

ANEXO IV.

INTRODUCCIÓN AL CÁLCULO VARIACIONAL.

Problemas Variacionales en Física.

El principio de mínima acción.

En Mecánica clásica, cuando una partícula se mueve bajo la acción de un potencial $V(\mathbf{x})$, el movimiento real es el dado por las ecuaciones de Newton, que expresan la aceleración de la partícula en términos de las fuerzas. Cuando las fuerzas derivan de un potencial $V(\mathbf{x})$, el movimiento real $t \rightarrow \mathbf{x}(t)$ satisface la ecuación diferencial:

$$m \frac{d^2 \mathbf{x}(t)}{dt^2} = - \frac{\partial V(\mathbf{x}(t))}{\partial \mathbf{x}}.$$

Cuya solución determina el movimiento *real* que sigue una partícula que en un instante inicial t_1 sale del punto \mathbf{x}_1 , se mueve bajo la acción del potencial, y llega en un instante final t_2 al punto \mathbf{x}_2 .

Una pregunta interesante es: ¿Podemos singularizar el movimiento *real* dado por las soluciones de esta ecuación, entre todos los movimientos que la partícula *podría seguir*, para ir desde el punto inicial \mathbf{x}_1 en el instante t_1 al punto final \mathbf{x}_2 en el instante t_2 ?

La respuesta a esta pregunta es un principio básico en Física, que en Mecánica se denomina principio de Hamilton, o principio de mínima acción. Este principio caracteriza a los movimientos *reales* entre todos los movimientos *imaginables* que llevarían a la partícula del estado inicial (posición \mathbf{x}_1 en el instante t_1) al estado final (posición \mathbf{x}_2 en el instante t_2), ambos dados.

La caracterización dada por el principio de Hamilton asocia una cantidad, denominada *acción* a cada movimiento imaginable. La acción es una cantidad de naturaleza bastante diferente a las cantidades que usualmente describen el *estado* de la partícula, como posición y/o velocidad. A diferencia de ellas, la acción no se asocia al estado, sino a la *historia completa* de la partícula entre dos instantes inicial y final. Para cada movimiento imaginable, descrito por $t \rightarrow \mathbf{x}(t)$ con las condiciones $\mathbf{x}(t_1) = \mathbf{x}_1$, $\mathbf{x}(t_2) = \mathbf{x}_2$, la *acción* de ese movimiento se define como

$$S[x(t)] = \int_{t_1}^{t_2} \left\{ \frac{1}{2} m \left(\frac{dx(t)}{dt} \right)^2 - V(x(t)) \right\} dt$$

El principio de *mínima acción* dice: entre todos los movimientos imaginables, la propiedad que distingue al movimiento *real* es que el valor de la acción $S[x(t)]$ es *menor* para el movimiento real que para cualquier otro.

¿Cuál es la relación entre este principio y la forma newtoniana de plantear las ecuaciones del movimiento? Resulta que ambas maneras de describir el movimiento son equivalentes. Para verlo, necesitamos abordar el problema de la *búsqueda* de la función $x(t)$ con las condiciones requeridas, que *minimice* el valor de la acción. No se trata de un problema ordinario de mínimo, ya que la acción depende del movimiento como un todo, esto es, depende de la función $x(t)$.

Principio de Fermat.

Según la Óptica Geométrica, la luz se propaga a lo largo de *rayos*. Entre todos los rayos *posibles* que unen dos puntos dados, ¿cual es el escogido *realmente* por la luz? En la antigüedad clásica se observó que en *ciertas* circunstancias la luz viaja a lo largo del camino geoméricamente más corto entre dos puntos extremos A, B .

Pero basta observar la propagación de la luz en una interfase entre aire y agua (en un río), para concluir que la luz *no siempre* sigue el camino más corto; el ejemplo más evidente es la refracción, pero hay otros, como los espejismos.

Se atribuye a Fermat el primer enunciado del principio general que la trayectoria *real* seguida por un rayo de luz entre dos puntos dados en un medio posiblemente inhomogéneo es aquella que hace *mínimo* el *tiempo total invertido*. Se trata de un enunciado notable, ya que cuando Fermat lo formuló, se comprendían aún muy mal los elementos implicados en el proceso de propagación. Por ejemplo, el que la velocidad de la luz en un medio material es *siempre* menor que la velocidad de la luz en el vacío sólo se decidió experimentalmente en el S. XIX.

Con la perspectiva actual podemos traducir a ecuaciones el principio de Fermat, así: La velocidad de la luz en el vacío es constante c . En un medio material, su velocidad v es *menor* que c , y el cociente c/v es igual al índice de refracción n del

medio; para medios no homogéneos, este índice es una función de la posición $n(\mathbf{x})$.

Supongamos una trayectoria posible para un rayo luminoso en un medio inhomogéneo, en el que el índice de refracción dependerá de la posición. Tomando la coordenada z como parámetro a lo largo del rayo (cuya dirección supondremos cercana al eje z), podemos describir tal trayectoria como $z \rightarrow (x(z), y(z), z)$. La longitud del rayo entre los puntos de parámetro z y $z + dz$ es:

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

El tiempo requerido para viajar entre estos dos puntos está dado por:

$$d\tau = \frac{ds}{v} = \frac{ds}{\frac{c}{n(z;x,y)}} = \frac{n(z;x,y)}{c} ds = \frac{n(z;x,y)}{c} \sqrt{x'(z)^2 + y'(z)^2 + 1} dz$$

y el tiempo total invertido en viajar desde un punto inicial (x_1, y_1, z_1) hasta otro final (x_2, y_2, z_2) , a lo largo del rayo descrito por $z \rightarrow (x(z), y(z), z)$ (que debe satisfacer las condiciones $x(z_1) = x_1, y(z_1) = y_1; x(z_2) = x_2, y(z_2) = y_2$) es

$$T = \frac{1}{c} \int_{z_1}^{z_2} n(z;x,y) \sqrt{x'(z)^2 + y'(z)^2 + 1} dz$$

Así pues, el principio de Fermat reduce el problema de encontrar la trayectoria seguida por un rayo luminoso al problema de encontrar, entre todas las curvas $z \rightarrow (x(z), y(z), z)$ que unan los puntos dados (x_1, y_1, z_1) y (x_2, y_2, z_2) , aquella para la cual el valor de esta integral sea mínimo.

En el vacío, o en cualquier medio que sea homogéneo, el índice de refracción es constante, y no depende de la posición. En este caso, el principio de tiempo mínimo se reduce al principio de longitud mínima, y las trayectorias seguidas por los rayos son, en el espacio euclídeo, líneas rectas.

El problema de la braquistocrona.

Sean dos puntos P y Q situados en el mismo plano vertical, P más alto que Q y no directamente sobre Q . Un punto material se mueve sin fricción entre P y Q a lo largo de una curva determinada que une P con Q , bajo la acción de la fuerza de la gravedad, que supondremos uniforme, y partiendo de P con velocidad inicial nula. De entre todas las curvas posibles que unen P con Q , ¿sobre cuál de ellas el tiempo que tarda la partícula en ir desde P hasta Q es el menor posible? Esta curva tiene un nombre especial: braquistocrona.

Denotemos por z la altura, y por x la coordenada horizontal sobre el plano.

Cualquier curva que una P con Q estará descrita por una función $x \rightarrow z(x)$, que deberá satisfacer las dos condiciones $z(x_P) = z_P$, $z(x_Q) = z_Q$. En el punto inicial, la energía de la partícula vale $E = mgz_P$. Cuando la partícula se encuentra en el punto genérico $(x, z(x))$ sobre la curva, su velocidad $v(x)$ está determinada por el principio de conservación de la energía

$$E = mgz_P = mgz(x) + \frac{1}{2}mv(x)^2,$$

de donde resulta

$$v(x) = \sqrt{2g(z_P - z(x))}$$

El tiempo invertido en llegar desde el punto de coordenada x al punto de coordenada $x + dx$ es:

$$dt = \frac{ds}{v(x)} = \frac{ds}{\sqrt{2g(z_P - z(x))}} = \frac{\sqrt{1 + z'(x)^2}}{\sqrt{2g(z_P - z(x))}} dx,$$

y el tiempo total invertido en llegar desde el punto P al punto Q a lo largo de la curva dada vale:

$$T = \frac{1}{\sqrt{2g}} \int_{x_P}^{x_Q} \frac{\sqrt{1 + z'(x)^2}}{\sqrt{z_P - z(x)}} dx$$

La catenaria.

¿Cómo cuelga un hilo inextensible y flexible, de longitud total L , suspendido entre dos torres con separación horizontal d , y alturas dadas, A y B ? Claramente, el principio que determina la forma de equilibrio del hilo es que su energía potencial sea la menor posible. Cada forma posible del hilo está descrita por una función $x \rightarrow z(x)$ que debe satisfacer las condiciones $z(a) = A$, $z(b) = B$ (donde a , b son las coordenadas horizontales de las torres, $d = b - a$, y además otra condición importante, a saber, la longitud total del hilo debe ser L ; esta condición se traduce en:

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

Veamos ahora cómo se expresa la energía potencial del hilo cuando su forma es la función $z(x)$. Suponiendo el hilo de densidad lineal ρ constante, la masa del elemento entre las coordenadas x y $x+dx$ es $\rho\sqrt{1 + z'(x)^2} dx$, y la energía potencial de ese elemento es $z(x)g \rho\sqrt{1 + z'(x)^2} dx$. Así pues, la energía potencial total es:

$$E = \rho g \int_a^b z(x) \sqrt{1 + z'(x)^2} dx.$$

Los problemas isoperimétricos clásicos

A los problemas variacionales con ligaduras se les suele denominar problemas isoperimétricos. Su prototipo son los problemas clásicos de Dido (“te concederé tanto terreno cuanto puedas encerrar con la piel de este buey”), que fueron planteados y resueltos en la antigüedad clásica. Consideremos una curva cerrada en el plano, de longitud total L . ¿Qué figura de la curva hace máxima el área encerrada por la curva? O, inversamente, consideremos una curva cerrada de área dada S .

¿Cuándo la longitud de la curva es mínima? En ambos casos la respuesta es un círculo. En tres dimensiones, la superficie cerrada de área dada que encierra mayor volumen es la esfera, y la forma de volumen dado que tiene menor área superficial es también la esfera; estos resultados subyacen a la explicación de que la forma de la esfera se encuentre por doquier en la Naturaleza.

Podemos plantear formalmente el problema isoperimétrico buscando, entre las curvas que unan dos puntos P, Q situados sobre el eje real (de coordenadas $(a, 0)$ y $(b, 0)$), una curva $y = y(x)$, tal que $y(a) = y(b) = 0$, que no corte en otros puntos intermedios al eje real (por ejemplo, imponiendo $y(x) > 0$ en todo el intervalo $[a, b]$), y que tenga longitud total dada L . Esta última condición se expresa por la ecuación

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

y el problema a resolver es: entre todas las curvas que satisfagan esta condición, encuéntrase aquella que maximiza el área comprendida entre ella y el eje real, es decir, la que proporcione un valor mayor para la integral

$$\int_a^b y(x) dx.$$

Recapitulación.

En todos los casos, el problema propuesto se reduce a buscar, entre todas las funciones $f: x \rightarrow f(x)$ definidas en un intervalo $[a, b]$, y con condiciones del tipo $f(a) = A, f(b) = B$ (además de otras condiciones de continuidad, regularidad, etc. que se precisarán a su tiempo), aquellas que minimizan o maximizan una expresión del tipo

$$\int_a^b \Phi(x, f(x), f'(x)) dx.$$

En algunos casos, la función $f: x \rightarrow f(x)$ debe satisfacer ciertas condiciones adicionales, que pueden imaginarse como ligaduras; en todos los casos que hemos discutido las ligaduras están expresadas también por condiciones del tipo

$$\int_a^b \Xi(x, f(x), f'(x)) dx = \text{cte}$$

Derivación 'a la Feynmann' de las ecuaciones de Euler-Lagrange para el principio de menor acción

Imaginemos el caso más sencillo de una partícula que se mueve en una dimensión (coordenada posición x) bajo un potencial $V(x)$. Por ejemplo, un objeto que sube y baja verticalmente en el campo gravitatorio terrestre bajo la acción de la gravedad. Para cada movimiento imaginable, descrito por $t \rightarrow x(t)$ con las condiciones $x(t_1) = x_1$, $x(t_2) = x_2$, la acción de ese movimiento se define como:

$$S[x(t)] = \int_{t_1}^{t_2} \left\{ \frac{1}{2} m \left(\frac{dx(t)}{dt} \right)^2 - V(x(t)) \right\} dt$$

El principio de mínima acción dice: entre todos los movimientos imaginables, la propiedad que distingue al movimiento real es que el valor de la acción $S[x(t)]$ es menor para el movimiento real que para cualquier otro. Antes mencionamos que aunque no lo parezca, esta manera de singularizar el movimiento real entre todos los posibles es equivalente a las leyes de Newton, que en este caso serían

El principio de mínima acción dice: entre todos los movimientos imaginables, la propiedad que distingue al movimiento real es que el valor de la acción $S[x(t)]$ es menor para el movimiento real que para cualquier otro. Antes mencionamos que aunque no lo parezca, esta manera de singularizar el movimiento real entre todos los posibles es equivalente a las leyes de Newton, que en este caso serían

$$m \frac{d^2 x(t)}{dt^2} = - \frac{\partial V(x(t))}{\partial x}.$$

Vamos ahora a presentar un argumento heurístico, siguiendo a Feynmann, para convencer al lector de que el movimiento que minimice la acción debe satisfacer las ecuaciones de Newton. Comenzamos con una situación familiar que debe ser bien conocida; la búsqueda de mínimos de funciones de varias variables. Sea una función $x \rightarrow F(x)$ de varias variables $x = (x_1, x_2, \dots, x_i, \dots)$, supuesta continua y con derivadas continuas. Queremos encontrar un punto x_0 en el cual se verifique

$$F(x) > F(x_0)$$

para cualquier x cercano a x_0 . Una condición necesaria es que todas las derivadas parciales de F se anulen en x_0 ; imaginemos que no lo supiéramos y veamos cómo podríamos obtener tal condición partiendo sólo del conocimiento

más básico de la condición de mínimo para funciones de una variable,

a saber $\left. \frac{df}{dx} \right|_{x^0} = 0$.

¿Cómo podríamos aprovechar tal conocimiento? La idea básica, que en los apartados siguientes trasladaremos al caso de minimizar la acción, es: supongamos que realmente el punto x_0 es un mínimo de F cuando x varía en las cercanías de x_0 . Entonces también x_0 es un mínimo de F cuando x varía sólo a lo largo de la recta $\ell: x_i = x_i^0 + \epsilon h_i$, donde los h_i son arbitrarios. La restricción de F a la recta ℓ nos da una función de una sola variable ϵ :

$$f_{\mathbf{h}}(\epsilon) := F(x^0 + \epsilon \mathbf{h})$$

que debe tener un mínimo en $\epsilon = 0$. La condición para ello es que la derivada $\frac{df_{\mathbf{h}}}{d\epsilon}$ se anule en $\epsilon = 0$. La derivada se calcula mediante la regla de la cadena:

$$\frac{df_{\mathbf{h}}}{d\epsilon} = \frac{\partial F}{\partial x_1} h_1 + \frac{\partial F}{\partial x_2} h_2 + \dots$$

de manera que la condición de mínimo es:

$$0 = \left. \frac{df_{\mathbf{h}}}{d\epsilon} \right|_{\epsilon=0} = \left. \frac{\partial F}{\partial x_1} \right|_{x^0} h_1 + \left. \frac{\partial F}{\partial x_2} \right|_{x^0} h_2 + \dots$$

Como esta condición debe satisfacerse para \mathbf{h} arbitrario, basta con tomar sucesivamente $\mathbf{h} = (0, 0, \dots, h_i = 1, 0, \dots, 0)$ con $i = 1, 2, \dots$ para obtener

$$\left. \frac{\partial F}{\partial x_i} \right|_{x^0} = 0.$$

Naturalmente desde el momento en que todas las derivadas parciales se anulan en x_0 , entonces es claro que la ecuación anterior se verifica para \mathbf{h} arbitrario.

Así lo que hemos (re)encontrado es una condición necesaria para que una función de varias variables tenga un mínimo en x_0 : todas las derivadas parciales deben anularse en ese punto.

Repitamos el mismo proceso con la acción. Para evitar complicaciones inesenciales supondremos el movimiento en una dimensión (esto es la función incógnita $x(t)$ es una función de una variable) y buscamos un movimiento $t \rightarrow x_0(t)$ dado por una función continua, con derivada continua que lleve de x_a en el instante t_a

a x_b en el instante t_b y tal que para cualquier otro movimiento cercano se verifique

$$S[x(t)] > S[x_0(t)], \quad S[x(t)] = \int_{t_a}^{t_b} \left\{ \frac{1}{2}m \left(\frac{dx(t)}{dt} \right)^2 - V(x(t)) \right\} dt$$

La idea es construir una familia uniparamétrica de movimientos cercanos a $x_0(t)$, en la cual cada movimiento esté etiquetado por un sólo parámetro ", imitando a la expresión usada para funciones de varias variables; esto se consigue definiendo:

$$x(t) = x_0(t) + \epsilon h(t)$$

donde $h(t)$ debe ser una función fijada, suficientemente regular (continua, con derivada continua) satisfaciendo las condiciones $h(t_a) = 0$, $h(t_b) = 0$. Ahora restringimos la acción a los movimientos de esta familia unidimensional, obteniendo una función que depende de una sola variable \mathcal{E} :

$$S_{h(t)}(\epsilon) := S[x_0(t) + \epsilon h(t)]$$

y que debe tener un mínimo en $\mathcal{E} = 0$, lo que implica:

$$0 = \left. \frac{dS_{h(t)}}{d\epsilon} \right|_{\epsilon=0} = \frac{d}{d\epsilon} \int_{t_a}^{t_b} \left\{ \frac{1}{2}m \left(\frac{d(x_0(t) + \epsilon h(t))}{dt} \right)^2 - V(x_0(t) + \epsilon h(t)) \right\} dt \Big|_{\epsilon=0}$$

Efectuando la derivación con respecto a \mathcal{E} y evaluando en $\mathcal{E} = 0$ lo que se encuentra es

$$0 = \int_{t_a}^{t_b} \left\{ m \frac{dx_0(t)}{dt} \frac{dh(t)}{dt} - \left. \frac{dV}{dx} \right|_{x_0(t)} h(t) \right\} dt$$

Ahora la idea clave es realizar una integración por partes para transformar el término que involucra la derivada $\frac{dh(t)}{dt}$ de $h(t)$ en un término que involucre directamente a $h(t)$, y otro que desaparece por las condiciones de frontera. En efecto, efectuando el cambio estándar $u = m \frac{dx_0(t)}{dt}$, $v = h(t)$ en el primer término de la integral, resulta:

$$\int_{t_a}^{t_b} m \frac{dx_0(t)}{dt} \frac{dh(t)}{dt} dt = m \frac{dx_0(t)}{dt} h(t) \Big|_{t_a}^{t_b} - \int_{t_a}^{t_b} m \frac{d^2x_0(t)}{dt^2} h(t) dt = - \int_{t_a}^{t_b} m \frac{d^2x_0(t)}{dt^2} h(t) dt$$

y en consecuencia la condición de que la función $S_{h(t)}(\epsilon)$ tenga un mínimo en $\mathcal{E} = 0$ es:

$$\int_{t_a}^{t_b} \left\{ m \frac{d^2x_0(t)}{dt^2} + \left. \frac{dV}{dx} \right|_{x_0(t)} \right\} h(t) dt = 0.$$

Esta condición debe satisfacerse para todo $h(t)$ con las condiciones de contorno adecuadas, lo que sólo puede ocurrir si la parte entre llaves del integrando se anula idénticamente. Demostrar esto con rigor requiere cierto cuidado —lo haremos,

después—, pero podemos de todas maneras ver que si el término entre llaves del integrando fuera diferente de 0 en un cierto instante, entonces podríamos tomar una función $h(t)$ que fuera diferente de 0 sólo en un entorno muy pequeño de dicho instante, lo que llevaría a un valor no nulo para la integral; como esto no debe ocurrir, parece claro que el término entre llaves debe anularse siempre.

Aunque este argumento sea poco riguroso, la conclusión a se llega es correcta: de la exigencia de anulación de la integral anterior para todo $h(t)$ se concluye una condición necesaria de mínimo para el funcional de acción, que es que se verifique la llamada *ecuación de Euler-Lagrange del problema variacional*:

$$m \frac{d^2 x_0(t)}{dt^2} + \left. \frac{dV}{dx} \right|_{x_0(t)} = 0,$$

que por supuesto es simplemente la ecuación de Newton que gobierna el movimiento de la partícula en el campo de potencial $V(x)$. Así pues, matemáticamente la descripción a través del principio de mínima acción y a través de las ecuaciones de Newton son equivalentes.

Introducción a las Matemáticas del Cálculo Variacional

Ejemplos como el principio de mínima acción, el principio de Fermat, y problemas como la determinación de la forma de equilibrio de un hilo flexible e inextensible o la determinación de la curva que encierra un área máxima con perímetro dado, muestran la necesidad de considerar, junto con las funciones de varias variables, un tipo más general de aplicaciones, que a cada función de un determinado conjunto de funciones le asocian un número real. Desde un punto de vista formal, se trata simplemente de aplicaciones de ciertos conjuntos de funciones en la recta real R , y por tanto encajan dentro de la definición general de función, como aplicación entre conjuntos. Pero es tradicional y resulta conveniente usar en estos casos el nombre de funcional, para enfatizar aquellos aspectos en los que el cálculo con funciones de varias variables difiera del cálculo con este nuevo tipo de “funciones” definidas en espacios de funciones cuya dimensión es infinita.

La idea de diferenciabilidad para funcionales puede desarrollarse de manera semejante a cómo se hace para funciones de varias variables en el caso de que el espacio de funciones en el que el funcional está definido tenga estructura de espacio de Banach (espacio vectorial, en general de dimensión infinita, normado y completo).

Afortunadamente, en muchos de los casos de interés en Física, incluyendo todos los ejemplos presentados más arriba, se da tal circunstancia. Supondremos en lo sucesivo que los funcionales que vamos a considerar están definidos en un espacio de Banach. [Recordemos que un espacio de Banach es un espacio vectorial V , a cuyos elementos $f \in V$ se les puede dotar de una norma $\|f\|$ que satisface ciertas condiciones que son familiares en el ejemplo de la norma natural del espacio R^n : Para todo elemento f , $\|f\|$ es un número real positivo, nulo sólo cuando $f = 0$, con la propiedad de homogeneidad $\|\lambda f\| = |\lambda| \|f\|$ y satisfaciendo la desigualdad triangular. Con la topología asociada a esa norma el espacio es completo, es decir, toda sucesión de Cauchy tiene límite en el espacio V . Es importante tener presente que en cuanto se consideran situaciones en donde intervienen espacios de dimensión infinita, los problemas asociados con los dominios de definición no pueden dejarse de lado].

Definición. Sea V un cierto conjunto de funciones, que supondremos con estructura de espacio de Banach de dimensión infinita sobre los reales.

Un funcional real F es una aplicación $F: D(F) \rightarrow R$, que a cada función f del dominio $D(F)$ le asocia un valor real.

El dominio de F , $D(F)$ es un cierto subconjunto del espacio V , en el que $F(f)$ está definido.

Diremos que F es un funcional lineal si para cada par de funciones $f, g \in D(F)$, se verifica $F(\lambda f + \mu g) = \lambda F(f) + \mu F(g)$; $\forall \lambda, \mu \in R$.

Un ejemplo importante de funcional

Sea $C^1[a, b]$ el espacio de las funciones continuas de $[a, b]$ en R que admitan derivada primera continua en todo el intervalo $[a, b]$. $C^1[a, b]$ admite la estructura de espacio de Banach cuando le dotamos de la norma definida por:

$$\|f\| = \sup_{x \in [a, b]} |f(x)| + \sup_{x \in [a, b]} |f'(x)|.$$

En la topología inducida por esta norma, dos funciones son próximas cuando en todo el intervalo $[a, b]$, tanto las funciones como sus derivadas toman valores próximos; la distancia entre dos funciones f, g se define como la norma de la diferencia $f - g$.

Vamos ahora a introducir un funcional F de un tipo particular, que ha sido sugerido por los ejemplos presentados en la introducción. A este tipo de funcionales se referirá la mayor parte de los resultados concretos expuestos más adelante.

Comenzamos por definir su dominio, es decir, el conjunto sobre el que el funcional está definido:

$$\mathcal{D}(F) = \{f \in C^1[a, b] / f(a) = A; f(b) = B\} \quad (2)$$

donde A, B son constantes fijas. Consideremos a continuación una función de R^3 a R que denotaremos como $\Phi(x, y, z)$. Supondremos que esta función es continua y admite derivadas parciales continuas de primer orden. Es por lo tanto una función de clase uno en R^3 , propiedad que denotaremos como $\Phi(x, y, z) \in C^1(R^3)$. La forma explícita del funcional F está dada por:

$$\mathcal{F}(f) = \int_a^b \Phi(x, f(x), f'(x)) dx, \quad \forall f \in \mathcal{D}(F), \quad (3)$$

donde $f'(x)$ indica la derivada de la función $f(x)$. La integral está bien definida

$f \in D(F)$, pues el integrando es una función continua de x en el intervalo compacto $[a, b]$, y por lo tanto la correspondiente integral de Riemann siempre existe.

Nótese que pueden existir funcionales de muchos otros tipos. Por ejemplo, el funcional podría estar dado por una expresión en donde la función f no aparezca bajo una integral (ejemplo, el funcional $F(f) = f(c)$, donde c es un punto dado del intervalo $[a, b]$; tales funcionales aparecen en relación con la teoría de distribuciones y la delta de Dirac) o bien podría estar dado por una integral cuyo integrando dependiera de la derivada segunda o incluso de derivadas de órdenes superiores de la función f . A lo largo de estas notas, aunque daremos las definiciones en la forma general, nos restringiremos a la consideración de funcionales como el definido en (2)-(3).

En general $D(F)$ no es un subespacio vectorial de $C^1[a, b]$, salvo que $f(a) = 0$ y $f(b) = 0$. Sin embargo, el dominio $D(F)$ está asociado muy directamente a un cierto subespacio vectorial M de $C^1[a, b]$: $M = \{h \in C^1[a, b] / h(a) = h(b) = 0\}$.

Es claro que M sí que es un subespacio vectorial. Todo entorno de f en $D(F)$ es de la forma $h + U$ donde U es un entorno de 0 en M . El subespacio lineal M es denso en $C^1[a, b]$; análogamente, el dominio $D(F)$ es denso en $C^1[a, b]$.

Diferenciabilidad de Funcionales

Comencemos recordando la definición de diferenciabilidad para funciones reales de varias variables. Se dice que una función $f: R^n \rightarrow R$ es diferenciable en el punto $x = (x_1, x_2, \dots, x_n)$ si existe una función lineal denotada $df_x: R^n \rightarrow R$ tal que en un cierto entorno de x se tenga

$$f(x+h) - f(x) = df_x(h) + \epsilon(x, h)\|h\|,$$

donde $\epsilon(x, h) \rightarrow 0$ cuando $\|h\| = \sqrt{h_1^2 + h_2^2 + \dots + h_n^2} \rightarrow 0$.

geoméricamente, supuesto el punto x fijo, la función afín dada en un cierto entorno de x por $x+h \rightarrow f(x) + df_x(h)$ es la aplicación tangente a la función f en el punto $f(x)$, y su gráfica es el hiperplano tangente a la gráfica de f en el punto $(x, f(x))$. En cualquier texto de Análisis Matemático pueden encontrarse las demostraciones de las siguientes propiedades importantes:

- Si f es diferenciable en x , entonces df_x es única.

- Si la función f tiene un mínimo en el punto x (o un máximo, o en general un valor estacionario), entonces en el punto x la diferencial $df_x(h)$ se anula, $df_x(h) = 0$.
- La expresión explícita de la diferencial df_x es:

$$df_x(h) = \left. \frac{\partial f}{\partial x_1} \right|_x h_1 + \left. \frac{\partial f}{\partial x_2} \right|_x h_2 + \cdots + \left. \frac{\partial f}{\partial x_n} \right|_x h_n,$$

es decir, df_x es una función lineal de $h = (h_1, h_2, \dots, h_n)$ cuyos coeficientes son las derivadas parciales de f en el punto x . En particular, esta relación implica que si x es un mínimo de la función f (o en general un punto estacionario de f), todas las derivadas parciales de f en el punto x se anulan.

Los puntos en los que la diferencial de la función diferenciable f se anula se denominan puntos críticos de la función f . Para funciones de una variable los puntos críticos aislados son, bien máximos relativos estrictos, bien mínimos relativos estrictos, bien puntos de inflexión con tangente horizontal. Para funciones de dos variables puede haber puntos críticos aislados (bien máximos relativos estrictos, mínimos relativos estrictos o puntos de ensilladura), o líneas críticas, formadas por puntos críticos no aislados (típicamente, cuando la función presenta un comportamiento tipo bañera, con una línea de mínimos relativos no estrictos).

Extremales de funcionales

Para funciones de varias variables, los puntos críticos incluyen aquellos en los que la función alcanza bien un máximo o bien un mínimo relativo. La idea análoga para funcionales es la de función extremal de un funcional, que también se denominan puntos críticos o funciones críticas del funcional. Las extremales más simples son los máximos y mínimos, cuyas definiciones son evidentes:

Definición. Diremos que el funcional F tiene un mínimo absoluto en la función $f \in D(F)$ si para cualquier función $g \in D(F)$ se verifica $F(g) > F(f)$. Diremos que el funcional F tiene un mínimo relativo en la función $f \in D(F)$ si para cualquier función $g \in D(F)$ en un cierto entorno de f se verifica $F(g) > F(f)$.

Las definiciones de máximo absoluto y relativo son análogas. No vamos a entrar en las modificaciones para distinguir entre mínimos estrictos ($<$) o no estrictos (\leq) que son completamente análogas a las pertinentes para funciones de varias variables.

Ahora podemos investigar los análogos de los tres resultados básicos enunciados antes para funciones de varias variables. Vamos a enunciarles primero, y luego pasaremos a su demostración.

- Si el funcional F es diferenciable en la función f , la variación primera $\delta Ff(h)$ es única.
- Si el funcional F tiene un mínimo (o un máximo) relativo estricto en la función f , entonces la variación primera $\delta Ff(h)$ del funcional F en la función f se anula idénticamente, $\delta Ff(h) = 0$.
- Para el funcional dado en (2-3), la expresión explícita de la variación primera $\delta Ff(h)$ es el funcional lineal:

$$\delta \mathcal{F}_f(h) = \int_a^b \left(\frac{\partial \Phi}{\partial f} - \frac{d}{dx} \frac{\partial \Phi}{\partial f'} \right) h(x) dx,$$

Variación primera y ecuación de Euler-Lagrange

Vamos ahora a calcular la variación primera del funcional del tipo importante que hemos presentado en (3). Se trata del funcional definido en $C^1[a, b]$, espacio de funciones continuas de $[a, b]$ en R que admitan derivada primera continua en todo el intervalo $[a, b]$, con dominio y definición siguientes:

$$\mathcal{D}(\mathcal{F}) = \{f \in C^1[a, b] / f(a) = A; f(b) = B\}, \quad \mathcal{F}(f) = \int_a^b \Phi(x, f(x), f'(x)) dx,$$

donde A, B son constantes fijas y la función $\Phi(x, y, z) : R^3 \rightarrow R$ admite derivadas parciales continuas de primer orden. Vamos a:

- Demostrar que este funcional siempre es diferenciable.
- Calcular su variación primera.

Comencemos dando una expresión auxiliar importante. Para ello tomaremos un punto (x, y, z) de R^3 . Puesto que $\Phi(x, y, z)$ es continua y admite derivadas parciales continuas en todo R^3 , es diferenciable en todos los puntos. Existirá entonces un entorno de (x, y, z) , tal que si (x_1, y_1, z_1) pertenece a este entorno se verifica:

$$\Phi(x_1, y_1, z_1) - \Phi(x, y, z) = \frac{\partial \Phi}{\partial x} \Big|_{\mathbf{x}} (x_1 - x) + \frac{\partial \Phi}{\partial y} \Big|_{\mathbf{x}} (y_1 - y) + \frac{\partial \Phi}{\partial z} \Big|_{\mathbf{x}} (z_1 - z) + \epsilon \|(x_1 - x, y_1 - y, z_1 - z)\| \quad (4)$$

donde las derivadas parciales están evaluadas en el punto (x, y, z) y \mathcal{E} es una función dependiente de $(x, y, z; x_1, y_1, z_1)$ y que tiende a cero cuando $(x_1, y_1, z_1) \rightarrow (x, y, z)$.

La norma $\|(x_1 - x, y_1 - y, z_1 - z)\|$ se refiere a la norma canónica del espacio \mathbb{R}^3 , esto es $\|(x_1 - x, y_1 - y, z_1 - z)\| = \sqrt{(x_1 - x)^2 + (y_1 - y)^2 + (z_1 - z)^2}$.

Vamos ahora a usar esta fórmula para evaluar la diferencia entre el valor del funcional en la función f y en otra función $f + h$ en un cierto entorno de f (o lo que es lo mismo, para h en un cierto entorno de cero en M):

$$\mathcal{F}(f+h) - \mathcal{F}(f) = \int_a^b \{ \Phi(x, f(x) + h(x), f'(x) + h'(x)) - \Phi(x, f(x), f'(x)) \} dx.$$

Ahora llevamos (4) a esta fórmula, a condición de que $h(x)$ este en un cierto entorno de 0 en M . De esta manera, para h en dicho entorno:

$$\mathcal{F}(f+h) - \mathcal{F}(f) = \int_a^b \left\{ \frac{\partial \Phi}{\partial f} h(x) + \frac{\partial \Phi}{\partial f'} h'(x) \right\} dx + \int_a^b \epsilon(x) \sqrt{(h(x))^2 + (h'(x))^2} dx,$$

en donde las derivadas de Φ están evaluadas ambas en el punto $(x, f(x), f'(x))$, y en el término de la derecha $\epsilon(x) \equiv \epsilon(x, f(x) + h(x), f'(x) + h'(x); x, f(x), f'(x))$ tiende a cero cuando $h \rightarrow 0$. Nótese que para f, h fijas, podemos considerar ϵ como una función de x . Denotemos ahora:

$$\varphi_f(h) = \int_a^b \left\{ \frac{\partial \Phi}{\partial f} h + \frac{\partial \Phi}{\partial f'} h' \right\} dx, \quad C(f, h) = \int_a^b \epsilon(x) \sqrt{(h(x))^2 + (h'(x))^2} dx.$$

En cuanto a $\varphi_f(h)$, se trata de un funcional lineal que está bien definido en todo M . En efecto, la integral que lo define existe ya que el integrando es una función continua de x en el intervalo $[a, b]$. Su linealidad es evidente.

Si demostramos que cuando $h \rightarrow 0$ el término restante $C(f, h)$ tiende a 0 más rápidamente que la norma de h , entonces, identificando con la descomposición en la definición de funcional diferenciable, habremos demostrado que el funcional F es diferenciable, y de paso habremos obtenido una expresión para la variación primera de F en f .

Así pues, queda probado que F es diferenciable y que su variación primera en f viene dada por el término $\varphi_f(h)$, es decir

$$\delta \mathcal{F}_f(h) = \int_a^b \left\{ \frac{\partial \Phi}{\partial f} h + \frac{\partial \Phi}{\partial f'} h' \right\} dx$$

donde, insistimos, las derivadas parciales se evalúan en el punto $(x, f(x), f'(x))$.

Vamos ahora a transformar esta expresión, haciendo uso del recurso de eliminar la derivada h' mediante una integración por partes:

$$\int_a^b \frac{\partial \Phi}{\partial f'} h' dx = \frac{\partial \Phi}{\partial f'} h \Big|_a^b - \int_a^b \frac{d}{dx} \frac{\partial \Phi}{\partial f'} h dx = - \int_a^b \frac{d}{dx} \frac{\partial \Phi}{\partial f'} h dx.$$

Sustituyendo en la expresión para la variación primera del funcional F se obtiene

$$\delta F_f(h) = \int_a^b \left\{ \frac{\partial \Phi}{\partial f} - \frac{d}{dx} \frac{\partial \Phi}{\partial f'} \right\} h(x) dx \quad (5)$$

Tenemos ya los ingredientes necesarios para determinar las funciones que hacen máximo o mínimo el funcional F . En dichas funciones, la primera variación del funcional debe ser idénticamente nula, lo que significa que si en la función f el funcional F es mínimo o máximo, debe satisfacerse la condición

$$\delta F_f(h) = \int_a^b \left\{ \frac{\partial \Phi}{\partial f} - \frac{d}{dx} \frac{\partial \Phi}{\partial f'} \right\} h(x) dx = 0$$

para cualquier $h \in M$, esto es, satisfaciendo las condiciones $h(a) = h(b) = 0$.

Vamos a demostrar rigurosamente que esto implica que el termino entre corchetes en el integrando debe ser nulo. La demostración se basa en dos lemas:

Lema I. Sea $\gamma(x)$ una función continua en el intervalo $[a, b]$ tal que para toda función $h(x) \in C^1[a, b]$ con $h(a) = h(b) = 0$ se satisface la condición

$$\int_a^b \gamma(x) h'(x) dx = 0$$

Entonces $\gamma(x)$ es una constante.

Lema II. Sean $\alpha(x)$ y $\beta(x)$ dos funciones continuas en el intervalo $[a, b]$ tales que para toda función $h(x) \in C^1[a, b]$ con $h(a) = h(b) = 0$ se satisface la condición:

$$\int_a^b \{ \alpha(x) h(x) + \beta(x) h'(x) \} dx = 0.$$

Entonces $\beta(x)$ es diferenciable y $\beta'(x) = \alpha(x)$; $\forall x \in [a, b]$.

Vamos ahora a usar estos dos lemas para demostrar la condición necesaria conocida como *ecuación de Euler-Lagrange* para que un funcional del tipo (2-3) tenga un máximo o un mínimo en f . En la expresión obtenida antes,

$$\delta F_f(h) = \int_a^b \left\{ \frac{\partial \Phi}{\partial f} h + \frac{\partial \Phi}{\partial f'} h' \right\} dx = 0,$$

aplicamos el Lema II, con $\alpha(x)$ en el papel de $\partial \Phi / \partial f$ mientras que $\beta(x)$ hace el papel de $\partial \Phi / \partial f'$. Obtenemos así que $\partial \Phi / \partial f$ es diferenciable con respecto a x y que su derivada con respecto a x debe ser igual a $\partial \Phi / \partial f$. Esto es, el funcional F del tipo (2)(3) tiene un máximo o un mínimo en

$$f \Rightarrow \frac{d}{dx} \frac{\partial \Phi}{\partial f'} - \frac{\partial \Phi}{\partial f} = 0$$

Esta es la llamada ecuación de Euler-Lagrange, ecuación diferencial para la función $f(x)$ en la cual el funcional F alcanza mínimos o máximos relativos.

En el caso de las funciones $f(x)$, en un máximo o un mínimo la diferencial en el punto x , df_x se anula. Pero de la anulación de la diferencial en x no se sigue que la función tenga en x un máximo o un mínimo, sino sólo que el punto x es un punto crítico. Para funcionales, la situación es semejante, y mientras que la primera variación del funcional F se anula en un mínimo o un máximo, de la anulación de la primera variación no se sigue que el funcional F tenga un máximo o un mínimo. Se definen en general las extremales del funcional diferenciable F como aquellas funciones en las que la primera variación $\delta F f$ se anula idénticamente. Además de las funciones en las que el funcional tiene un mínimo o un máximo relativo, las extremales incluyen otras funciones en las que el funcional es estacionario, de manera análoga al caso de los puntos críticos que además de los mínimos o máximos incluyen puntos de inflexión con tangente horizontal, puntos de silla con diferentes signaturas, etc.

Podemos formular el resultado importante obtenido en esta sección mediante:

el funcional F tiene una extremal en $f \Leftrightarrow \frac{d}{dx} \frac{\partial \Phi}{\partial f'} - \frac{\partial \Phi}{\partial f} = 0$

Extremales que no sean de clase C^1

La exigencia que hemos hecho a las funciones de ser continuas y con derivada continua es una exigencia técnica, que permite dar una condición necesaria muy simple —la ecuación de Euler-Lagrange— para la existencia de extremales. Pero conviene mencionar que en muchos casos los problemas variacionales nos obligan a salirnos de este marco. Es decir, hay casos en los que no existen

extremales que sean suficientemente regulares, pero hay extremales que no son regulares.

Variación segunda de un funcional

La anulación de la variación primera de un funcional en la función f es una condición necesaria para que el funcional F sea extremal en f . Pero esa anulación no garantiza que el funcional tenga un mínimo relativo en f . Se trata de la misma situación que ocurre para las funciones, donde un mínimo relativo en el punto x requiere la anulación de la diferencial df_x en x como condición suficiente, pero tal condición se da también en un máximo y en un punto estacionario; si deseamos que la función tenga un mínimo debemos exigir la condición de que la diferencial segunda $df^2_x(h)$ sea una forma cuadrática definida positiva. Aunque en la mayor parte de las aplicaciones lo que resulta ser realmente importante es la condición de extremalidad (y no la de ser precisamente mínimo), resulta conveniente conocer el análogo de la diferencial segunda de una función de varias variables, que para funcionales se denomina *variación segunda*.

Para acabar esta sección, vamos a obtener una fórmula para la variación segunda del funcional (2-3). Ya que en el contexto que nos interesa aquí la segunda variación sólo se necesita para discriminar entre los diversos tipos de extremales (máximos, mínimos o tipo silla), basta con calcular la segunda variación en funciones f para las que la primera variación es ya idénticamente nula.

De esta manera nos queda la siguiente expresión para la variación segunda de un funcional del tipo usual en una función f que anule la variación primera, en estos términos:

$$\delta^2 \mathcal{F}_f(h) = \frac{1}{2} \int_a^b \left\{ \left(\frac{\partial^2 \Phi}{\partial f^2} - \frac{d}{dx} \frac{\partial^2 \Phi}{\partial f \partial f'} \right) h^2 + \frac{\partial^2 \Phi}{\partial (f')^2} (h')^2 \right\} dx, \text{ cuando } \delta \mathcal{F}_f = 0$$

Problemas variacionales con ligaduras (Problemas isoperimétricos).

Se presenta con frecuencia el problema de encontrar los puntos críticos (mínimos, máximos, etc) de una función de varias variables que no son independientes sino que están sujetas a una o varias condiciones adicionales, conocidas como ligaduras. El ejemplo más sencillo y fácil de visualizar es el de la búsqueda del máximo (o mínimo) de una función de dos variables $f(x, y)$ sobre una determinada curva Γ en el plano x, y ; tales máximos y mínimos condicionados ocurren en puntos en los que la función considerada como función de dos variables independientes no tiene máximos ni mínimos. Un ejemplo de comprensión inmediata: cuando se sigue un camino en la ladera de una montaña, la altura puede tener máximos o mínimos relativos a lo largo del camino, que en general no corresponden a máximos o mínimos de la función que da la altura de la superficie en cada punto de la montaña.

La condición de anulación de la diferencial (o la equivalente de anulación de todas las derivadas parciales) no resulta aplicable en tales casos; geoméricamente esto es claro, ya que un máximo o mínimo a lo largo de la curva sólo debe traducirse en la anulación de la derivada direccional de la función a lo largo de la dirección de la curva.

Procedimiento de fuerza bruta: usemos la condición adicional para eliminar una de las dos variables, y consideremos la función, ya restringida a la curva, como una función de una variable independiente, a la que se le puedan aplicar las condiciones usuales de máximo o mínimo. Este método de fuerza bruta dista de ser práctico. Aunque se pueda eliminar la variable (o variables) que debido a las ligaduras resultan dependientes, las expresiones que se obtienen en términos de variables independientes pueden ser poco manejables. Y además puede ocurrir que las ligaduras estén dadas en forma implícita, que no admita la eliminación explícita.

Se debe a Lagrange un método de determinación de máximos y mínimos de funciones sometidas a ciertas condiciones adicionales que se conoce como método de los multiplicadores de Lagrange, y consiste esencialmente en que si un punto $x = (x_1, x_2, \dots, x_n)$ es un máximo o mínimo de la función $f(x_1, x_2, \dots, x_n)$ sobre la subvariedad determinada por las condición adicional $g(x_1, x_2, \dots, x_n) = 0$, entonces el punto $x = (x_1, x_2, \dots, x_n)$ es un máximo o mínimo de la

función $f(x_1, x_2, \dots, x_n) + \lambda g(x_1, x_2, \dots, x_n)$, considerada como función de n variables independientes.

La constante λ , conocida como multiplicador de Lagrange, queda determinada junto con la posición de los posibles puntos estacionarios, al resolver las ecuaciones que establecen que en tales puntos todas las derivadas parciales de la función $f(x_1, x_2, \dots, x_n) + \lambda g(x_1, x_2, \dots, x_n)$ deben anularse. Los detalles de este método pueden consultarse en cualquier texto de Análisis Matemático.

En el cálculo variacional aparecen también naturalmente problemas con condiciones adicionales. Hemos visto dos. En el problema de la catenaria, es evidente por razones físicas que sin ninguna condición adicional, el funcional “energía potencial” de un hilo en un campo gravitatorio uniforme no presenta mínimos (entre dos puntos dados, para cualquier hilo con energía potencial dada, siempre podemos tender un hilo más largo, cuya energía potencial sea menor). Pero si consideramos hilos de longitud prefijada, entonces sí que debemos esperar un mínimo para cierta forma del hilo, especificada por cierta función $z = f(x)$, que deberá satisfacer la condición adicional de tener la longitud dada. En el problema de determinar la curva que encierre un área máxima, es de nuevo claro que sin ninguna condición adicional, podemos encerrar áreas cada vez mayores y mayores. Sólo si imponemos una condición extra (longitud prefijada) debemos esperar que cierta forma de la curva encierre un área máxima. En ambos casos, la condición adicional está dada por otro funcional. Vamos a discutir este problema en el caso más sencillo de que tanto el funcional a minimizar como la ligadura sean del tipo (2-3).

Consideremos pues un funcional en el espacio $C^1[a, b]$, del tipo:

$$\mathcal{F}(f) = \int_a^b \Phi(x, f(x), f'(x)) dx, \quad \mathcal{D}(\mathcal{F}) = \{f \in C^1[a, b] / f(a) = A; f(b) = B\}.$$

y consideremos otro funcional G , cuyo dominio supondremos el mismo que el de F y dado por:

$$\mathcal{G}(f) = \int_a^b \Gamma(x, f(x), f'(x)) dx$$

donde Γ es una función continua y con derivadas continuas que juega, para el funcional G , un papel análogo al que Φ juega para F .

Problema isoperimétrico. Entre todas las funciones que satisfagan la condición $G(f) = G$ donde G es una constante real, encontrar las extremales del funcional F . En estas condiciones puede demostrarse el siguiente:

Teorema. Sea $f \in D(F)$ una función extremal del funcional F satisfaciendo la condición $G(f) = G$. Supongamos además que la primera variación del funcional G en f no es idénticamente nula. Entonces existe un número real λ de manera que f es un extremal del nuevo funcional

$$\tilde{F}(f) = \int_a^b \{ \Phi(x, f(x), f'(x)) + \lambda \Gamma(x, f(x), f'(x)) \} dx,$$

en el que ya no se considera ninguna condición subsidiaria.

La demostración de este teorema, así como su extensión para el caso de que existan varias condiciones de ligadura puede consultarse en el libro de Troutman.

Problemas Variacionales con varios grados de libertad

Hasta ahora nos hemos limitado a considerar funcionales definidos en espacios de funciones reales de una variable, $f : R \rightarrow R$. Pero pueden darse funcionales de tipos más generales, por ejemplo funcionales del tipo (2-3) en las que Φ dependa de funciones con varias componentes (vectoriales) pero de una sola variable, o bien funcionales definidos sobre espacios de funciones de más de una variable. En tanto intervengan funciones de una sola variable, posiblemente con varias componentes, esto es, funciones de $f : R \rightarrow R^n$, la mayor parte de las técnicas y resultados descritos en estas notas se extienden directamente y de manera casi inmediata. Por ejemplo, el principio de mínima acción determina el movimiento real que sigue una partícula en un potencial externo $V(x, t)$; este movimiento es una función $x : R \rightarrow R^3$, que puede describirse mediante tres funciones componentes, $x(t), y(t), z(t)$. En este caso el funcional que se pretende minimizar es siempre del tipo (2-3), donde ahora la función Φ depende de t y de las tres componentes $x(t), y(t), z(t)$, así como de las tres derivadas $dx(t)/dt, dy(t)/dt, dz(t)/dt$. En el principio de Fermat, la trayectoria seguida por un rayo de luz esta descrita también por una función que podemos describir dando las dos funciones $y(x), z(x)$, y el funcional a minimizar involucra las dos funciones $y(x), z(x)$ y sus derivadas $dy(x)/dx, dz(x)/dx$.

Estos ejemplos sugieren extender el tipo usual de funcionales (2-3) de la siguiente manera: Denotemos $C^1([a, b], \mathbb{R}^n)$ el espacio de las funciones definidas en un intervalo $[a, b]$ y con valores en \mathbb{R}^n que sean continuas y todas cuyas funciones componentes admitan derivada primera continua en todo el intervalo $[a, b]$. Este espacio admite la estructura de espacio de Banach cuando le dotamos de la norma definida por:

$$\|f\| = \sup(\|f_1\|, \|f_2\|, \dots, \|f_n\|),$$

donde para cada función componente la norma es la usada en (1). Es fácil demostrar que se trata de una norma, y menos fácil de demostrar aunque también cierto, que dotado de esta norma, el espacio $C^1([a, b], \mathbb{R}^n)$ es un espacio de Banach. En la topología inducida por esta norma, dos funciones f, g son próximas cuando en todo el intervalo $[a, b]$, tanto cada una de las componentes de f, g como sus derivadas toman valores próximos.

Vamos ahora a definir un funcional F que es la extensión del tipo descrito en (2-3). Comenzamos por definir su dominio, es decir, el conjunto sobre el que el funcional está definido:

$$\mathcal{D}(F) = \{f(x) \in C^1([a, b], \mathbb{R}^n) / f_i(a) = A_i; f_i(b) = B_i; i = 1, 2, \dots, n\}.$$

Notemos que este dominio no es un subespacio vectorial salvo en el caso $A_i = B_i = 0, i = 1, 2, \dots, n$. Sea ahora:

$$\mathcal{M} = \{h(x) \in C^1([a, b], \mathbb{R}^n) / h_i(a) = h_i(b) = 0, i = 1, 2, \dots, n\}.$$

Obviamente $\mathcal{D}(F) = \mathcal{M} + \mathcal{M}$ para todo $f \in \mathcal{D}(F)$

Sea ahora $\Phi(x_1, x_2, \dots, x_{2n+1})$ una función de \mathbb{R}^{2n+1} a \mathbb{R} que sea continua y admita derivadas parciales continuas. Escribamos

$$F(f) = \mathcal{F}(f_1, f_2, \dots, f_n) = \int_a^b \Phi(x, f_1(x), \dots, f_n(x), f_1'(x), \dots, f_n'(x)) dx.$$

F está bien definido $\forall f \in \mathcal{D}(F)$, puesto que la función bajo el signo integral es continua en todo el intervalo $[a, b]$.

Repitiendo lo hecho en el caso de una función de una componente, se demuestra que este funcional es diferenciable en todos los puntos de su dominio. Sea entonces $f = (f_1, f_2, \dots, f_n) \in \mathcal{D}(F)$ y $h = (h_1, h_2, \dots, h_n) \in \mathcal{M}$. Como Φ es una función continua con derivadas parciales continuas, es diferenciable.

Razonando como en la derivación de (5) existirá un entorno de cero en M tal que

$$\begin{aligned} & \mathcal{F}(f_1 + h_1, \dots, f_n + h_n) - \mathcal{F}(f_1, \dots, f_n) = \\ & \int_a^b \{ \Phi(x, f_1 + h_1, \dots, f_n + h_n, f'_1 + h'_1, \dots, f'_n + h'_n) - \Phi(x, f_1, \dots, f_n, f'_1, \dots, f'_n) \} dx = \\ & \int_a^b \sum_{i=1}^n \left\{ \frac{\partial \Phi}{\partial f_i} h_i + \frac{\partial \Phi}{\partial f'_i} h'_i \right\} dx + \int_a^b \epsilon \|(0, h_1, \dots, h_n, h'_1, \dots, h'_n)\| dx, \end{aligned}$$

para todo h en dicho entorno. La primera integral está bien definida para todo $h = (h_1, \dots, h_n) \in M$ y es una aplicación lineal de M en R . La segunda puede ponerse en la forma $E(f, h) \|h\|$, con $E(f, h) \rightarrow 0$ si $h \rightarrow 0$. De esta manera, para h en un entorno de 0 en M ,

$$\mathcal{F}(f_1 + h_1, \dots, f_n + h_n) - \mathcal{F}(f_1, \dots, f_n) = \varphi_f(h) + \mathcal{E}(f, h) \|h\|,$$

con

$$\varphi_f(h) = \int_a^b \sum_{i=1}^n \left\{ \frac{\partial \Phi}{\partial f_i} h_i + \frac{\partial \Phi}{\partial f'_i} h'_i \right\} dx$$

Esta expresión muestra que F es diferenciable, y su diferencial primera viene dada por:

$$\delta \mathcal{F}_f(h) = \int_a^b \sum_{i=1}^n \left\{ \frac{\partial \Phi}{\partial f_i} h_i + \frac{\partial \Phi}{\partial f'_i} h'_i \right\} dx.$$

Supongamos que F admite un extremal en $f = (f_1, \dots, f_n) \in D(F)$. Entonces, la variación primera de F en f se ha de anular. Esto significa que para todo $h \in M$, tenemos:

$$\int_a^b \sum_{i=1}^n \left\{ \frac{\partial \Phi}{\partial f_i} h_i + \frac{\partial \Phi}{\partial f'_i} h'_i \right\} dx = 0$$

Las derivadas parciales se evalúan en $(x, f_1(x), \dots, f_n(x), f'_1(x), \dots, f'_n(x))$. Escogiendo $h_i(x) = 0$ para todo i salvo para $i = j$, queda

$$\int_a^b \left\{ \frac{\partial \Phi}{\partial f_j} h_j + \frac{\partial \Phi}{\partial f'_j} h'_j \right\} dx = 0.$$

Esto implica, tras el Lema II visto antes que

$$\frac{d}{dx} \frac{\partial \Phi}{\partial f'_j} - \frac{\partial \Phi}{\partial f_j} = 0.$$

Realizando la misma operación para todos los $j = 1, 2, \dots, n$ resulta un sistema de n ecuaciones en las n funciones incógnitas f_1, f_2, \dots, f_n entre cuyas soluciones están los extremales $f = (f_1, f_2, \dots, f_n)$ del funcional F . A dicho sistema de ecuaciones se le conoce como sistema de Euler-Lagrange, o simplemente, ecuaciones de Euler-Lagrange:

$$f \text{ es un extremal del funcional } \mathcal{F} \Leftrightarrow \frac{\partial \Phi}{\partial f_i} - \frac{d}{dx} \frac{\partial \Phi}{\partial f'_i} = 0, \quad i = 1, 2, \dots, n.$$

Problemas Variacionales con varias variables: superficies mínimas.

Otros problemas variacionales involucran funciones de dos o más variables como los objetos primitivos de los cuales depende algún funcional que se trata de minimizar.

Tal situación resulta ser mucho más complicada que el caso de funciones de una variable. El prototipo es el problema de las superficies mínimas: De entre todas las superficies en el espacio R^3 con un borde dado, encontrar aquellas que tengan área mínima. En este caso, el funcional a minimizar depende de una función de dos variables.

Históricamente, es notable que las ideas básicas del cálculo variacional, en la forma que las hemos expuesto, aparecieran por primera vez en un trabajo de Lagrange (1760) dedicado precisamente al estudio del problema nada trivial de las superficies mínimas. Este trabajo despertó el interés de Euler, dando lugar a un desarrollo por parte de ambos autores, que culminó en la sistematización de las condiciones hoy llamadas de Euler-Lagrange.

Vamos a limitarnos a derivar, de manera directa, la ecuación diferencial que debe satisfacer cualquier superficie mínima, y lo vamos a hacer poniendo solamente el énfasis en las ideas relevantes desde el punto de vista del cálculo variacional, eludiendo discutir detalles adicionales.

Localmente cualquier superficie puede describirse en la forma denominada de Monge, como la gráfica de una función $(x, y) \rightarrow (x, y, f(x, y))$, pero posiblemente tal representación puede no cubrir la superficie “completa”. Por ejemplo un plano puede representarse de manera completa en forma de Monge: $(x, y) \rightarrow (x, y, f(x, y) = z_0)$, pero para una esfera la región máxima representable de esta manera es un hemisferio, excluido su borde ecuatorial:

$$(x, y) \rightarrow (x, y, f(x, y) = \sqrt{R^2 - x^2 - y^2}).$$

Para simplicidad, nos limitaremos a estudiar porciones de superficie que sean representables de dicha forma, lo que no constituye ninguna limitación importante, ya que como veremos la condición de superficie mínima se traduce en una ecuación diferencial que determina f localmente.

El área de la porción de superficie que corresponde a un dominio D del plano de parámetros x, y es:

$$A = \iint_D dx dy \sqrt{1 + (f_x)^2 + (f_y)^2}$$

con la integral doble extendida al dominio D y en donde para abreviar la escritura denotamos $f_x \equiv \partial f / \partial x$, etc. Nótese la analogía de esta expresión con la que da la longitud de una curva plana descrita en la forma $x \rightarrow (x, f(x))$, dada por $\mathcal{L} = \int dx \sqrt{1 + (f_x)^2}$.

Consideremos una curva Γ dada en espacio R^3 . Esta curva se supondrá cerrada, sin autointersecciones y suficientemente regular, y es quien va a jugar el papel que los dos extremos $t_a, x_a; t_b, x_b$ jugaban para problemas variacionales del tipo implicado en el principio de mínima acción. La proyección de Γ sobre el plano x, y es una curva plana, que llamaremos γ , que también supondremos cerrada, sin autointersecciones y suficientemente regular. La propia curva Γ puede describirse como el conjunto de puntos $(x, y, z_\gamma(x, y))$ en donde se supone que $(x, y) \in \gamma$ donde $z(x, y)$ es la función fija, definida solamente en γ y que describe la altura de la curva Γ . Denotemos D el dominio del plano cuyo borde es γ : este dominio es homeomorfo a un disco ya que la curva γ no tiene autointersecciones.

La forma general de la descripción de Monge de una superficie que tenga a Γ como borde está dada por una función de dos variables, suficientemente regular, en la forma:

$$(x, y) \in D \rightarrow (x, y, z(x, y)), \quad \text{donde} \quad z(x, y) = z_\gamma(x, y) \text{ para } (x, y) \in \gamma$$

La idea esencial de la derivación de Lagrange es la siguiente. Supongamos que la función $f(x, y)$ (aún desconocida) corresponde a una superficie Σf con borde Γ y de área mínima entre todas las que satisfagan las condiciones anteriores. Sea $h(x, y)$ una función fija, suficientemente regular, definida en el dominio D , y a la que exigimos satisfacer la condición

$$h(x, y) = 0 \quad \text{para} \quad (x, y) \in \gamma$$

En estas condiciones, tenemos una familia de superficies, que podemos denotar mediante $\Sigma f(h, \mathcal{E})$ cuya descripción de Monge es:

$$(x, y) \in D \rightarrow (x, y, f(x, y) + \epsilon h(x, y)),$$

que se construyen a partir de la superficie Σf (aún desconocida), tomando como dato de deformación la función $h(x, y)$; aquí \mathcal{E} juega el papel de un parámetro, de manera que esta familia es una familia uniparamétrica de superficies, todas las cuales tienen a la curva Γ como borde, ya que para cualquier valor del parámetro \mathcal{E} se verifica la condición

$$f(x, y) + \epsilon h(x, y) = z(x, y) \quad \text{para} \quad (x, y) \in \gamma$$

El área de la superficie $\Sigma f(h, \mathcal{E})$ está dada por:

$$A_{h, \epsilon} = \iint dxdy \sqrt{1 + (f_x + \epsilon h_x)^2 + (f_y + \epsilon h_y)^2}$$

Si la superficie Σf (descrita por $f(x, y)$) tiene realmente área mínima entre todas las superficies con el mismo borde, también debe tener área mínima entre las de la familia uniparamétrica anterior $\Sigma f(h, \mathcal{E})$. Esto significa que la función $A_{h, \epsilon}$ debe tener un mínimo en $\mathcal{E} = 0$, es decir

$$0 = \left. \frac{dA_{h, \epsilon}}{d\epsilon} \right|_{\epsilon=0}.$$

Derivando con respecto a \mathcal{E} en $A_{h, \mathcal{E}}$ y evaluando en $\mathcal{E} = 0$ la condición anterior se transforma en:

$$\iint_D dx dy \frac{f_x h_x + f_y h_y}{\sqrt{1 + (f_x)^2 + (f_y)^2}} = 0 \quad (6)$$

Así pues, si la superficie Σf es mínima, la condición (6) debe satisfacerse para *cualquier* elección de la función auxiliar h que satisfaga la condición de anulación sobre γ .

Por analogía con lo estudiado anteriormente, el paso siguiente debe ser transformar la integral en (6) en otra integral que sea *lineal* en h , pero en donde no aparezcan las derivadas de h . La manera más clara de hacerlo es la siguiente. Consideremos la integral en (6) (que debe anularse) como una suma de dos sumandos:

$$\iint_D dx dy \frac{f_x h_x}{\sqrt{1+(f_x)^2+(f_y)^2}} + \iint_D dx dy \frac{f_y h_y}{\sqrt{1+(f_x)^2+(f_y)^2}}$$

Vamos a realizar la transformación de manera ligeramente diferente, aunque perfectamente análoga, sobre cada uno de estos dos sumandos. Comencemos con

$$\iint_D dx dy \frac{f_x h_x}{\sqrt{1+(f_x)^2+(f_y)^2}}$$

que escribiremos como (hágase un diagrama que aclare el uso de los límites de integración):

$$\int_{y_{\min}}^{y_{\max}} dy \int_{x_a(y)}^{x_b(y)} \frac{f_x}{\sqrt{1+(f_x)^2+(f_y)^2}} h_x dx.$$

donde y_{\min} , y_{\max} son los valores mínimo y máximo de la coordenada y sobre la curva γ , mientras que $x_a(y)$, $x_b(y)$ son los valores mínimo y máximo de x sobre el segmento de recta paralela al eje x y que tiene ordenada y . (Nota: para simplificar la discusión estamos suponiendo que el dominio es convexo, y que la intersección con las rectas paralelas a los ejes tiene sólo dos puntos; esta restricción simplifica la discusión pero no es esencial al resultado).

Hacemos ahora la integración en x por partes, tomando

$$u = \frac{f_x}{\sqrt{1+(f_x)^2+(f_y)^2}}, \quad v = h. \text{ Así obtenemos para la integral en } x \text{ lo siguiente:}$$

$$\frac{f_x}{\sqrt{1+(f_x)^2+(f_y)^2}} h \Big|_{x_a(y)}^{x_b(y)} - \int_{x_a(y)}^{x_b(y)} \left(\frac{\partial}{\partial x} \frac{f_x}{\sqrt{1+(f_x)^2+(f_y)^2}} \right) h dx.$$

El término de borde no contribuye debido a que los dos puntos $(x_a(y), y)$, $(x_b(y), y)$ están por construcción sobre el borde γ y la función $h(x, y)$ se anula sobre γ . Integrando ahora con respecto a y lo que obtenemos es que el término que implicaba a la derivada con respecto a x de h puede reescribirse como una integral en la que es la propia función h (y no su derivada) quien aparece como factor:

$$\iint_D dx dy \frac{f_x h_x}{\sqrt{1+(f_x)^2+(f_y)^2}} = - \iint_D dx dy \left(\frac{\partial}{\partial x} \frac{f_x}{\sqrt{1+(f_x)^2+(f_y)^2}} \right) h$$

La derivada parcial que aparece ahora en el integrando se calcula fácilmente: conviene recordar que tanto f_x como f_y son funciones de x, y . El resultado es:

$$\frac{\partial}{\partial x} \frac{f_x}{\sqrt{1 + (f_x)^2 + (f_y)^2}} = \frac{f_{xx}(1 + f_y^2) - f_x f_y f_{xy}}{\left(\sqrt{1 + (f_x)^2 + (f_y)^2}\right)^3}$$

de manera que finalmente, lo que encontramos es:

$$\iint dx dy \frac{f_x h_x}{\sqrt{1 + (f_x)^2 + (f_y)^2}} = - \iint dx dy \frac{f_{xx}(1 + f_y^2) - f_x f_y f_{xy}}{\left(\sqrt{1 + (f_x)^2 + (f_y)^2}\right)^3} h(x, y)$$

Para el otro sumando que involucra h_y se procede de manera análoga, pero intercambiando y por x . Es bastante evidente que tal procedimiento conduce a:

$$\iint dx dy \frac{f_y h_y}{\sqrt{1 + (f_x)^2 + (f_y)^2}} = - \iint dx dy \frac{f_{yy}(1 + f_x^2) - f_y f_x f_{xy}}{\left(\sqrt{1 + (f_x)^2 + (f_y)^2}\right)^3} h(x, y)$$

Así pues, la condición de que la superficie sea mínima, contenida en la ecuación (6), se convierte en:

$$\iint dx dy \frac{f_{xx}(1 + f_y^2) - 2f_x f_y f_{xy} + f_{yy}(1 + f_x^2)}{\left(\sqrt{1 + (f_x)^2 + (f_y)^2}\right)^3} h(x, y) = 0$$

y como esta ecuación debe satisfacerse para cualquier función $h(x, y)$ (con la sola exigencia de anularse sobre el borde γ), parece claro que la única posibilidad de que tal cosa ocurra es que el integrando se anule, esto es, que la función $f(x, y)$ satisfaga la ecuación:

$$f_{xx}(1 + f_y^2) - 2f_x f_y f_{xy} + f_{yy}(1 + f_x^2) = 0$$

que se conoce como ecuación de Lagrange para las superficies mínimas. A pesar de su aspecto superficialmente inocente, como ecuación diferencial es bastante complicada: es muy no lineal y se conocen muy pocas soluciones explícitas. La búsqueda efectiva de superficies mínimas requiere el uso de técnicas mucho más avanzadas y elaboradas. Un dominio cualquiera de un plano es evidentemente una superficie mínima, cuyo borde es una curva plana. Escogiendo adecuadamente las coordenadas, esta porción de superficie está descrita por $f(x, y) = z_0$, que satisface trivialmente la ecuación de Lagrange. Es decir, si la curva Γ es una curva plana, la superficie mínima con borde Γ es una porción de plano. Este ejemplo es absolutamente trivial.

A finales del S. XVIII se obtuvieron otras dos superficies mínimas relativamente sencillas. Una es el catenoide, que es la única superficie mínima de revolución.

El otro ejemplo de superficie mínima es el helicoide recto, que es la superficie engendrada por una recta "horizontal" que se desliza a velocidad constante a lo

largo de un “eje” vertical al tiempo que gira alrededor de dicho eje, en un plano “horizontal” y a velocidad angular constante.

Durante más de 200 años el catenoide y el helicoide han sido las únicas superficies mínimas conocidas que satisfacen las condiciones de ser embebidas en \mathbb{R}^3 , completas y sin autointersecciones. Por ello ha resultado una agradable noticia para la comunidad matemática el descubrimiento a principios de los 80 del siglo pasado de una nueva superficie mínima que satisface la exigencias anteriores: *la superficie de Costa*. Las técnicas, apoyadas en el análisis de funciones de variable compleja, que han llevado a este descubrimiento han abierto la puerta a una auténtica eclosión de un mundo fascinante y mucho más rico de superficies mínimas. Una descripción puede verse en el libro *El turista matemático* de I. Petersen, y sobre *la superficie de Costa* hay un artículo de C. J. Costa en *La Gaceta Matemática*, 4 (1999). Actualmente se conocen multitud de nuevos ejemplos. La portada del *Notices of the American Mathematical Society de Diciembre de 2000* se dedica a una de ellas.

En <http://www.susqu.edu/brakke> hay cantidad de información sobre superficies mínimas triplemente periódicas.

En particular, merece la pena indicar que para una porción de superficie arbitraria Σ_f pero con borde fijo Γ , la variación primera del funcional área está dada por:

$$A(\Sigma_f) = \iint_D dx dy \sqrt{1 + (f_x)^2 + (f_y)^2} \quad \delta A_{\Sigma_f}(h) = \iint_{\Sigma_f} d\Sigma_f H_{\Sigma_f} h$$

donde la función H_{Σ_f} es la llamada curvatura media de la superficie, Σ_f , definida como la semisuma de las dos curvaturas principales, que a su vez son las curvaturas máxima y mínima de las curvas planas que se obtienen como secciones normales de la superficie. La interpretación geométrica de *la condición de Lagrange*, dada por vez primera por *Meusnier*, es que las superficies mínimas tienen curvatura media igual a cero en todos sus puntos, lo que evidentemente garantiza la anulación del funcional variación primera.

Para acabar, conviene mencionar que las ideas básicas (diferenciabilidad de funcionales, funcional lineal variación primera, anulación de dicho funcional como condición de extremalidad, etc.) se extienden a estos problemas. Aunque no hemos escrito de manera general el problema, puede comprobarse que para un funcional del tipo

$$\mathcal{F}(f) = \iint_D dx dy \Phi(x, y; f(x, y), f_x(x, y), f_y(x, y)) dx,$$

sobre un dominio D del plano y con condiciones de frontera sobre el borde y de D del tipo

$$(x, y) \in D \rightarrow (x, y, z(x, y)), \quad \text{donde} \quad f(x, y) = f_\gamma(x, y) \text{ para } (x, y) \in \gamma$$

las ecuaciones de Euler-Lagrange que se obtienen son:

$$\frac{\partial \Phi}{\partial f} - \frac{\partial}{\partial x} \frac{\partial \Phi}{\partial f_x} - \frac{\partial}{\partial y} \frac{\partial \Phi}{\partial f_y} = 0.$$

forma de la que la extensión a más variables independientes, o a funciones vectoriales de varias variables resulta ya evidente.