si $y_0 \in X$ satisface la ecuación de Euler Lagrange (2.50), entonces $\delta\Phi(y_0; y - y_0) = 0, \ \forall \ y \in X$. El Corolario 2.15 prueba que y_0 es mínimo global de Φ en X y, además, que el mínimo global se alcanza en un único elemento de X.

La ecuación de Euler-Lagrange (2.50) se cumple si y solamente si

$$\frac{y_0'(x)}{(2gx(1 + y_0'(x)^2))^{1/2}} = k, \tag{2.51}$$

para alguna constante $k \in \mathbb{R}$. Además, si $c \in \mathbb{R}^+$ satisface

$$cb < 1, \tag{2.52}$$

entonces cualquier función $y_0 \in X$ tal que

$$y_0'(x) = \left(\frac{cx}{1 - cx}\right)^{1/2}, \ \forall \ x \in [0,b], \tag{2.53}$$

satisface (2.51) con $c = 2gk^2$. Un cambio de variable adecuado (¿cuál?) permite integrar la ecuación anterior, obteniéndose

$$y_0(x) = \frac{\arcsin(cx)^{1/2} - (cx)^{1/2}(1 - cx)^{1/2}}{c} \tag{2.54}$$

Observemos que $y_0(0) = 0$. Sólo queda comprobar que se puede elegir c satisfaciendo (2.52) y tal que $y_0(b) = b'$. Técnicas elementales prueban que esto es posible si

$$\frac{b'}{b} < \frac{\pi}{2}. \tag{2.55}$$

Finalmente, la función y_0 definida en (2.54) se puede escribir en forma paramétrica como

$$\begin{aligned} x(t) &= \tfrac{1}{2c}(1 - \cos t), \\ y_0(t) &= \tfrac{1}{2c}(t - \operatorname{sen} t), \ t \in [0, \pi], \end{aligned} \tag{2.56}$$

que es una cicloide.

2.4. **Condiciones suficientes en problemas clásicos del cálculo de variaciones.**

Hasta ahora hemos establecido condiciones necesarias que han de satisfacer los mínimos locales. Vamos a ocuparnos a continuación del establecimiento de condiciones suficientes para tener la existencia de mínimos locales. Esto es bastante más complicado, y pueden consultarse [3] y [18] para los detalles. Si $L \in \mathcal{C}^3([a,b] \times \mathbb{R}^2, \mathbb{R})$ y $q_0 \in X \bigcap C^2[a,b]$, notemos

$$\begin{aligned} A^{(q_0)}(t) &\equiv L_{qq}(t, q_0(t), \dot{q}_0(t)) - \tfrac{d}{dt} L_{q\dot{q}}(t, q_0(t), \dot{q}_0(t)) \\ B^{(q_0)}(t) &\equiv L_{\dot{q}\dot{q}}(t, q_0(t), \dot{q}_0(t)) \end{aligned} \tag{2.57}$$

Teorema 2.20. *(Principio de mínima acción)*

Sea $q_0 \in X$ tal que:

(1) *q_0 verifica la ecuación de Euler-Lagrange, es decir,*

$$\frac{d}{dt} L_{\dot{q}}(t, q_0(t), \dot{q}_0(t)) = L_q(t, q_0(t), \dot{q}_0(t)), \ \forall t \in [a,b]$$

.

(2) $A^{(q_0)} \in C([a,b], \mathbb{R}), \ B^{(q_0)} \in C^1([a,b], \mathbb{R})$.

(3) $B^{(q_0)}(t) > 0, \ \forall \ t \in [a,b]$.

(4) *$A^{(q_0)}$ y $B^{(q_0)}$ satisfacen la condición de Jacobi en $[a,b]$, esto es: la ecuación diferencial ordinaria lineal de segundo orden*

$$\frac{d}{dt}(B^{(q_0)}(t)\dot{u}(t)) = A^{(q_0)}(t)u(t), t \in [a,b] \tag{2.58}$$

tiene alguna solución $u \in C^2[a,b]$ que no se anula en $[a,b]$ (por lo tanto, la ecuación anterior tiene alguna solución estrictamente positiva en $[a,b]$).

Entonces q_0 es un mínimo local estricto de Φ en X.

Para la demostración detallada de este resultado pueden consultarse [3] y [18]. No obstante, comentamos las ideas principales, que tienen interés en sí mismas, por lo que suponen sobre el estudio de formas cuadráticas en dimensión infinita.

Las etapas fundamentales de la demostración son las siguientes:

(1) **Regularidad de los puntos estacionarios.** Si $q_0 \in X$ es un punto estacionario (es decir, satisface la ecuación de Euler-Lagrange), y tal que $B^{(q_0)}(t) > 0$, $\forall \, t \in [a,b]$, entonces $q_0 \in C^2[a,b]$. En realidad, para tener este hecho, sólo necesitamos que la función $L_{\dot{q}}(t, q_0(t), \dot{q}_0(t))$ sea de clase $C^1[a,b]$ ([3]).

(2) **Expresión conveniente de la derivada segunda** $W''(0)(h,h)$. Por el paso previo, las hipótesis de nuestro teorema garantizan que la función $Lq\dot{q}(t,q_0(t),\dot{q}_0(t))$ es $C^1[a,b]$. Realizando una integración por partes en la expresión de la derivada segunda (véase (2.47)), obtenemos

$$\int_a^b 2L_{q\dot{q}}(t,q_0(t),\dot{q}_0(t))h(t)\dot{h}(t) =$$

$$\int_a^b L_{q\dot{q}}(t,q_0(t),\dot{q}_0(t))\tfrac{d}{dt}h^2(t) = -\int_a^b \tfrac{d}{dt}L_{q\dot{q}}(t,q_0(t),\dot{q}_0(t))h^2(t)$$

de donde se deduce

$$W''(0)(h,h) = \tag{2.59}$$

$$\int_a^b [L_{qq}(t,q_0(t),\dot{q}_0(t)) - \tfrac{d}{dt}L_{q\dot{q}}(t,q_0(t),\dot{q}_0(t))]h^2(t) + L_{\dot{q}\dot{q}}(t,q_0(t),\dot{q}_0(t))\dot{h}^2(t) =$$

$$\int_a^b A^{(q_0)}(t)h^2(t) + B^{(q_0)}(t)\dot{h}^2(t)$$

(3) **Uso de la condición de Jacobi en la forma cuadrática anterior.**

Sea u una solución positiva de (2.58). Entonces, cualquier función $h \in X_0$ (X_0 se ha definido en 2.39) puede escribirse de la forma $h = u\xi$, con $\xi \in X_0$. De esta manera

$$W''(0,0)(h,h) = \int_a^b A^{(q_0)}(t)h^2(t) + B^{(q_0)}(t)\dot{h}^2(t) =$$

$$\int_a^b A^{(q_0)}(t)u^2(t)\xi^2(t) + B^{(q_0)}(t)\dot{u}^2(t)\xi^2(t) +$$

$$\int_a^b B^{(q_0)}(t)u^2(t)\dot{\xi}^2(t) + 2B^{(q_0)}(t)(\dot{u}(t)\xi(t) + u(t)\dot{\xi}(t))$$

Realizando una integración por partes con el segundo miembro de la expresión anterior, tenemos

$$\int_a^b B^{(q_0)}(t)\dot{u}^2(t)\xi^2(t) = -\int_a^b B^{(q_0)}(t)u(t)(u''(t)\xi^2(t) + 2\dot{u}\xi(t)\dot{\xi}(t))$$

Así,

$$W''(0)(h,h) = \tag{2.60}$$

$$\int_a^b \xi^2(t)u(t)[-\tfrac{d}{dt}(B^{(q_0)}(t)\dot{u}(t)) + A^{(q_0)}(t)u(t)] + \int_a^b B^{(q_0)}(t)u^2(t)\dot{\xi}^2(t) =$$

$$\int_a^b B^{(q_0)}(t)u^2(t)\dot{\xi}^2(t).$$

Teniendo en cuenta la hipótesis sobre $B^{(q_0)}$ y el hecho de que u es estrictamente positiva, tenemos

$$W''(0)(h,h) > 0,\ \forall\ h \in X_0 \setminus \{0\}. \tag{2.61}$$

(4) Puede probarse aún más. Para ello, si ε es cualquier constante positiva suficientemente pequeña, la condición de Jacobi se sigue verificando para funcionales de la forma

$$\int_a^b A^{(q_0)}(t)h^2(t) + B^{(q_0)}(t)\dot{h}^2(t) - \varepsilon\dot{h}^2(t).$$

Para ver esto, pensemos que las soluciones de la ecuación diferencial

$$\frac{d}{dt}(B^{(q_0)}(t)\dot{u}(t) - \varepsilon\dot{u}(t)) = A^{(q_0)}(t)u(t), t \in [a,b]$$

dependen continuamente de las condiciones iniciales y del parámetro ε. Así, la ecuación diferencial anterior tiene alguna solución $v \in C^2[a,b]$ que no se anula en $[a,b]$. Por tanto, por el paso anterior, debe existir alguna constante $k_0 > 0$ verificando

$$W''(0)(h,h) \geq k_0 \int_a^b \dot{h}^2, \ \forall \ h \in X_0 \tag{2.62}$$

(5) **Uso del desarrollo de Taylor.**
Teniendo en cuenta que $W'(0) = 0$, tenemos que

$$\Phi(q_0 + h) - \Phi(q_0) = W(h) - W(0) = \tfrac{1}{2!}W''(0)(h,h) + R(h), \tag{2.63}$$

donde

$$\begin{gathered} R(h) = \textstyle\int_0^1 [(1-\psi)(W''(\psi h) - W''(0))(h,h)] \, d\psi = \\ \textstyle\int_a^b [A^{(q_0+\psi h)}(t) - A^{(q_0)}(t)]h^2(t) + [B^{(q_0+\psi h)}(t) - B^{(q_0)}(t)]\dot{h}^2(t) \end{gathered} \tag{2.64}$$

(6) **Uso de la desigualdad de Poincaré.** Como para cualquier función $h \in X_0$ se tiene

$$\int_a^b h^2 \leq \frac{(b-a)^2}{2} \int_a^b \dot{h}^2,$$

usando la regularidad de L y (2.64), obtenemos que para δ positivo y suficientemente pequeño y $h \in X_0$ satisfaciendo $\|h\|_1 \leq \delta$, entonces

$$|R(h)| \leq \frac{k_0}{2} \int_a^b \dot{h}^2(t) \tag{2.65}$$

(7) Por último, las relaciones (2.62) y (2.65) demuestran

$$\Phi(q_0 + h) - \Phi(q_0) > 0, \ \forall \ h \in X_0 \setminus \{0\} : \|h\|_1 \leq \delta \tag{2.66}$$

Nota 3. Puede verse en [3] y [18] que la condición de Jacobi es "casi necesaria" para que la forma cuadrática $W''(0)(h,h)$ sea definida positiva. No obstante, la condición de Jacobi no es sencilla de comprobar. Esto hace que las condiciones suficientes mostradas con anterioridad sean muy difíciles de aplicar en casos concretos.

2.5. **Otros problemas clásicos del cálculo de variaciones.** En un curso de estas características, debemos proporcionar al alumno los conocimientos suficientes para que resuelva por sí sólo situaciones similares. Las ideas expuestas con anterioridad deben ser suficientes para resolver algunas otras situaciones del cálculo de variaciones, como aquellas donde $q(t)$ es una función vectorial o L depende de derivadas de orden superior a uno. Comentamos algunas a continuación ([7]).

Funcionales que dependen de varias funciones.

Sea $X = \left\{ q \in \mathcal{C}^1([a,b],\mathbb{R}^n) \quad / \quad q(a) = q_1, q(b) = q_2 \right\}$ con (a,q_1) y (b,q_2) dados. Notemos $q = (q_1,\ldots,q_n),\ \dot{q} = (\dot{q_1},\ldots,\dot{q_n})$.

Sea $L \in \mathcal{C}^2([a,b]\times\mathbb{R}^n\times\mathbb{R}^n,\mathbb{R})$ y

$$\Phi : X \longrightarrow \mathbb{R}$$

$$\Phi(q) = \int_a^b L(t,q(t),\dot{q}(t))dt \tag{2.67}$$

La condición necesaria de Euler-Lagrange se expresaría ahora como un sistema de ecuaciones diferenciales ordinarias de la forma

$$\frac{d}{dt}(L_{\dot{q}_i}(t,q_0(t),\dot{q}_0(t)) = L_{q_i}(t,q_0(t),\dot{q}_0(t)),\ 1 \le i \le n \tag{2.68}$$

Funcionales que dependen de derivadas de orden superior.

Sea $X = \left\{ q \in \mathcal{C}^k([a,b],\mathbb{R}) \quad / \quad q^{j)}(a) = q_1^j, q^{j)}(b) = q_2^j,\ 0 \le j \le k-1 \right\}$ con $(a,q_1) \in \mathbb{R}^{k+1}$ y $(b,q_2) \in \mathbb{R}^{k+1}$ dados.

Sea $L \in \mathcal{C}^2([a,b]\times\mathbb{R}^{k+1},\mathbb{R})$ y

$$\Phi : X \longrightarrow \mathbb{R}$$

$$\Phi(q) = \int_a^b L(t,q(t),q^{1)}(t)\ldots,q^{k)}(t))dt \tag{2.69}$$

La condición necesaria de Euler-Lagrange se expresaría ahora como una ecuación diferencial de orden $2k$ dada por

$$\sum_{i=1}^{k}(-1)^{i+1}\frac{d^i}{dt^i}(L_{q^{i)}}(t,q_0(t),q_0^{1)}(t),\ldots,q_0^{k)}(t)) = L_q(t,q_0(t),q_0^{1)}(t),\ldots,q_0^{k)}(t)). \tag{2.70}$$

La condición en un extremo es libre.

Sea $X = \left\{ q \in \mathcal{C}^1([a,b],\mathbb{R}) \quad / \quad q(a) = q_1 \right\}$ con (a,q_1) dado.

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

$$\Phi : X \longrightarrow \mathbb{R}$$

$$\Phi(q) = \int_a^b L(t, q(t), \dot{q}(t))dt \tag{2.71}$$

La condición necesaria de Euler se expresaría ahora como

$$\begin{gathered} \tfrac{d}{dt}(L_{\dot{q}}(t, q_0(t), \dot{q}_0(t)) = L_q(t, q_0(t), \dot{q}_0(t)), \\ L_{\dot{q}}(b, q_0(b), \dot{q}_0(b)) = 0. \end{gathered} \tag{2.72}$$

Observemos que, al haber tomado ahora un conjunto X que incluye estrictamente al del Teorema 2.16, aparece una condición adicional para $t = b$, a la que se le suele llamar "condición natural de frontera."

No hay condiciones en los extremos: ambos son libres.
Sea $X = \left\{ q \in \mathcal{C}^1([a,b], \mathbb{R}) \right\}$.

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

$$\Phi : X \longrightarrow \mathbb{R}$$

$$\Phi(q) = \int_a^b L(t, q(t), \dot{q}(t))dt \tag{2.73}$$

La condición necesaria de Euler se expresaría ahora como

$$\begin{gathered} \tfrac{d}{dt}(L_{\dot{q}}(t, q_0(t), \dot{q}_0(t)) = L_q(t, q_0(t), \dot{q}_0(t)), \\ L_{\dot{q}}(b, q_0(b), \dot{q}_0(b)) = 0, \\ L_{\dot{q}}(a, q_0(a), \dot{q}_0(a)) = 0. \end{gathered} \tag{2.74}$$

Observemos que, al haber tomado ahora un conjunto X que incluye estrictamente al del Teorema 2.16 y al del caso anterior, aparece una condición adicional para $t = b$ y otra para $t = a$, a las que se le suele llamar "condiciones naturales de frontera."

Hay restricciones adicionales (multiplicadores de Lagrange).
Sea $X = \left\{ q \in \mathcal{C}^1([a,b], \mathbb{R}),\ q(a) = q_1,\ q(b) = q_2,\ \int_a^b g(t, q(t), \dot{q}(t))\ dt = 0 \right\}$ con $(a, q_1), (b, q_2)$ y la función $g \in C^2([a,b] \times \mathbb{R} \times \mathbb{R}, \mathbb{R})$ dados.

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

$$\Phi : X \longrightarrow \mathbb{R}$$

$$\Phi(q) = \int_a^b L(t, q(t), \dot{q}(t))dt \tag{2.75}$$

La condición necesaria de Euler se expresaría ahora diciendo que existe alguna constante real λ (multiplicador de Lagrange) tal que

(2.76)
$$\frac{d}{dt}(L_{\dot{q}}(t,q_0(t),\dot{q}_0(t)) - L_q(t,q_0(t),\dot{q}_0(t)) = \lambda[\frac{d}{dt}(g_{\dot{q}}(t,q_0(t),\dot{q}_0(t)) - g_q(t,q_0(t),\dot{q}_0(t))].$$