

Capítulo 1: Cálculo Funcional

Cálculo de Variaciones se refiere a la optimización de funcionales. Las ideas fundamentales fueron desarrolladas por Euler (1701-1783) y por Lagrange (1736-1813).

El calculo variacional esta basado en una generalización del problema básico de cálculo diferencial en el cual estamos interesados en encontrar mínimos (o máximos) de funcionales. Para nuestros intereses, los funcionales asignan un numero real a cada función de una clase bien definida. Para ser mas precisos definamos el conjunto A de funciones. El funcional S asigna a cada miembro $f \in A$ un numero real definido por $S(f)$

$$S : A \rightarrow R$$

El problema básico que vamos a encontrar en la Mecánica Analítica es encontrar la función $f \in A$ que permite un mínimo (mínimo local) de S . Este tipo de problemas se da en Geometría, Física, Ingeniería, Teoría de Control Óptimo, etc.

En pocas palabras re-formularemos la mecánica Newtoniana de fuerzas en términos de minimizar un funcional S . Entonces una trayectoria esta dada por la función única que minimiza este funcional S .

Notación: En este libro usaremos la siguiente notación:

1. font **negrita x** para describir una lista de números, que por lo general representan los componentes de un vector en una representación o sistema de referencia dado.
2. font **negrita con flecha \vec{x}** para describir el objeto geométrico vector que es independiente del sistema de referencia o representación.
3. font **normal** para describir un escalar que podría ser por ejemplo un componente x_i de una lista.
4. la magnitud de un vector se escribe como $|x|$

Índice

1. Ejemplos de Funcionales	3
2. Calculo Funcional en una dimensión	5
2.1. Cálculo diferencial	5
2.2. Derivadas Funcionales	6
2.3. La ecuación de Euler-Lagrange	8
2.3.1. Constantes de movimiento	10
2.3.2. Existencia de una solución	10
3. Generalizaciones	10
3.1. Caso 1: Derivadas de alto orden	10
3.2. Caso 2: Varias funciones	11
3.3. Caso 3: Múltiples integrales	11
3.4. Condiciones naturales de borde	12
4. Teorema para transformación de coordenadas	13
4.1. Restricciones	14
5. Solución Numérica	18
5.1. Shooting method	18
5.2. Elementos finitos en 1D	19
5.3. Elementos finitos en 2D	21

1. Ejemplos de Funcionales

Hay muchos tipos de funcionales. Por ejemplo,

$$S[f] = f(1)e^5$$

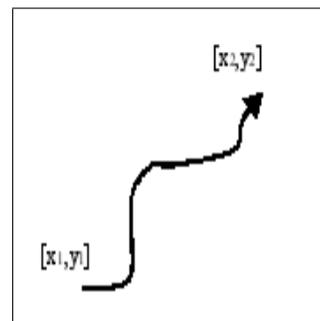
Pero en general estamos interesados en funcionales dados por integrales de funciones.

Ejemplo: Cual es la distancia mínima entre dos puntos?

Este problema implica encontrar la función $y_o(x)$ que minimiza el funcional

$$S[y] = \int_{[x_1, y_1]}^{[x_2, y_2]} ds = \int_{x_1}^{x_2} \sqrt{1 + y'(x)^2} dx$$

entre los dos puntos en el plano (x_1, y_1) y (x_2, y_2)



El camino mas corto entre dos puntos

Sabemos intuitivamente que la solución de este problema es la linear recta entre los dos puntos. La idea es encontrar un método general para probar que este es el caso.

Ejemplo: Cual es el camino que toma menos tiempo bajo la fuerza de gravedad?

Este es un problema de gran importancia en el siglo XVII. Una masa se mueve sin fricción bajo la fuerza de gravedad desde el punto $[x_1, y_1]$ al punto $[x_2, y_2]$ por un alambre definido por la curva $y(x)$. ¿Cual es la curva que toma menos tiempo? Este problema implica encontrar la función $y_o(x)$ que minimiza el tiempo definido por el funcional

$$S[y] = \int_0^T dt = \int_{x_1, y(x_1)}^{x_2, y(x_2)} \frac{ds}{v} = \int_{x_1}^{x_2} \frac{\sqrt{1 + y'(x)^2}}{\sqrt{2g(y_1 - y(x))}} dx$$

Esta relación se obtiene gracias a la conservación de energía, ósea

$$\frac{1}{2}mv^2 + mgy = mgy_1$$

Ejemplo: Principio de Fermat

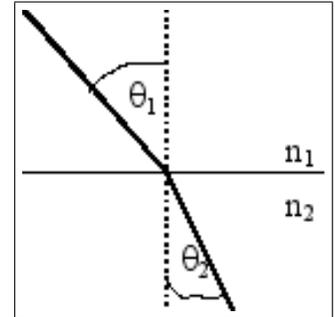
El principio de Fermat (la aproximación WKB de las ecuaciones de Maxwell's) dice que las ondas electromagnéticas, definidas por su normal o rayo, satisfacen el principio de mínimo tiempo. La velocidad de la onda se define como $v = c/n$, por lo tanto el rayo se define por el mínimo del funcional

$$S = \int_P^Q \frac{ds}{v}$$

Ejemplo: Un problema entretenido es el de la propagación en una fibra óptica que tiene un índice de refracción dado por $n(x) = n_o(1 - y(x))^2$. El rayo parte del centro de la fibra.

Problema: Use el principio de Fermat para probar la relación de óptica

$$n_1 \sin \theta_1 = n_2 \sin \theta_2$$

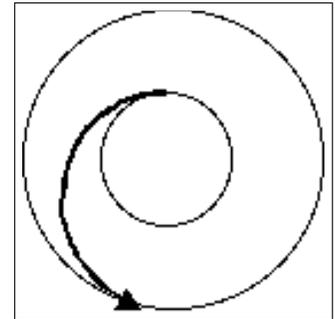


La Ley de Snell

Ejemplo: Camino que un satélite consumiría la menor cantidad de combustible bajo la fuerza de gravedad? Cual es la función $r(\theta)$ que minimiza el gasto en combustible para llegar de $r_1(\theta_1)$ a $r_2(\theta_2)$.

$$S[r(\theta)] = \int_{M_o}^{M_1} dm = \int_0^\pi \frac{dm(\theta)}{d\theta} d\theta$$

$$r(0) = r_o \quad r(\pi) = r_1$$



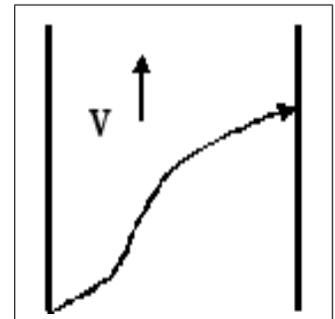
La Orbita menos costosa

Ejemplo: Tiempo mas corto para cruzar un río que fluye con una velocidad $v(x, y) = v(x)\hat{x}$?

La rapidez del bote es c . El funcional es ($c^2 > v^2$)

$$S[y] = \int_0^b \frac{dx}{c_x} = \int_0^b \frac{1 + y'(x)^2}{\sqrt{c^2(1 + y'(x)^2) - v^2 + vy'(x)}} dx$$

$$y(0) = 0 \quad y(b) \text{ no especificada}$$



Cruzar el río

Ejemplo: Cual es el camino mas corto entre dos puntos sobre la superficie de una esfera.

En este caso hay que minimizar el funcional

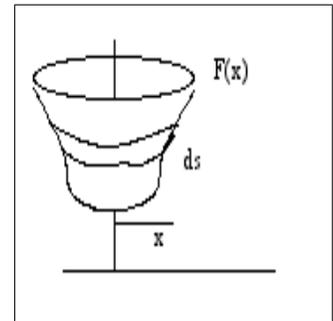
$$ds^2 = r^2(d\theta^2 + \sin^2\theta d\phi^2)$$

$$S = r \int_1^2 \sqrt{\left(\frac{d\theta}{d\phi}\right)^2 + \sin^2\theta} d\phi$$

Ejemplo: Encontrar la función $y(x)$ que minimiza la superficie de revolución.

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

$$y(a) = y_a \quad y(b) = y_b$$



Volumen de revolución

2. Calculo Funcional en una dimensión

Antes de empezar con el calculo funcional revisemos como encontramos el mínimo o máximo de una función.

2.1. Cálculo diferencial

En cálculo diferencial usualmente estamos interesados en el mínimo de una función en un intervalo (abierto). Osea, la función $f(x)$ tiene un mínimo local si existe un $\delta > 0$, tal que

$$f(x_o) \leq f(x) \quad \forall x \quad |x - x_o| < \delta$$

para funciones diferenciables (C^1), esto se traduce a

$$\frac{df(x_o)}{dx} = \lim_{\Delta x \rightarrow 0} \frac{f(x_o + \Delta x) - f(x_o)}{\Delta x} = 0 \quad (1)$$

Es importante notar que Ec. (1) es una condición necesaria, pero solo

$$\frac{df(x_o)}{dx} = 0 \quad \text{and} \quad \frac{d^2f(x_o)}{dx^2} > 0$$

garantiza un mínimo local.

Una muy buena analogía para lo que necesitamos hacer en calculo funcional se puede encontrar en calculo diferencial en R^n . La función $f : R^n \rightarrow R$ tiene un extremo en \mathbf{x} si

$$\lim_{|\mathbf{y}| \rightarrow 0} \frac{f(\mathbf{x} + \mathbf{y}) - f(\mathbf{x})}{|\mathbf{y}|} = 0$$

para todas las posibles variaciones $|\mathbf{y}| \rightarrow 0$. Esto es equivalente a la ecuación que estamos buscando pero en el espacio de las funciones A. Si la función es suficientemente suave, es interesante darse cuenta que esto significa que el extremo satisface $\nabla f = 0$.

Nosotros estamos interesados en establecer el equivalente diferencial de ∇f para extremos (máximos o mínimos) de funcionales. Es muy importante darse cuenta que se necesitan dos cosas.

1. Primero, para establecer una noción de limite (continuidad), ósea cercanía en A, necesitamos definir el concepto de espacio lineal con norma (Normed Linear Spaces).
2. Segundo, necesitamos definir el concepto de derivada en A si esta existe.

El primer punto se satisface con el concepto de norma y espacio lineal, tal que funciones como $f(x) + \varepsilon h(x) \in A$ para todo ε suficientemente pequeño. Estos detalles matemáticos son importantes, pero vamos suponer que nuestro conjunto A satisface las necesidades de diferenciabilidad que ameriten.

2.2. Derivadas Funcionales

La derivada funcional de

$$S[f] = \int_a^b \mathcal{L}(x, f, f') dx$$

con $f \in A$ (subset de un espacio lineal normado) en la dirección de $h \in A$ se define como

$$\delta S(f, h) = \lim_{\varepsilon \rightarrow 0} \frac{S(f + \varepsilon h) - S(f)}{\varepsilon} = \left. \frac{dS(f + \varepsilon h)}{d\varepsilon} \right|_{\varepsilon=0}$$

De aquí en adelante supongo que todas las condiciones de continuidad y diferenciabilidad son satisfechas por las funciones f y h. Por ejemplo, para garantizar la existencia de este limite, se necesita que el funcional sea uniformemente continuo, etc., etc., etc. Estos detalles matemáticos son importantes, pero vamos suponer que nuestro conjunto A satisface las necesidades de diferenciabilidad que ameriten. De esta forma podemos avanzar para los temas que necesitamos tocar.

Si la función $f_o \in A$ es un mínimo local de S (dada una norma), entonces

$$\delta S(f_o, h) = 0 \tag{2}$$

para todas las variaciones $h(x)$ admisibles. En general vamos a exigir que las variaciones satisfagan

$$h(a) = h(b) = 0$$

ya que en este subset tenemos suficiente flexibilidad para encontrar la función óptima.

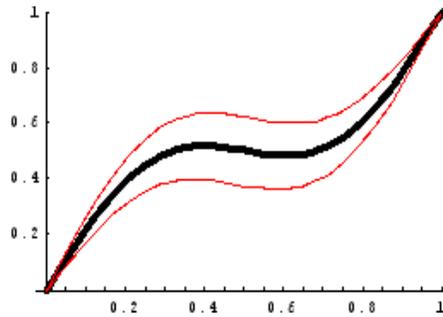


Figura 1: La función y dos variaciones.

La importancia de esta declaración, Ec. (2), es que veremos muy luego que f_o satisface una ecuación diferencial muy específica. Esta ecuación diferencial (la ecuación de Euler-Lagrange) es la base de la mecánica analítica.

Ejemplo: primer intento:

Considere el funcional

$$S(f) = \int_0^1 \sqrt{1 + f'(x)^2} dx$$

$$f(0) = 0 \quad f(1) = 1$$

Intuitivamente sabemos que la solución debería ser $f_o(x) = x$. Así que calculemos el valor del funcional para la familia $f(x, \varepsilon) = f_o(x) + \varepsilon h(x) = x + \varepsilon x(1 - x)$. Es importante notar que en este caso A son las funciones (C^1) que satisfacen las condiciones en $x = 0$ y $x = 1$, ósea $h(0) = h(1) = 0$. Obtenemos

$$S[f(x, \varepsilon)] = \int_0^1 \sqrt{1 + (f'(x, \varepsilon))^2} dx \cong \sqrt{2} + \frac{\varepsilon^2}{12\sqrt{2}} + O(\varepsilon^4)$$

como debería ser para un mínimo. Por lo tanto tenemos

$$\delta S(f, h) = 0$$

Es muy importante darse cuenta que hemos probado que la función $f_o(x) = x$ satisface $\delta S(f_o, h) = 0$ solo en la dirección $h(x) = x(1 - x)$, pero aun no hemos probado que esta función minimiza este funcional para toda función $h \in A$ con $h(0) = h(1) = 0$.

Ahora desarrollaremos un método general para encontrar la función que optimiza un funcional. De pasada probaremos que efectivamente $y(x) = x$ es la función que optimiza el funcional de este ejemplo. Por lo tanto tenemos que encontrar una ecuación para la función $f(x)$ que garantice que

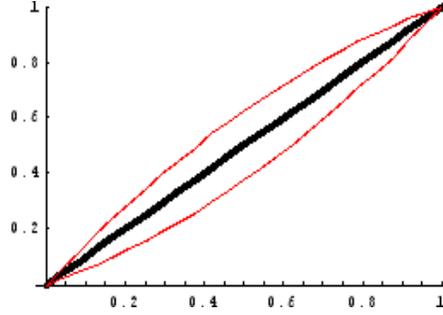


Figura 2: Línea recta y dos variaciones.

$$\delta S[f, h] = 0 \text{ con } f, h \in A$$

para todas las posibles variaciones $\forall f(x) + \varepsilon h(x) \in A$
tal que $h(0) = h(1) = 0$

2.3. La ecuación de Euler-Lagrange

Estamos interesados en optimizar funcionales del siguiente tipo

$$S[f] = \int_{x_1}^{x_2} \mathcal{L}(x, f, f') dx \quad f(x_1) = f_1 \quad f(x_2) = f_2$$

donde L es denominado el Lagrangiano. Usaremos la noción de que tanto f, h son por lo menos C^2 , tal que $f + \varepsilon h$ también es C^2 y que satisfacen las condiciones de borde $h(x_1) = h(x_2) = 0$. Expandiendo

$$S[f + \varepsilon h] - S[f] \approx \int_{x_1}^{x_2} [(\mathcal{L}_f(x, f, f')h + \mathcal{L}_{f'}(x, f, f')h')] \varepsilon dx$$

podemos encontrar que

$$\lim_{\varepsilon \rightarrow 0} \left(\frac{S[f + \varepsilon h] - S[f]}{\varepsilon} \right) = \int_{x_1}^{x_2} (\mathcal{L}_f(x, f, f')h + \mathcal{L}_{f'}(x, f, f')h') dx = 0$$

Integramos por partes para deshacernos de h' , utilizando $(\Phi\Psi)' = \Phi'\Psi + \Phi\Psi'$, podemos obtener

$$\int_{x_1}^{x_2} \left[\mathcal{L}_f(x, f, f') - \frac{d}{dx} \mathcal{L}_{f'}(x, f, f') \right] h dx + \mathcal{L}_{f'}(x, f, f')h \Big|_{x=x_1}^{x=x_2} = 0$$

Es importante darse cuenta que esta última derivada es completa, no es una derivada parcial como las otras. Dado que $h(x_1) = h(x_2) = 0$, tenemos que

$$\int_{x_1}^{x_2} \left[\mathcal{L}_f(x, f, f') - \frac{d}{dx} \mathcal{L}_{f'}(x, f, f') \right] h dx = 0$$

Esta relación debe de ser satisfecha para todo $h \in C^2$ que cumpla con $h(x_1) = h(x_2) = 0$. Notemos que en particular si tenemos que

$$\int_{x_1}^{x_2} g(x)h(x) dx = 0$$

para todo $h \in C^2$ con $h(x_1) = h(x_2) = 0$, es fácil demostrar que

$$g(x) = 0$$

por contradicción. Supongamos que $g(x) > 0$ en el intervalo (x_1, x_2) , entonces podemos encontrar una función h que tiene soporte positivo en (x_1, x_2) (por ejemplo $h = (x_1 - x)^3(x - x_2)^3$) y que es cero fuera de este intervalo. Entonces la integral daría un valor positivo, lo cual no es posible. Por lo tanto $g(x) = 0$.

Usamos esta última demostración para probar el siguiente teorema. La función $f(x) \in C^2$ que optimiza el funcional

$$S[f] = \int_{x_1}^{x_2} \mathcal{L}(x, f, f') dx$$

con las condiciones de borde $f(x_1) = f_1$ y $f(x_2) = f_2$ satisface la ecuación de Euler-Lagrange

$$\boxed{\frac{d}{dx} \mathcal{L}_{f'}(x, f, f') - \mathcal{L}_f(x, f, f') = 0}$$

para $x \in [x_1, x_2]$. Se dice que la función $f(x)$ que satisface la ecuación de Euler-Lagrange, es estacionaria para el funcional S

Solución Ejemplo

Usemos la ecuación de Euler-Lagrange para el Lagrangiano

$$\mathcal{L} = \sqrt{1 + y'(x)^2} \quad y(x_1) = y_1 \quad y(x_2) = y_2$$

con las expresiones

$$\begin{aligned} \mathcal{L}_y(x, y, y') &= 0 \\ \mathcal{L}_{y'}(x, y, y') &= \frac{2y'(x)}{\sqrt{1 + y'(x)^2}} \end{aligned}$$

Podemos ahora mostrar que la función $y(x)$ satisface la ecuación de Euler-Lagrange

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

Esto implica que

$$\frac{2y'(x)}{\sqrt{1+y'(x)^2}} = \text{const} \quad \rightarrow \quad y'(x) = C_2$$

De esta forma concluimos que la distancia mas corta entre los dos puntos $[x_1, y_1]$ y $[x_2, y_2]$ es la línea recta dada por la ecuación

$$y(x) = y_1 + (x - x_1) \frac{(y_2 - y_1)}{(x_2 - x_1)}$$

2.3.1. Constantes de movimiento

De la solución de este problema, aparece naturalmente el concepto de las integrales o constantes del movimiento. Las ecuaciones de Euler Lagrange son de segundo orden en general, lo que implica que hay dos constantes de integración, las cuales están determinadas por las dos condiciones de borde en $x = x_1, x_2$.

Por ejemplo,

$\mathcal{L}(x, f) \rightarrow \mathcal{L}_f(x, f) = 0$
$\mathcal{L}(x, f') \rightarrow \mathcal{L}_{f'}(x, f') = \text{const}$
$\mathcal{L}(f, f') \rightarrow \mathcal{L}(f, f') - f' \mathcal{L}_{f'}(f, f') = \text{const}$

2.3.2. Existencia de una solución

Sabemos que existe una solución para problemas en que L depende en forma cuadrática en f' , como será en mecánica.

3. Generalizaciones

En esta sección nos dedicaremos a la generalización de las ecuaciones de Euler-Lagrange

3.1. Caso 1: Derivadas de alto orden

Definamos el funcional $S : A \rightarrow R^1$, con $A \in C^{2n}$

$$S[f] = \int_{x_1}^{x_2} \mathcal{L}(x, f, f', f'', \dots, f^{(n)}) dx$$

$$f^{(i)}(x_1) = f_1^{(i)} \quad f^{(i)}(x_2) = f_2^{(i)} \quad i = 0, \dots, n - 1$$

Es fácil demostrar que la ecuación de Euler-Lagrange es

$$\mathcal{L}_f(x, f, f') + \sum_{m=1}^n (-1)^m \frac{d^m}{dx^m} \mathcal{L}_{f^{(m)}} = 0$$

para $x \in [x_1, x_2]$.

3.2. Caso 2: Varias funciones

Definamos el funcional $S : A \rightarrow \mathbb{R}^1$, con $A \in C^2$

$$S[f_1, f_2, \dots] = \int_{x_1}^{x_2} \mathcal{L}(x, f_1, \dots, f_n, f'_1, \dots, f'_n) dx$$

$$f_i(x_1) = f_{i,1} \quad f_i(x_2) = f_{i,2} \quad i = 1, \dots, n$$

Para esta prueba se usa las variaciones $f_i(x) + \varepsilon h_i(x)$, con $h_i(x_1) = h_i(x_2) = 0$. Dado que las funciones h_i son arbitrarias e independientes, se puede probar que las funciones $f_i(x)$ que optimizan el funcional satisfacen las ecuaciones de Euler-Lagrange

$$\frac{d}{dx} \mathcal{L}_{f'_i} - \mathcal{L}_{f_i} = 0$$

para $x \in [x_1, x_2]$ y para $i = 1, \dots, n$.

3.3. Caso 3: Múltiples integrales

Definamos el funcional S sobre un volumen en $x \in \mathbb{R}^n$

$$S[f] = \int_{\Omega} dx^n \mathcal{L}[x, f, \{f_{,i}\}]$$

con la notación

$$dx^n = \prod_{i=1}^n dx_i$$

donde las condiciones de borde están dadas sobre la curva cerrada $\delta\Omega$, tal que

$$f(x)|_{\delta\Omega} = g(x)$$

Usemos la notación

$$f_{,i} = \frac{\partial f}{\partial x_i}$$

Usando la variación $f + \varepsilon h$, con

$$h(x)|_{\delta\Omega} = 0$$

tenemos

$$S[f + \epsilon h] - S[f] = \epsilon \int_{\Omega} dx \left[\mathcal{L}_f h + \sum_i^n \mathcal{L}_{f,i} h_{,i} \right]$$

Notemos que para un campo vectorial A en n dimensiones

$$\sum_i \partial_i (A_i h) = \sum_i A_i \partial_i h + h \sum_i \partial_i A_i$$

(esto es equivalente a $\nabla \cdot (Ah) = A \cdot \nabla h + h \nabla \cdot A$) y además

$$\int_{\Omega} \nabla \cdot A dx = \int_{\partial\Omega} A \cdot \hat{n} dS$$

Con lo cual ahora podemos integrar por parte para obtener

$$S[f + \epsilon h] - S[f] = \epsilon \int_{\Omega} dx \left[\mathcal{L}_f - \sum_i^n \frac{\partial}{\partial x_i} \mathcal{L}_{f,i} \right] h$$

usando las condiciones de borde para h . Por lo tanto obtenemos las ecuaciones de Euler-Lagrange

$$\boxed{\mathcal{L}_f - \sum_{i=1}^n \frac{\partial}{\partial x_i} \mathcal{L}_{f,i} = 0}$$

En el caso de varias funciones de varias variables tenemos $f_i(x)$ la ecuación de Euler-Lagrange es

$$\boxed{\mathcal{L}_{f_i} - \sum_j \frac{\partial}{\partial x_j} \mathcal{L}_{f_i,j} = 0}$$

para $i = 1, \dots, n$, con la notación

$$f_{i,j} = \frac{\partial f_i}{\partial x_j}$$

3.4. Condiciones naturales de borde

Hasta ahora hemos discutido problemas en los cuales las funciones deben de satisfacer una condición de contorno muy específica. Que pasa si la condición de borde no esta especificada, como en el ejemplo de arriba. Este caso particular se denomina, **condiciones naturales de borde**. Supongamos que no condición es especificada en $x = x_2$. La variación nuevamente esta dada por

$$\int_{x_1}^{x_2} \left[\mathcal{L}_f - \frac{d}{dx} \mathcal{L}_{f'} \right] h dx + \mathcal{L}_{f'} h \Big|_{x=x_1}^{x=x_2} = 0$$

excepto que las condiciones en x_2 no esta definida y por lo tanto el ultimo termino no es cero a priori. Pero por el otro lado, ya que $h(x_1) = 0$, tenemos

$$\int_{x_1}^{x_2} \left[\mathcal{L}_f - \frac{d}{dx} \mathcal{L}_{f'} \right] h dx + \mathcal{L}_{f'}(x_2, f(x_2), f'(x_2))h(x_2) = 0 \quad (3)$$

Es importante darse cuenta que esta relación esta definida para todas as variaciones h tal que $h(x_1) = 0$. En particular podemos elegir el subconjunto de h que satisfacen $h(x_2) = 0$,

$$\int_{x_1}^{x_2} \left[\mathcal{L}_f - \frac{d}{dx} \mathcal{L}_{f'} \right] h dx = 0$$

y por lo tanto tenemos nuevamente que f debe de satisfacer

$$\mathcal{L}_f - \frac{d}{dx} \mathcal{L}_{f'} = 0$$

pero ya que Ec. (3) debe de ser satisfecha para todo h , incluso $h(x_2) \neq 0$, tenemos que la condición de borde natural es

$$\boxed{\mathcal{L}_{f'}(x_2, f(x_2), f'(x_2)) = 0}$$

4. Teorema para transformación de coordenadas

Primero, es casi intuitivo darse cuenta que podemos sumarle a L la derivada completa de un funcional $M(x, f)$

$$\bar{\mathcal{L}} = \mathcal{L} + \frac{dM(x, f)}{dx}$$

y obtendremos las mismas ecuaciones de movimiento. Es importante también darse cuenta que esto no funciona si el funcional M depende de las velocidades. Esto se llama una **transformación de Gauge**. Es fácil probar que

$$\delta \bar{\mathcal{L}} = \delta \mathcal{L}$$

En el caso mas general supongamos una transformación uno a uno $G : x \rightarrow \bar{x}$ (tanto G como G^{-1} son diferenciables)

$$\bar{x} = G(x)$$

Asumamos que el Lagrangiano es un escalar, lo que implica

$$\mathcal{L} [x, f(x), f'(x)] = \mathcal{L} [G^{-1}(\bar{x}), \bar{f}(\bar{x}), \bar{f}'(\bar{x})\alpha(\bar{x})] = \bar{\mathcal{L}} [\bar{x}, \bar{f}(\bar{x}), \bar{f}'(\bar{x})]$$

con la definición

$$\bar{f}(\bar{x}) = f(x).$$

Dado que

$$\frac{df(x)}{dx} = \frac{d\bar{f}(\bar{x})}{d\bar{x}} \frac{d\bar{x}}{dx} = \frac{d\bar{f}(\bar{x})}{d\bar{x}} \alpha(\bar{x})$$

podemos escribir

$$\alpha(\bar{x}) = \left. \frac{dG(x)}{dx} \right|_{x=G^{-1}(\bar{x})}.$$

Por lo tanto si tenemos

$$\frac{\partial}{\partial x} \mathcal{L}_{f'} - \mathcal{L}_f = 0$$

entonces podemos calcular

$$\begin{aligned} \frac{d}{d\bar{x}} [\bar{\mathcal{L}}_{\bar{f}'}] - \bar{\mathcal{L}}_{\bar{f}} &= \frac{dG^{-1}}{d\bar{x}} \frac{d}{dx} [\alpha(\bar{x}) \mathcal{L}_{f'}] - \mathcal{L}_f \\ &= \left[\frac{dG^{-1}}{d\bar{x}} \alpha(\bar{x}) \right] \frac{d}{dx} \mathcal{L}_{f'} - \mathcal{L}_f \\ &= \left[\frac{dx}{d\bar{x}} \frac{d\bar{x}}{dx} \right] \frac{d}{dx} \mathcal{L}_{f'} - \mathcal{L}_f \\ &= \frac{d}{dx} \mathcal{L}_{f'} - \mathcal{L}_f = 0 \end{aligned}$$

usando la regla de la cadena sobre $G(G^{-1}(\bar{x})) = \bar{x}$. Vemos que $\bar{f}(\bar{x})$ es consistente con la trayectoria $f(x)$ y por lo tanto una transformación de coordenadas no afecta la solución de problemas, esto es decir que el problema es invariante bajo una transformación de coordenadas uno a uno. Mas importante aun, en los dos sistemas las ecuaciones de Euler-Lagrange describen las trayectorias del sistema. Este es probablemente el teorema mas importante de este capitulo.

Notemos que este análisis se generaliza rápidamente a a varias funciones de varias variables. Veremos mas adelante que esto es fundamental cuando utilicemos el **principio de Hamilton** para reformular la dinámica Newtoniana.

4.1. Restricciones

En general las variables en las cuales estamos interesados tienen restricciones. Podemos evadir este problema encontrando “variables generalizadas” que incluyen las restricciones, por ejemplo haciendo una transformada

de (x, y, z) a (θ, ϕ) sobre la superficie de una esfera. El análisis anterior nos garantiza que las soluciones en los dos sistemas de coordenadas son consistentes.

En esta sección veremos como optimizar funcionales con restricciones. Definimos las restricciones Holonomicas (Holonomic constraints) como las cuales pueden ser descritas por un conjunto de m de ecuaciones independiente dados por

$$g_j(x, f_1, \dots, f_n) = C_j \quad j = 1, \dots, m$$

No podemos aplicar el método de Euler-Lagrange a las variables dependientes, $f_i \quad i = 1, \dots, n$, ya que las variaciones h_i no son independientes. Supongamos primero que somos capaces de escribir las variables $f_i(x, q_1, \dots, q_{n-m})$ en función de variables generalizadas independientes $q_i \quad i = 1, \dots, n - m$, entonces podemos aplicar el método de Euler-Lagrange al Lagrangiano

$$\mathcal{L}(x, f_1, \dots, f_n) = \mathcal{L}(x, q_1, \dots, q_{n-m})$$

ya que las variaciones $h_i^{(q)} \quad i = 1, \dots, n - m$ son independientes. Por ejemplo, en la superficie de una esfera tenemos 3 variables dependientes, (x, y, z) que tienen la restricción $x^2 + y^2 + z^2 = r^2$ por lo cual uno utiliza dos variables independientes (θ, ϕ) , denominadas “variables generalizadas” y fuimos capaces de re-escribir L en termino de variables independientes a las cuales les podemos aplicar el formalismo de Euler-Lagrange

Supongamos que no podemos o no queremos re-escribir el Lagrangiano en termino de variables generalizadas. Nos damos cuenta de inmediato que el conjunto de variaciones $h_i(x)$ no es arbitrario y se deben regir por las restricciones.

Problema: Que pasa cuando uno trata el problema de minimizar una función $L(x, f_1, f_2, f'_1, f'_2)$ sujeta a una condición extra $g(x, f_1, f_2) = const$. Intuitivamente usamos la técnica de Multiplicadores de Lagrange. Minimizamos el siguiente funcional

$$\bar{\mathcal{L}} = \mathcal{L} - \lambda(g - C)$$

lo que nos da

$$\frac{d}{dx} \mathcal{L}_{f'_i} - \mathcal{L}_{f_i} = \lambda \frac{\partial g}{\partial f_i}$$

para $i = 1, 2$, además de la restricción. Tenemos 3 ecuaciones y tres variables que resolver, f_1, f_2, λ .

Miremos esto desde el punto de vista variacional evaluando las variaciones del funcional

$$S[f_1 + \epsilon h_1, f_2 + \epsilon h_2] - S[f_1, f_2] = \epsilon \int dx \sum_i h_i \left[\mathcal{L}_{f_i} - \frac{d}{dx} \mathcal{L}_{f'_i} \right]$$

Aquí no podemos forzar los paréntesis a cero porque las variaciones de h_i no son independientes pero las podemos relacionar como

$$0 = dg = \epsilon \sum_i \frac{\partial g}{\partial f_i} h_i$$

Por lo tanto resolviendo por h_2 tenemos

$$h_2 = -\frac{1}{\frac{\partial g}{\partial f_2}} \frac{\partial g}{\partial f_1} h_1 = -\frac{1}{g_{f_2}} g_{f_1} h_1$$

y por lo tanto

$$S[f_1 + \epsilon h_1, f_2 + \epsilon h_2] - S[f_1, f_2] = \epsilon \int dx h_1 \left[\mathcal{L}_{f_1} - \frac{d}{dx} \mathcal{L}_{f_1'} - \frac{1}{g_{f_2}} g_{f_1} \left(\mathcal{L}_{f_2} - \frac{d}{dx} \mathcal{L}_{f_2'} \right) \right]$$

y la ecuación de Euler-Lagrange es

$$\mathcal{L}_{f_1} - \frac{d}{dx} \mathcal{L}_{f_1'} - \frac{1}{g_{f_2}} g_{f_1} \left(\mathcal{L}_{f_2} - \frac{d}{dx} \mathcal{L}_{f_2'} \right) = 0$$

ya que h_1 si se puede elegir en forma independiente. Esto es lo mismo que obtenemos por el método de Multiplicadores de Lagrange observando que

$$\frac{1}{g_{f_1}} \left(\mathcal{L}_{f_1} - \frac{d}{dx} \mathcal{L}_{f_1'} \right) = const = \frac{1}{g_{f_2}} \left(\mathcal{L}_{f_2} - \frac{d}{dx} \mathcal{L}_{f_2'} \right)$$

Notemos que es relativamente fácil generalizar esto a mas funciones y mas dimensiones.

Problema: Que pasa si tenemos una restricción global? Por ejemplo, encontremos f que optimiza

$$S[f] = \int_{x_1}^{x_2} \mathcal{L}(x, f, f') dx \quad f(x_1) = f_1 \quad f(x_2) = f_2$$

sujeto a la condición

$$W(f) = \int_{x_1}^{x_2} g(x, f, f') dx = C(x_2 - x_1)$$

De nuevo, asumimos que estas funciones admiten todas las condiciones de continuidad y diferenciabilidad (ósea C^2) requeridas. Es fácil de probar que este problema es equivalente a minimizar el siguiente funcional

$$S[f] = \int_{x_1}^{x_2} [\mathcal{L}(x, f, f') + \lambda (C - g(x, f, f'))] dx$$

La función $f(x)$ que minimiza esta integral esta dada por la ecuación de Euler-Lagrange

$$\begin{aligned}\mathcal{L}' &= \mathcal{L} - \lambda g \\ \mathcal{L}'_f - \frac{d}{dx} \mathcal{L}'_{f'} &= 0\end{aligned}$$

Esta ecuación es segundo orden, por lo tanto requiere 2 constantes mas λ . Estos valores se pueden calcular de las dos condiciones de borde y de la función de restricción. Es muy interesante darse cuenta que la ecuación de restricción se puede derivar directamente de la Ec. (3) haciendo una variación en λ . Esto último parece sugerir que λ tiene un valor físico.

Ejemplo Encontrar la función que optimiza el funcional

$$S[\psi] = \int \left[\frac{\hbar^2}{2m} (\nabla\psi(x))^2 + V(x)\psi(x)^2 \right] dx$$

Bajo las condiciones

$$\int \psi^2 dx = 1$$

Este problema implica que necesitamos minimizar el funcional

$$L = \frac{\hbar^2}{2m} (\nabla\psi)^2 + V\psi^2 - \lambda\psi^2$$

Del cual obtenemos

$$-\frac{\hbar^2}{2m} \nabla^2\psi + V\psi = \lambda\psi$$

Este es el problema característico de mecánica cuántica. Es muy interesante darse cuenta que los problemas de la física se pueden formular como un problema de optimización. Esto es fundamental para dar el paso a la física moderna. También es relevante notar que λ adquiere un significado físico. Veremos mas de esto cuando re-formulemos la mecánica en el próximo capítulo.

Para el caso general de n variables $f(x, y_1, \dots, y_n)$ con m condiciones de restricción $g_j(x, y_1, \dots, y_n) = C_j$ tenemos

$$\begin{aligned}\mathcal{L}' &= \mathcal{L}(x, y_1, \dots, y_n) + \lambda_j(x)(g_j(x, y_1, \dots, y_n) - C_j) & i &= 1, \dots, n \\ g_j(x, y_1, \dots, y_n) - C_j &= 0 & j &= 1, \dots, m\end{aligned}\tag{4}$$

5. Solución Numérica

En general muchos de estos problemas quedan explícitos como problemas con condiciones de borde, no como un problema inicial que se puede integrar con métodos estándares.

5.1. Shooting method

Para esto debemos formular un “shooting method”.

Problema: Supongamos que tenemos que resolver

$$\frac{d^2\Psi}{dx^2} = -4\pi\rho(x)$$

que en principio se puede encontrar como

$$\Psi(x) = \Psi(0) + \Psi'(0)x - 4\pi \int_0^x dz \int_0^z \rho(y)dy$$

Podemos notar que para integrar esta ecuación necesitamos ser capaces de hacer la integral y además tener las condiciones de borde en $x = 0$.

Muchas veces la integral no se puede resolver en forma analítica y en otras no tenemos las condiciones en el mismo punto, sino mas bien tenemos condiciones de borde $\Psi(0)$ y $\Psi(L)$. Para estos casos, podemos utilizar un “método de disparo” para encontrar la solución en forma numérica.

Por ejemplo, si queremos resolver el problema de

$$\rho(x) = \sin(\cos[x]) \quad \Psi(0) = 1 \quad \Psi(1) = 4$$

Para eso definimos una función que encuentre $\Psi(1)$ numericamente a partir de un valor inicial $\Psi'(0)$. Ahora tenemos que encontrar el cero de la función

$$g(x) = fun[x] - \Psi(1)$$

Para encontrar el valor de $\Psi'(0)$ que de $\Psi(1) = 4$. Esto se puede hacer con el método de la secante dados dos valores iniciales $s_0 = 1$, $s_1 = 3$,

$$s_{n+1} = s_n - g(s_n) \frac{s_n - s_{n-1}}{g(s_n) - g(s_{n-1})}$$

el cual converge rápidamente a $s \rightarrow 7,2$, y la solución se muestra en la Fig. 3.

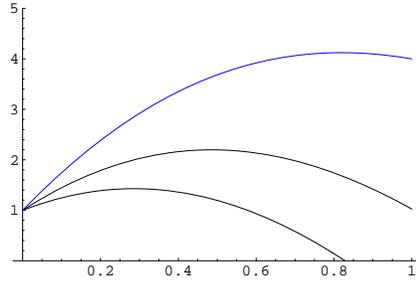


Figura 3: Solución al problema en una dimensión

5.2. Elementos finitos en 1D

También es posible construir directamente una aproximación numérica del funcional

$$S[f] = \int \mathcal{L} dx$$

discretizando la función $f(x)$ en termino de funciones localizadas

$$f(x) = \sum_i f_i h_i(x)$$

donde $h_i(x)$ es el elemento finito alrededor del punto de la grilla (x_i). Por ejemplo,

$$h_i(x) = \begin{cases} 1 - \frac{|x - x_i|}{\Delta x} & |x - x_i| \leq \Delta x \\ 0 & |x - x_i| > \Delta x \end{cases}$$

con $x_i = i\Delta x$. Pero hay otras posibilidades como una cuadrática centrada en x_i o una gaussiana centrada en x_i . Podemos calcular la integral de arriba en función de f_i . Luego de optimizar esta expresión sobre los valores desconocidos de f_i , obtenemos la solución esperada. En MATHEMATICA podemos utilizar NSolve si el problema es polinomial, o FindMinimum si el problema es no-lineal en f_i . Para discretizaciones mas avanzadas, es posible que tengamos que usar un método mas global como algoritmos genéticos o CSA, ya que pueden existir múltiples mínimos.

Ejemplo: Supongamos que queremos resolver el mismo problema de arriba

$$\frac{d^2\Psi}{dx^2} = -4\pi\rho(x)$$

con

$$\rho(x) = \sin(\cos[x]) \quad \Psi(0) = 1 \quad \Psi(1) = 4$$

En este caso sabemos que este problema se puede escribir como un funcional

$$\mathcal{L} = \frac{1}{2}\Psi'^2 - 4\pi\rho\Psi$$

Por lo tanto re-escribmos la acción como

$$S[\{\Psi_i\}] = \frac{1}{2} \sum_{i,j=0}^N \Psi_i \Psi_j \int_0^1 dx h'_i(x) h'_j(x) - 4\pi \sum_{i=0}^N \Psi_i \int_0^1 dx h'_i(x) \rho(x)$$

Con la definición

$$A_{i,j} = \int_0^1 dx h'_i(x) h'_j(x)$$

y

$$b_i = 4\pi \int_0^1 dx h'_i(x) \rho(x)$$

donde $i = 0, \dots, N$ y $j = 0, \dots, N$. Entonces se puede escribir como una matriz actuando sobre el vector $\bar{\Psi} = (\Psi_0, \dots, \Psi_N)$,

$$S[\bar{\Psi}] = \frac{1}{2} \bar{\Psi}^T A \bar{\Psi} - \bar{\Psi}^T b$$

Notemos que las condiciones de borde tienen que ser satisfechas por

$$\Psi(0) = \sum_i \Psi_i h_i(0) \quad \Psi(1) = \sum_i \Psi_i h_i(1)$$

que se usan para despejar dos Ψ_i . Para nuestros elementos finitos triangulares tenemos

$$\Psi_0 = 1 \quad \Psi_N = 4$$

Con estos elementos finitos podemos calcular inmediatamente

$$A = \frac{1}{\Delta x} \begin{bmatrix} 1 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & & -1 & 1 \end{bmatrix}$$

por lo tanto

$$S = \frac{1}{2} \sum_{n,m=0}^N \Psi_n A_{n,m} \Psi_m - \sum_{n=0}^N b_n \Psi_n$$

Si ahora buscamos el óptimo

$$\frac{dS}{d\Psi_i} = 0 \quad i = 1, \dots, N-1$$

obtenemos

$$\sum_{n=0}^N A_{i,n} \Psi_n - b_i = 0 \quad i = 1, \dots, N - 1$$

porque la matriz A es simétrica. Notemos que las condiciones de borde están incluidas aquí ya que la sumatoria va de $n = 0$ hasta $m = N$. Separemos los términos correspondientes a condiciones de borde

$$\sum_{n=1}^{N-1} A_{i,n} \Psi_n = b_i - A_{i,0} \Psi_0 - A_{i,N} \Psi_N \quad i = 1, \dots, N - 1$$

Podemos re-escribir entonces esto como

$$\bar{A} \bar{\Psi} = \bar{b}$$

y la solución es entonces

$$\bar{\Psi} = \bar{A}^{-1} \bar{b}$$

Problemas mas complicados también se pueden enfrentar.

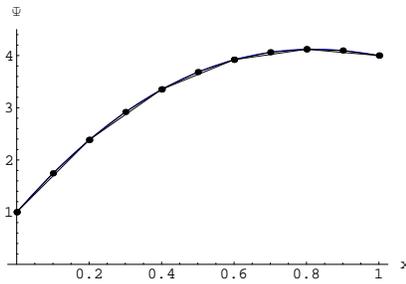


Figura 4: Solución al problema en una dimensión con elementos finitos

5.3. Elementos finitos en 2D

En realidad el método de elementos finitos tiene que ver con un procedimiento mas general para resolver problemas de este tipo. Notemos que la ecuación de Laplace viene de un principio de optimización

$$L[\Phi] = \int ((\nabla\Phi)^2 - 4\pi\rho\Phi) dx^3 \quad \rightarrow \quad \delta L = 0 \rightarrow \nabla^2\Phi = -4\pi\rho$$

Por lo tanto uno puede asumir una forma de interpolación

$$\Phi(x, y) = \sum_{i,j} \Phi_{i,j} h(x, y, x_i, y_j)$$

donde $h(x, y, x_i, y_i)$ es el elemento finito alrededor del punto de la grilla (x_i, y_j) . Usando una aproximación lineal

$$h(x, y, x_i, y_i) = h(x, x_i)h(y, y_i)$$

donde

$$h(x, x_i) = \begin{cases} 1 - \frac{|x-x_i|}{\Delta x} & |x - x_i| \leq \Delta x \\ 0 & |x - x_i| > \Delta x \end{cases}$$

con $x_i = i\Delta x$, se puede calcular la integral de arriba. Luego optimizamos esta expresión sobre los valores desconocidos de $\Phi_{i,j}$. Aquí lo importante es ordenar apropiadamente los índices para poder escribir este problema como

$$\bar{A}\bar{\Phi} = \bar{b}$$

y la solución es entonces

$$\bar{\Phi} = \bar{A}^{-1}\bar{b}$$

donde \bar{b} contiene información sobre las condiciones de borde, y \bar{A} es una matriz conocida.

Si tenemos condiciones de borde donde otro sistema de coordenadas es razonable, tenemos que evaluar el funcional en forma apropiada.

Cuando las condiciones de borde no son regulares o rectangulares, es conveniente usar una grilla no rectangular. Para estos casos es muy común utilizar una grilla triangular (la mayoría de las veces se utiliza una grilla variable con en la Fig. 5

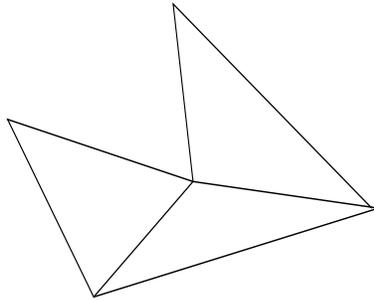


Figura 5: Grilla triangular variable

Para estos casos, conviene utilizar una interpolación lineal dentro de cada triángulo

$$\phi(x, y) \approx a_0 + a_1x + a_2y$$

donde los coeficientes van variando en cada triángulo y dependen de la posición de los tres puntos que definen cada triángulo. Es posible hacer una transformación de una interpolación dentro de los triángulos a una descripción en termino de elementos finitos sobre cada punto de la grilla, pero estos claramente dependen de la posición de los otros puntos del triángulo. El caso de una grilla regular es mas fácil.