INSTITUTO DE INGENIERÍA UNAM

MÉTODOS VARIACIONALES PARA APLICACIONES EN INGENIERÍA Y FÍSICA

PARTE 1. FORMULACIÓN GENERAL DE LOS MÉTODOS VARIACIONALES

ISMAEL HERRERA

FEBRERO 1975 349

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

ABSTRACT

This is the first part of a monograph in which a general formulation of variational principles is presented and applied to different fields of Engineering and Physics. In this part the general formulation is developed incorporating in an unified manner variational extremum and dual principles, which have received much attention in research carried out in recent years. This kind of principles have been used extensively in obtaining approximate solutions to many problems, incorporated as a basic ingredient of methods such as the finite element method.

PROLOGO

Durante los últimos años, los métodos variacionales han sido importantes en el progreso de las técnicas numéricas. En particular, la formulación del método de elementos finitos (refs 1 y 2) los ha utilizado ampliamente.

En México también se han utilizado bastante. Así, en el Instituto de Ingeniería gran número de problemas de estructuras, mecánica de suelos e hidráulica, se han tratado empleando esa clase de resultados. Simultáneamente, el Centro de Investigación en Matemáticas Aplicadas y Sistemas, ha realizado estudios tendientes a aclarar su fundamentación y a simplificar su empleo. Ambas dependencias universitarias acordaron llevar a cabo un trabajo en forma conjunta, cuya finalidad ha sido difundir en nuestro país los resultados obtenidos en el ámbito internacional en esta materia, complementándolos con investigaciones que permitieran ampliar su aplicabilidad y simplificar su formulación. El presente trabajo es el resultado de este esfuerzo, que in tenta mostrar algunos de los progresos en métodos variacionales, en forma accesible para quienes los utilizan en México.

El cálculo de variaciones abarca tópicos de carácter muy diverso, por lo que se ha preferido no tener ninguna pretensión de exhaustividad. El tema central es la metodología para formular problemas de interés en ingeniería y en física, en términos variacionales; es decir, se considera con mayor amplitud la forma en que es posible obtener para cada problema un principio variacional equivalente y las diversas clases que existen de estos.

En esta primera parte del trabajo, se establece la formulación general en que se basa el resto del mismo. En las partes subsecuentes se resumirán los principios variacionales que son aplicables a diversas ramas de ingeniería y física.

1. INTRODUCCION

En esta primera parte se presenta una formulación general de principios variacionales y los temas que se abordan, ilustrándolos con ejemplos en espacios vectoriales de dimensión finita.

Se consideran problemas definidos en un espacio vectorial. Cuando el espacio es de dimensión finita, el problema que se trata corresponde a la determinación de un punto en el cual un campo vectorial alcanza un valor fijo dado. Un principio variacional es una afirmación que establece que la variación o gradiente de una función de valores reales se anula en un punto, si y solo si el mismo es solución del problema.

Los fundamentos teóricos de los métodos variacionales se encuentran en el cálculo diferencial en espacios de Banach (ref 3), y con mayor precisión, en los resultados de esa teoría relativos a operadores potenciales. Las ideas básicas son una extensión de las correspondientes a espacios de dimensión finita;

una condición suficiente para que una ecuación admita una formulación variacional es que el operador (es decir, el campo vectorial) sea potencial, y este es el caso si y solo si las derivadas del campo forman una matriz simétrica. Cuando el campo vectorial no es derivable de un potencial se puede siempre, cuando menos en principio, transformar el problema en otro equivalente para el cual el campo vectorial u operador sí es potencial. Esta observación ha sido empleada con éxito para tratar algunos problemas de condiciones iniciales (refs 4 y 5).

Un principio extremal es el que establece la equivalencia entre una ecuación y el hecho de que alguna función alcance un valor extremo, es decir, un máximo o un mínimo. Cuando la función es diferenciable, una condición necesaria para la existencia de un máximo o un mínimo es que su gradiente o variación se anule. Esta condición no basta para la existencia de un máximo o mínimo; sin embargo, si la matriz de las segundas derivadas es positiva, es suficiente para garantizar un mínimo, e inversamente, si esa matriz es negativa, la condición es suficiente para un máximo.

A funciones que tienen la propiedad de que en todo punto la matriz de sus segundas derivadas es positiva, se les llama convexas, y aquellas para las cuales en todo punto ella es negativa se les llama cóncavas. Por tanto, para funciones convexas, todo principio variacional es un principio de mínimo y para funciones cóncavas todo principio variacional es un principio de máximo.

Es interesante particularizar los resultados anteriores al caso en que el operador o campo vectorial es lineal. Entonces la condición de potencialidad es equivalente a la simetría del operador. Además, si el operador es positivo, el potencial del campo vectorial es convexo, y si es negativo, el potencial es cóncavo. Como consecuencia, para operadores positivos se construyen principios de mínimo y para operadores negativos se construyen principios de máximo.

Existe gran número de operadores lineales simétricos de interés que no son positivos ni negativos, por lo que es deseable construir principios variacionales aplicables a ellos. Para el caso en que el espacio D es de dimensión finita, el operador se caracteriza por una matriz simétrica que tiene una base de vectores propios. Sean D_+ el subespacio generado por los vectores propios correspondientes a valores propios no negativos y D_- el subespacio generado por los vectores propios correspondientes a valores propios negativos. Entonces D_+, D_- constituyen una descomposición de D con la propiedad de que el operador es no negativo en D_+, mientras que es negativo en D_-. La condición de que el gradiente de una función se anule en D, equivale a que su componente en D_+ sea cero y su componente en D_- también lo sea. Así, la condición de que la variación se anule, se descompone en dos ecuaciones que pueden emplearse para definir una clase de principios variacionales que se acostumbra llamar duales. Ellos establecen que la función alcanza un valor extremo entre los puntos que satisfacen una de las ecuaciones solo cuando se satisface la otra.

Esta clase de principios se puede generalizar a operadores no lineales, introduciendo la noción de función *silla* y es fácil derivar de ellos los principios de Hamilton generalizados (refs 6 y 11), que tienen gran número de aplicaciones, lo mismo que los de Lagrange (ref 11), que se pueden derivar de ellos empleando transformadas de Legendre.

Aunque los principios de Hamilton y los de Lagrange son bien conocidos (ref 11) la clase más amplia que aquí se presenta es un resultado derivado en el curso de una investigación realizada en la UNAM (ref 12). La formulación que se utiliza se originó en dicho proyecto de investigación y tiene las siguientes características:

a) Se desarrolla en términos de operadores definidos en espacios vectoriales, en los que no es necesario definir un producto interior o una norma. Por ello, los resultados son de mayor aplicabilidad que los obtenidos en otras formulaciones como la presentada por Vainberg (ref 13). Esto la hace adecuada para aplicaciones en mecánica y con mayor generalidad en sistemas de ecuaciones diferenciales ordinarias o parciales.

b) El esquema general es sencillo y constituye una unidad bien definida, debido a la introducción de una clase de principios duales recientemente descubierta+ (ref 12).

c) La teoría se basa en hechos matemáticos sencillos. Es de esperarse que esto permitirá que resulte accesible a un gran número de personas interesadas en sus aplicaciones.

Finalmente, conviene mencionar que la terminología y notación para las diferentes clases de derivadas de funcionales y operadores cambia mucho con los diferentes autores. En este trabajo se utiliza ampliamente la presentada por Nashed (ref 3) en su revisión del tema.

\+ Cuando este trabajo estaba en prensa, el autor adquirió conciencia de que esta clase de principios duales fue formulada originalmente por Sewell, en la publicación:
M. J. Sewell, *The Governing Equations and Extremum Principles of Elasticity and Plasticity Generated from a Single Functional. Parte I*, J. Struct. Mech., Vol 2 (1973), pp 1-32
El propio Sewell hizo algunas aplicaciones de estos principios en:
M. J. Sewell, *The Governing Equations and Extremum Principles of Elasticity and Plasticity Generated from a Single Functional. Parte II*, J. Struct. Mech., Vol 2 (1973), pp 1-35-158

2. NOCIONES PRELIMINARES

En este capítulo se introducirán las nociones matemáticas que servirán de base para el desarrollo de la teoría general de los principios variacionales.

2.1 *Espacios vectoriales*

Se consideran exclusivamente espacios vectoriales en los que los escalares son números reales. Por tanto, la situación más generalizada en que los coeficientes son elementos de un campo algebraico arbitrario no se discute aquí, dado que carece de interés para los propósitos del presente estudio.

2.1.1 Definición

Un espacio vectorial D es un conjunto de elementos en que se han definido dos operaciones: de suma, que asocia con cada pareja de vectores $x, y \varepsilon D$ un vector $x + y \varepsilon D$ y la de multiplicación por un escalar que asocia con cada vector $x \varepsilon D$ y con cada escalar a un vector ax. El conjunto D y las operaciones satisfacen los siguientes postulados:

1. $x + y = y + x$ (Ley conmutativa)

2. $(x + y) + z = x + (y + z)$ (Ley asociativa)

3. Hay un vector nulo θ en D tal que $x + \theta = x$ para toda x de D

4. $a(x + y) = ax + ay$

5. $(a + b)x = ax + bx$

(Leyes distributivas)

6. $(ab)x = a(bx)$ (Ley asociativa)

7. $0x = \theta$, $1x = x$

Frecuentemente se empleará el mismo símbolo para el vector nulo θ, que para el cero de los números reales. En muchos casos este uso no da lugar a ambigüedad, ya que el contexto indica a cuál de los dos se hace referencia, pero cuando ocurra se procurará ser más explícito.

Se supone que el lector tiene alguna familiaridad con los espacios vectoriales. Sin embargo, a continuación se presentan algunas de sus propiedades más fundamentales.

En cualquier espacio vectorial:

1. $x + y = x + z$ implica $y = z$

$$\left.\begin{array}{l} 2.\quad ax = ay \text{ y } a \neq 0 \text{ implica } x = y \\ \\ 3.\quad ax = bx \text{ y } x \neq \theta \text{ implica } a = b \end{array}\right\} \quad \text{(Leyes de cancelación)}$$

$$\left.\begin{array}{l} 4.\quad (a - b)x = ax - bx \\ \\ 5.\quad a(x - y) = ax - ay \end{array}\right\} \quad \text{(Leyes distributivas)}$$

6. $a\theta = \theta$

Ejemplo 1. Tal vez el ejemplo más sencillo de un espacio vectorial es el conjunto de los números reales. El mismo constituye un espacio vectorial si la suma se toma en forma ordinaria y el producto por un escalar como la multiplicación usual entre reales. En tal caso, el vector nulo es el cero. Es fácil comprobar que los postulados 1 a 7 se satisfacen. A este espacio vectorial se le llama la línea real y se le denota por R^1.

Ejemplo 2. Otro ejemplo de espacio vectorial es el llamado espacio cartesiano n-dimensional de coordenadas reales, que se suele representar por R^n. Los vectores en este espacio son sucesiones (eneadas ordenadas) de n números reales, por lo que cada vector tiene la forma $x = (x_1, x_2,\dots,x_n)$. Al número real $x_k (k = 1,\dots,n)$ se le denomina el k-ésimo componente del vector x. Dos vectores son iguales si y solo si cada uno de sus componentes son iguales. El vector nulo se define como la eneada $(0, 0,\dots,0)$. Si $x = (x_1, x_2,\dots,x_n)$ y $y = (y_1, y_2,\dots,y_n)$, el vector $x + y$ se define como la eneada cuyo componente k-ésimo es $x_k + y_k$. El vector ax, donde a es un escalar, es la eneada cuyo k-ésimo componente es $a\xi_k$. Las operaciones de suma y multiplicación así definidas satisfacen los postulados de la definición de un espacio vectorial, y se puede comprobar verificando la igualdad componente por componente. Por ejemplo, $x + \theta = x$ se satisface porque $x_k + 0 = x_k$, para cada $k = 1, 2,\dots,n$.

Ejemplo 3. Sea G un conjunto cualquiera. La colección de todas las funciones de valores reales definidas en G forma un espacio vectorial. Si x y y son dos funciones, se consideran idénticas si y solo si $x(t) = y(t)$ para todos los elementos $t\varepsilon G$. El vector nulo es la función idénticamente cero. La suma y multiplicación por un escalar se definen en la forma habitual para funciones, es decir, $(x + y)(t) = x(t) + y(t)$ y $(ax)(t) = ax(t)$.

Conviene observar que el ejemplo 3 es de gran generalidad, ya que abarca gran número de situaciones de interés en aplicaciones. El conjunto G puede ser un intervalo a, b de la línea real o una región del plano, o con mayor generalidad del espacio euclideano n-dimensional E^n. Muchos problemas de medios continuos, difusión, transporte y, en general, de ecuaciones en derivadas parciales, se plantean en espacios vectoriales que son casos particulares del ejemplo 3.

2.1.2 Definición

A un subconjunto no vacío E de un espacio vectorial D se le llama subespacio de D , si E es un espacio vectorial con respecto a las mismas operaciones que D.

Ejemplo 4. Sea D la colección de todas las funciones definidas en un intervalo a, b con valores reales. De acuerdo con el ejemplo 3, D es un espacio vectorial. Sea E el conjunto de las funciones continuas definidas en a, b . En tal caso, E es un subconjunto de D. Además, se puede comprobar que E satisface los siete postulados que definen un espacio vectorial. Por lo mismo, E es un subespacio de D.

Los subespacios de un espacio vectorial se caracterizan por una propiedad sencilla. Este hecho tiene importancia práctica, ya que para determinar si un subconjunto de un espacio vectorial es subespacio, basta determinar si posee dicha propiedad, la cual se enuncia a continuación.

2.1.3 Teorema

Un subconjunto no vacío E de un espacio vectorial D, es un subespacio si y solo si $ax + by$ está en E, siempre que x y y están en E, y que a, b son reales.

Demostración. Basta comprobar que E satisface los siete postulados de espacio vectorial, si y solo si E tiene la propiedad enunciada en el teorema.

Ejemplo 5. Tome D como en el ejemplo 4 y sea E el conjunto de las funciones de D que tienen primera derivada continua. Entonces E forma un subespacio de D, porque si f y g tienen derivada continua, $af + bg$ también la tienen.

Ejemplo 6. En vez de tomar E como en el ejemplo anterior, tómese como el conjunto de las funciones de D que se anulan en el extremo inferior del intervalo $[a, b]$. Entonces E forma un subespacio de D, porque si $f(a) = 0$ y $g(a) = 0$, entonces $af(a) + bg(a) = 0$.

El ejemplo 6 es un caso particular de una situación más generalizada, aplicable a una clase de condiciones iniciales o de frontera, llamadas homogéneas, que satisfacen la propiedad del teorema 2.1.3.

Ejemplo 7. Para ilustrar esta clase de condiciones de frontera, considérese una región G del plano, limitada por una curva cerrada B. Se descompone B en dos partes ajenas, B_1 y B_2. Sea D el conjunto de todas las funciones que tienen primeras derivadas continuas en la cerradura de G y sea E el conjunto de las funciones que satisfacen:

$$u(\underset{\sim}{x}) = 0, \text{ en } B_1 \tag{2.1.1a}$$

$$\frac{\partial u}{\partial n}(\underset{\sim}{x}) = 0, \text{ en } B_2 \tag{2.1.1b}$$

Entonces D es un espacio vectorial y E un subespacio de D.

2.2 *Suma y producto cartesiano de espacios vectoriales*

Las operaciones que se mencionan a continuación desempeñan un papel importante en la formulación de principios variacionales.

2.2.1 Definición

Dados dos subespacios D_1 y D_2 de D, la suma $D_1 + D_2$ se define como el conjunto de elementos de la forma $x + y$ con $x \varepsilon D_1$ y $y \varepsilon D_2$.

Ejemplo 8. Sea D el conjunto de las funciones reales definidas en el intervalo $\left[-1, 1\right]$. Obsérvese que de acuerdo con el ejemplo 3, D es un espacio lineal. Tómese el subespacio D_1 como la colección de funciones f con la propiedad de que

$$f(t) = f(-t) \tag{2.2.1a}$$

y D_2 como la colección de funciones f, con la propiedad de que

$$f(t) = -f(-t) \tag{2.2.1b}$$

A las funciones que satisfacen la propiedad 2.2.1a se les llama simétricas y a las que satisfacen la 2.2.1b antisimétricas. La colección D_1 de funciones simétricas es efectivamente un subespacio de D_1, ya que es fácil comprobar que D_1 tiene la propiedad enunciada en el teorema 2.1.3. En forma semejante se ve que D_2 también es subespacio. En este caso, la suma $D_1 + D_2$ es el espacio total D. Para comprobarlo obsérvese que dada cualquier función u de D, se pueden definir dos funciones u_s y u_a por las ecuaciones

$$u_s(t) = \frac{1}{2}\left[u(t) + u(-t)\right] \tag{2.2.2a}$$

y

$$u_a(t) = \frac{1}{2}\left[u(t) - u(-t)\right] \tag{2.2.2b}$$

Entonces u_s es simétrica mientras que u_a es antisimétrica. Además, dada cualquier u de D.

$$u = u_s + u_a \tag{2.2.3}$$

2.2.2 Definición

Se dice que dos subespacios D_1, D_2 de D constituyen una descomposición de D, si todo elemento $x \varepsilon D$ se puede escribir como

$$x = x_1 + x_2 \tag{2.2.4}$$

con $x_1 \varepsilon D_1$ y $x_2 \varepsilon D_2$, y la representación 2.2.4 es única.

Ejemplo 9. Los subespacios D_1 y D_2 del ejemplo 8, formados por las funcio_ nes simétricas y antisimétricas, constituyen una descomposición del espacio D, de las funciones definidas en el intervalo [-1, 1]. Para probarlo basta demostrar que la representación 2.2.3 es única.

En efecto, si

$$u(t) = u_1(t) + u_2(t) \tag{2.2.5a}$$

con $u_1 \varepsilon D_1$ y $u_2 \varepsilon D_2$, entonces

$$u(-t) = u_1(t) - u_2(t) \tag{2.2.5b}$$

El sistema de ecs 2.2.5 implica que u_1 y u_2 están dadas, respectivamente, por las ecs 2.2.2a y b.

2.2.3 Definición

Dados dos espacios lineales D_1 y D_2, se define un espacio lineal $D = D_1 \times D_2$ de la manera siguiente:

i) $D = \{(x_1, x_2) \mid x_1 \varepsilon D_1, x_2 \varepsilon D_2\}$

ii) Si $a \varepsilon R^1$, $x = (x_1, x_2)$ y $y = (y_1, y_2)$

entonces

$$x + y = (x_1 + y_1, x_2 + y_2) \tag{2.2.6a}$$

$$ax = (ax_1, ax_2) \tag{2.2.6b}$$

Al espacio así definido se le llama producto cartesiano de D_1 y D_2, y se le representará por $D_1 \times D_2$. Es fácil comprobar que es un espacio lineal.

Si D_1, D_2 y D_3 son espacios lineales, se tiene:

$$(D_1 \times D_2) \times D_3 = D_1 \times (D_2 \times D_3) \tag{2.2.7}$$

Es decir, la operación producto entre espacios es asociativa. Por lo mismo, si $D_1, D_2, \ldots, D_n$ son n espacios lineales se escribirá $D_1 \times D_2 x \ldots x D_n$ para el espacio que se obtiene al efectuar el producto sucesivo de cada uno de ellos.

Considérese el espacio $D = D_1 \times D_2$. Defínase

$$\hat{D}_1 = \{(x_1, 0) \mid x_1 \varepsilon D_1\} \tag{2.2.8}$$

donde $0 \varepsilon D_2$ es el cero de dicho espacio lineal. Entonces $\hat{D}_1 \subset D$ es subespacio de D isomorfo a D_1. Por ello, D_1 se identificará con $\hat{D}_1$ y se escribirá $D_1 \subset D$. En forma semejante se sumergirá el espacio D_2 en D.

Se adoptará la notación

$$D^n = \underbrace{D x \ldots x D}_{n \text{ términos}} \tag{2.2.9}$$

2.3 *Operadores y funcionales*

2.3.1 Definición

Sean D y D' dos espacios lineales. Se llama operador a cualquier correspondencia

$$P : D \to D' \tag{2.3.1}$$

que a cada elemento $x \varepsilon D$ le asigna un elemento $P(x) \varepsilon D'$.

Ya que R^1 es un espacio lineal, esta definición es aplicable en particular si $D' = R^1$. En este caso se le llama *funcional* al operador P.

Si una función L es tal que

$$L(ax + by) = aL(x) + bL(y) \tag{2.3.2}$$

Para toda $x,y \varepsilon D$ y $a,b \varepsilon R^1$, se dice que la función operador L es lineal y sus valores se representarán por Lx. En particular, cuando L es una funcional, se hablará de una funcional lineal.

2.3.2 Definición

Sea D un espacio lineal y

$$\alpha : D^n \to R^1$$

una funcional. Si $\alpha(x_1, x_2,\ldots,x_n)$ es lineal en cada uno de sus argumentos cuando los demás se mantienen fijos, se dice que α es n-lineal en D. A esta clase de funcionales se les llama multilineales.

En tal caso, sus valores $\alpha(x_1,\ldots,x_n)$ correspondientes a cada eneada $(x_1,\ldots,x_n)$ de elementos de D serán representados por $\langle\alpha, x_1, x_2,\ldots,x_n\rangle$, es decir

$$\langle\alpha, x_1,\ldots,x_n\rangle = \alpha(x_1,\ldots,x_n) \qquad (2.3.3)$$

A la colección de las funcionales n-lineales en D, se les representará por D^{n*}. Cuando $n = 1$, se obtiene la colección de las funcionales lineales con dominio en D, la cual se denotará por D^*. Además, se define $D^{\circ*}$ por

$$D^{\circ*} = R^1 \qquad (2.3.4)$$

Si α, $\beta\varepsilon D^{n*}$ (es decir, si α,β son funcionales n-lineales con dominio en D) es común definir $\alpha + \beta$ por

$$(\alpha + \beta)(x) = \alpha(x) + \beta(x) \qquad (2.3.5)$$

para cada $x\varepsilon D^n$. Para cada real $a\varepsilon R^1$, se define la funcional n-lineal $a\alpha$ por

$$(a\alpha)(x) = a\left[\alpha(x)\right] \qquad (2.3.6)$$

para cada $x \varepsilon D$. Con respecto a estas operaciones, D^{n*} forma un espacio li_ neal. En particular a D^* se le conoce como espacio dual[+].

Ejemplo 10. Sea $D = R^n$, entonces D^* es la colección de todas las funciona_ les lineales α cuyos valores en cada $x \varepsilon R^n$ están dados por

$$<\alpha, x> = a_1 x_1 + \ldots + a_n x_n$$

donde $x_1, \ldots, x_n$ son los componentes de x, y $a_1, \ldots, a_n$ es una colección de números reales que caracteriza la funcional α. Obsérvese que en este caso el espacio dual D^* es isomorfo a R^n.

2.4 Operadores con valores en funcionales

Ya que D^{n*} forma un espacio lineal, las definiciones anteriores son aplica_ bles a operadores:

$$P : D \rightarrow D^{n*} \tag{2.4.1}$$

Operadores de este tipo, cuyos valores son funcionales, desempeñarán un papel central en la teoría de principios variacionales que aquí se presenta.

Obsérvese que cuando $n = 1$ para cada x, $y \varepsilon D$, $<P(x), y>$ es un real. De hecho $<P(x), y>$ define una funcional en D^2 que es lineal en y. Inversamente, cada funcional con dominio D^2 que es lineal en su segundo argumento, define un operador P del tipo 2.4.1 cuando los valores de la funcional se interpretan

---

+ Sin embargo, muchos autores a las funcionales del espacio dual les exigen ser además contínuas, cosa que no hacemos en este trabajo.

como los valores de $\langle P(x), y\rangle$. En particular, si $P = L$ es lineal, $\langle Lx, y\rangle$ define una funcional bilineal. Considerando esta observación, toda funcional bilineal define un operador lineal con dominio en D y con valores en D^*.

Dado un operador lineal

$$L : D \to D^* \tag{2.4.2}$$

se define el operador adjunto por la relación

$$\langle L^*x, y\rangle = \langle Ly, x\rangle \tag{2.4.3}$$

Ya que el segundo miembro de la relación 2.4.3 es una funcional bilineal en D, entonces

$$L^* : D \to D^* \tag{2.4.4}$$

Para operadores con valores en funcionales es posible introducir la noción de continuidad.

2.4.1 Definición

Sea

$$P : D \to D^{n*} \tag{2.4.5}$$

Se dice que P es bidimensionalmente continuo en $x \varepsilon D$, si para cada

$$y,\ z,\ \xi_1, \ldots, \xi_n \varepsilon D$$

la función F de las variables reales ξ, η, definidas por

$$F(\xi, \eta) = P\langle(x + \xi y + \eta z),\ \xi_1, \ldots, \xi_n\rangle \tag{2.4.6}$$

es continua en $\xi = \eta = 0$

2.5 *Diferencial o derivada de un operador*

Sea D un espacio lineal y P un operador

$$P \,.\, D \to D^{n*} \tag{2.5.1}$$

Para x, y, $\xi_1,\ldots,\xi_n \varepsilon D$ considérese la función F de variable real t, definida por

$$F(t) = \langle P(x + ty),\ \xi_1,\ldots,\xi_n\rangle \tag{2.5.2}$$

Dada $x\varepsilon D$, si la derivada $F'(0)$ existe para cada y, $\xi_1,\ldots,\xi_n \varepsilon D$, es necesariamente lineal en $\xi_1,\ldots,\xi_n$, por lo que se puede escribir

$$F'(0) = \langle VP(x,\ y),\ \xi_1,\ldots,\xi_n\rangle \tag{2.5.3}$$

donde $VP(x,\ y)\varepsilon D^{n*}$. Sin embargo, $F'(0)$ no siempre es lineal en y. Cuando $F'(0)$ es lineal en y, se puede escribir

$$F'(0) = \langle P'(x),\ y,\ \xi_1,\ldots,\xi_n\rangle \tag{2.5.4}$$

donde $P'(x)\varepsilon D^{(n+1)*}$.

2.5.1 Definición

Dada $x\varepsilon D$, si una funcional lineal $P(x)\varepsilon D^{(n+1)*}$ existe tal que con la definición 2.5.2 la ec 2.5.4 se cumple para toda y, $\xi_1,\ldots,\xi_n \varepsilon D$, entonces se dice que la derivada de P existe en x. $P'(x)$ se llama derivada o diferencial* de P.

---

* A la derivada o diferencial empleada en este trabajo se llama habitualmente (ref 3) variación aditiva de Gateaux.

Con frecuencia se usará la notación

$$\frac{dP}{dx}(x) = P'(x) \tag{2.5.5}$$

Ejemplo 11. Sea D el espacio de funciones reales definidas en el intervalo $[0,\ 1]$, con segunda derivada continua. Defínase

$$\Omega : D \to D^{\circ *} = R^1 \tag{2.5.6a}$$

por

$$\Omega(u) = -\frac{1}{2}\int_0^1 \frac{du}{dx}\frac{du}{dx}\,dx + u(1)u'(1) - u(0)u'(0) \tag{2.5.6b}$$

Por otra parte, sea el operador

$$L : D \to D^* \tag{2.5.7a}$$

definido por

$$\langle Lu, v\rangle = \int_0^1 u''(x)v(x)\,dx + u(1)v'(1) - u(0)v'(0) \tag{2.5.7b}$$

entonces

$$u\Omega'(u) = Lu \tag{2.5.8}$$

Ejemplo 12. Sea D el espacio de funciones reales definidas en el intervalo $[0,1]$, con primera derivada continua. Defínase

$$\Omega : D \to D^{\circ *} = R^1 \tag{2.5.9a}$$

por

$$\Omega(u) = \frac{1}{2}\int_0^1 u'(1-\tau)u(\tau)\,d\tau + \frac{1}{2}u(0)u(1) \tag{2.5.9b}$$

Por otra parte, sea el operador

$$L : D \to D^* \tag{2.5.10a}$$

definido por

$$\langle Lu, v\rangle = \int_0^1 u'(\tau)v(1-\tau)d\tau + u(0)v(1) \tag{2.5.10b}$$

entonces

$$\Omega'(u) = Lu \tag{2.5.11}$$

Si los espacios lineales D_1, D_2 son una descomposición de D, para cada elemento $y \varepsilon D$ existen $y_1 \varepsilon D_1$ y $y_2 \varepsilon D_2$, tales que

$$y = y_1 + y_2 \tag{2.5.12}$$

Por otra parte, si la derivada de P existe en $x \varepsilon D$, la ec 2.5.4 se cumple para toda $y \varepsilon D$. Entonces se pueden definir dos funcionales $P_{,1}(x) \varepsilon D^{(n+1)*}$ y $P_{,2}(x) \varepsilon D^{(n+1)*}$ para cada $y, \xi_1, \ldots, \xi_n \varepsilon D$ mediante la ecuación

$$\langle P_{,1}(x), y, \xi_1, \ldots, \xi_n\rangle = \langle P'(x), y_1, \xi_1, \ldots, \xi_n\rangle \tag{2.5.13a}$$

$$\langle P_{,2}(x), y, \xi_1, \ldots, \xi_n\rangle = \langle P'(x), y_2, \xi_1, \ldots, \xi_n\rangle \tag{2.5.13b}$$

2.5.2 **Definición**

Cuando la derivada del operador P dado por la ec 2.4.1 existe en $x \varepsilon D$, se dice que la derivada parcial en $x \varepsilon D$ del operador P con respecto al subespacio D_1 es $P_{,1}(x) \varepsilon D^{(n+1)*}$. En forma análoga se define la derivada parcial con respecto al subespacio D_2 como $P_{,2}(x) \varepsilon D^{(n+1)*}$.

Es pertinente hacer algunas observaciones. Ya que $P'(x)$ es lineal, las ecs 2.5.13 implican que

$$P'(x) = P_{,1}(x) + P_{,2}(x) \qquad (2.5.14)$$

Por otra parte, dado $y = y_1 + y_2$ con $y_1 \varepsilon D_1$ y $y_2 \varepsilon D_2$ y $x, \xi_1, \ldots, \xi_n \varepsilon D$, defína se

$$F_1(t) = \langle P(x + ty_1), \xi_1, \ldots, \xi_n \rangle \qquad (2.5.15a)$$

y

$$F_2(t) = \langle P(x + ty_2), \xi_1, \ldots, \xi_n \rangle \qquad (2.5.15b)$$

Si la derivada de P existe en x, entonces

$$F_1'(0) = \langle P_{,1}(x), y, \xi_1, \ldots, \xi_n \rangle \qquad (2.5.16a)$$

$$F_2'(0) = \langle P_{,2}(x), y, \xi_1, \ldots, \xi_n \rangle \qquad (2.5.16b)$$

Para derivadas parciales se empleará la notación alternativa:

$$\frac{\partial P}{\partial x_1}(x) = P_{,1}(x) \qquad (2.5.17a)$$

$$\frac{\partial P}{\partial x_2}(x) = P_{,2}(x) \qquad (2.5.17b)$$

Ejemplo 13. Tómense D y Ω como en el ejemplo 11. En forma análoga como se hizo en el ejemplo 9, defínase D_1 como las funciones de D simétricas con respecto al punto medio del intervalo $[0,1]$ y D_2 como las antisimétricas. En tal caso, D_1, D_2 constituyen una descomposición de D. Además, para cada $u \varepsilon D$ se tiene

$$u_s(t) = \frac{1}{2}\left[u(t) + u(1 - t)\right] \tag{2.5.18a}$$

$$u_a(t) = \frac{1}{2}\left[u(t) - u(1 - t)\right] \tag{2.5.18b}$$

Las derivadas parciales $\Omega,_1$ y $\Omega,_2$ están dadas por

$$<\Omega,_1(u),\ v> = \int_0^1 (u'')_s(\tau)v(\tau)d\tau + u_s(0)\left[v'(1) - v'(0)\right] \tag{2.5.19a}$$

$$<\Omega,_2(u),\ v> = \int_0^1 (u'')_a(\tau)v(\tau)d\tau - u_a(0)\left[v'(1) + v'(0)\right] \tag{2.5.19b}$$

Ejemplo 14. Tómense D y Ω como en el ejemplo 12. Defínase D_1 y D_2 como las funciones simétricas y antisimétricas de D. En tal caso las derivadas parciales $\Omega,_1$ y $\Omega,_2$ están dadas por

$$<\Omega,_1(u),\ v> = \int_0^1 (u')_s(\tau)v(\tau)d\tau + u(0)v_s(1) \tag{2.5.20a}$$

$$<\Omega,_2(u),\ v> = \int_0^1 (u')_a(\tau)v(\tau)d\tau + u(0)v_a(1) \tag{2.5.20b}$$

Para funcionales con n derivadas continuas, se puede establecer el desarrollo de Taylor con residuo.

Para funcionales con n derivadas continuas, se puede establecer el desarrollo de Taylor con residuo.

2.5.3 Teorema

Supóngase

i) $\phi : D \rightarrow R^1$

ii) ϕ tiene derivadas hasta de orden $n + 1$ en D. Además, el operador $\phi^{(n+1)} : D \rightarrow D^{(n+1)^*}$ es bidimensionalmente continuo en D.

Entonces, dados $x, y \varepsilon D$ existe un elemento ξ del segmento que une a x con y tal que

$$\phi(y) = \phi(x) + \langle\phi'(x), y - x\rangle + \ldots + \frac{1}{n!} \langle\phi^{(n)}(x), y - x, \ldots, y - x\rangle +$$

$$\frac{1}{(n+1)!} \langle\phi^{(n+1)}(\xi), y - x, \ldots, y - x\rangle \qquad (2.5.21)$$

Demostración. Defínase la función $F : [0, 1] \rightarrow R^1$ para cada $t \varepsilon [0, 1]$ por la ecuación

$$F(t) = \phi(x + t(y - x)) \qquad (2.5.22)$$

Esta función tiene derivada continua de orden $n + 1$. Luego, por el desarrollo de Taylor con residuo

$$F(1) = F(0) + F'(0) + \ldots + \frac{1}{n!} F^{(n)}(0) + \frac{1}{(n+1)!} F^{(n+1)}(\lambda) \qquad (2.5.23)$$

para alguna $\lambda \varepsilon \left[0, 1\right]$. Calculando las derivadas sucesivas se obtiene la ec 2.5.21.

2.6 *Operadores potenciales*

Se introducirá en esta sección la noción de operadores potenciales y se establecerán condiciones necesarias y suficientes para que un operador sea potencial.

2.6.1 Definición

Sea un operador

$$N : D \to D^* \qquad (2.6.1)$$

Se dice que N es potencial, si existe una funcional

$$\phi : D \to R^1 = D^{o*} \qquad (2.6.2)$$

tal que para cada $x \varepsilon D$ se tiene

$$\frac{d\phi}{dx}(x) = N(x) \qquad (2.6.3)$$

2.6.2 Teorema

Supóngase

1. $N : D \to D^*$.
2. N tiene una diferencial $\frac{dN}{dx}(x)$ en cada $x \varepsilon D$.
3. La funcional bilineal $\frac{dN}{dx}(x)$ es bidimensionalmente continua en cada $x \varepsilon D$.

Entonces, para que N sea un operador potencial en D, es necesario y suficiente que para cada $x\varepsilon D$ la funcional bilineal $\frac{dN}{dx}(x)$ sea simétrica; es decir

$$<\frac{dN}{dx}(x),\ y,\ z> = <\frac{dN}{dx}(x),\ z,\ y> \qquad (2.6.4)$$

para cada y, $z\varepsilon D$.

Demostración. Para demostrar la necesidad, obsérvese que si $\phi : D \to R^1$ es la funcional cuya derivada es N en D, entonces dados x, y, $z\varepsilon D$ defínase la función de dos variables reales ξ y η por

$$f(\xi,\ \eta) = \phi(x + \xi y + \eta z) \qquad (2.6.5)$$

Con esta definición se tiene

$$\frac{\partial f}{\partial \xi}(\xi,\ \eta) = <N(x + \xi y + \eta z),\ y> \qquad (2.6.6a)$$

$$\frac{\partial f}{\partial \eta}(\xi,\ \eta) = <N(x + \xi y + \eta z),\ z> \qquad (2.6.6b)$$

Consecuentemente

$$\frac{\partial^2 f}{\partial \eta \partial \xi}(0,\ 0) = <\frac{dN}{dx}(x),\ y,\ z> \qquad (2.6.7a)$$

$$\frac{\partial^2 f}{\partial \xi \partial \eta}(0,\ 0) = <\frac{dN}{dx}(x),\ z,\ y> \qquad (2.6.7b)$$

Ya que $\frac{dN}{dx}$ es bidimensionalmente continua en D, se tiene

$$\frac{\partial^2 f}{\partial\xi\partial\eta}(0,\ 0) = \frac{\partial^2 f}{\partial\eta\partial\xi}(0,\ 0) \tag{2.6.8}$$

lo que implica 2.6.4.

A continuación se presenta la demostración de suficiencia. Defínase una función

$$G : D \times D \to R^1$$

para cada x, yεD por la ecuación

$$G(x,\ y) = \int_0^1 \langle N(x + t(y - x)),\ y - x\rangle\, dt \tag{2.6.9}$$

La integral evidentemente existe por ser el integrando diferenciable y consecuentemente continuo. Además, si x, y, zεD, entonces

$$G(x,\ y) + G(y,\ z) = G(x,\ z) \tag{2.6.10}$$

Esto puede verse considerando las funciones de dos variables reales

$$F_1(\xi,\ \eta) = \langle N(x + \xi(y - x) + \eta(z - y)),\ y - x\rangle \tag{2.6.11a}$$

$$F_2(\xi,\ \eta) = \langle N(x + \xi(y - x) + \eta(z - y)),\ z - y\rangle \tag{2.6.11b}$$

Estas funciones están definidas en todo el plano ξ, η. Considérese en él el triángulo formado por las rectas

$$\eta = 0;\ 0 \leq \xi \leq 1 \tag{2.6.12a}$$

$$\xi = 1;\ 0 \leq \xi \leq 1 \tag{2.6.12b}$$

$$\xi = \eta;\ 0 \leq \xi = \eta \leq 1 \tag{2.6.12c}$$

Además, las ecuaciones

$$G(x,\ y) = \int_0^1 F_1(\xi,\ 0)\ d\xi \tag{2.6.13a}$$

$$G(y,\ z) = \int_0^1 F_2(0,\ \eta) d\eta \tag{2.6.13b}$$

$$G(x,\ z) = \int_0^1 \left[F_1(t,\ t) + F_2(t,\ t)\right] dt \tag{2.6.13c}$$

muestran que $G(x,\ y)$, $G(y,\ z)$ y $G(x,\ z)$ son las integrales de línea a la larga de los tres segmentos orientados de la fig 1. Pero la ec 2.6.4 implica

$$\frac{\partial F_1}{\partial \eta}(\xi,\ \eta) = \frac{\partial F_2}{\partial \xi}(\xi,\ \eta) \tag{2.6.14}$$

en el triángulo y su interior. Consecuentemente, la integral es independiente de la trayectoria (fig 1). Por lo mismo, la ec 2.6.10 se cumple.

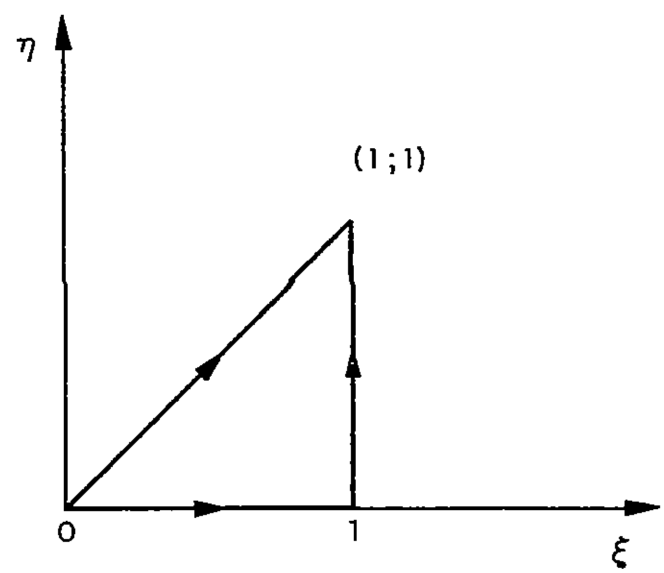

Fig 1

Sea $x_0 \varepsilon D$ fijo. Defina

$$\phi(x) = G(x_0, x) \tag{2.6.15}$$

para cada $x \varepsilon D$. Entonces, dada $y \varepsilon D$, defínase

$$F(t) = \phi(x + ty) = G(x_0, x + ty) \tag{2.6.16}$$

la cual se puede escribir, en vista de 2.6.10 como

$$F(t) = G(x_0, x) + G(x, x + ty) \tag{2.6.17}$$

Luego

$$F'(t) = \frac{d}{dt}\left[t\int_0^1 <N(x + t\tau y), y> d\tau\right] \tag{2.6.18}$$

si la derivada del miembro derecho existe. Pero ya que N es bidimensional_ mente continuo, el integrando es diferenciable y

$$F'(0) = <N(x), y> \tag{2.6.19}$$

Como esta relación es cierta para toda $y \varepsilon D$.

$$\frac{d\phi}{dx}(x) = N(x) \tag{2.6.20}$$

para toda $x \varepsilon D$.

3. FORMULACION GENERAL DE PRINCIPIOS VARIACIONALES

3.1 *Consideraciones generales*

Se estudiará el problema

$$N(x) = f \qquad (3.1.1)$$

donde N es un operador diferenciable, posiblemente no lineal

$$N : D \rightarrow D^*$$

Aquí se conserva la notación introducida en el cap 2, por lo que D es un espacio lineal, D^* su espacio dual, $x \varepsilon D$ y $f \varepsilon D^*$. Se considerará en D un subconjunto $\hat{E}$. Se llamarán estados a los elementos de D, y a los elementos de $\hat{E}$, estados admisibles.

Se supondrá sistemáticamente que $\hat{E}$ es un subespacio afín; es decir, que existe un subespacio lineal $E \subset D$ y un elemento fijo $w \varepsilon D$ tales que

$$\hat{E} = w + E \qquad (3.1.2)$$

Evidentemente, si $w = 0$ el subespacio lineal E coincide con $\hat{E}$. Se tratará exclusivamente el problema de encontrar soluciones de la ec 3.1.1 que sean estados admisibles.

Si una funcional Ω: $D \rightarrow D^*$ está dada, para cada $x \varepsilon \hat{E}$ y $y \varepsilon E$ la variación de Ω en x, se define como una funcional lineal $\delta(x) \varepsilon E^*$ dada para cada $y \varepsilon E$ por

$$\left| \frac{d}{d\lambda} \Omega\ (x + \lambda y) \right|_{\lambda = 0} = \langle \delta\Omega(x),\ y \rangle \qquad (3.1.3)$$

Es claro que

$$\langle \delta\Omega(x),\ y\rangle = \langle \Omega'(x),\ y\rangle \tag{3.1.4}$$

para toda $y\varepsilon E$. La diferencia entre $\Omega'(x)$ y $\delta\Omega(x)$ es que el dominio de la primera es más amplio (ya que es D) que el de la segunda (el cual es E). Consecuentemente, el hecho de que $\Omega'(x)$ se anule implica que $\delta\Omega(x)$ también se anula, pero no a la inversa.

En relación con la teoría que se va a desarrollar, el siguiente hecho resultará de utilidad.

3.1.1 Lema

Supóngase

i) $\Omega'(x) = N(x) - f$ (3.1.5)

ii) Para cada $x\varepsilon\hat{E}$

$\langle N(x) - f,\ y\rangle = 0$ para toda $y\varepsilon E$, implica

$$N(x) = f \tag{3.1.6}$$

Entonces para cada $x\varepsilon\hat{E}$

$$\Omega'(x) = 0 \text{ si y solo si } \delta\Omega(x) = 0 \tag{3.1.7}$$

Demostración. Ya se ha visto que si Ω' se anula entonces $\delta\Omega$ también. Falta demostrar la implicación inversa. Pero esto es obvio porque si $\delta\Omega(x) = 0$, entonces, para toda $y \varepsilon E$:

$$<N(x) - f, y> = <\Omega'(x), y> = <\delta\Omega(x), y> = 0$$

En consecuencia

$$\Omega'(x) = N(x) - f = 0$$

Se entenderá por un principio variacional una afirmación que establece que la ec 3.1.1 se satisface si y solo si la variación de cierta funcional se anula. Por otra parte, un principio de máximo es el que establece que cierta funcional alcanza su máximo en un punto si y solo si la ec 3.1.1 se satisface ahí. La definición de principio de mínimo es análoga. Por principios extremales se entienden tanto los de máximo como los de mínimo. El hecho bien conocido de que en un máximo y en un mínimo la derivada de una función necesariamente se anula, implica en este contexto que todo principio extremal es un principio variacional. Sin embargo, la afirmación recíproca no es cierta ya que la anulación de la derivada de una función no implica la presencia de un máximo o de un mínimo. Por lo mismo, la clase de los principios extremales es una subclase propia de los principios variacionales.

Para problemas lineales basta que el operador que se considere sea positivo o alternativamente negativo, para que el principio variacional correspondiente sea extremal. En general, en el caso de problemas no lineales, basta que la funcional asociada al principio variacional sea convexa para que el principio correspondiente sea extremal.

Se verá que la construcción de principios variacionales es inmediata si el operador es potencial, es decir si su derivada es simétrica. Para operado res lineales esta condición es equivalente a la de que el operador N sea simé trico. Consecuentemente, para todo operador lineal simétrico es posible construir principios variacionales.

Hay muchos operadores lineales simétricos que no son positivos ni negativos; para ellos, el principio variacional correspondiente no es extremal. Sin embargo, en condiciones generales el espacio D puede descomponerse en dos subespacios, D_1 y D_2, con la propiedad de que el operador es positivo en D_1 y negativo en D_2. En tal caso, la condición de que la derivada de la funcional se anule equivale a que, simultáneamente

a) la derivada parcial con respecto a D_1 se anule

b) la derivada parcial con respecto a D_2 se anule

Se probará que entre los elementos de D, que satisfacen *a*, la funcional alcanza su máximo cuando se satisface *b*, y entre los elementos de D que satisfacen *b*, la funcional alcanza su mínimo cuando se satisface *a*. A esta clase de principios duales se le llamará complementarios o recíprocos y puede formularse para operadores no lineales, introduciendo la noción de funcional *silla*.

Hay una clase importante de principios duales que ha recibido una amplia atención en la literatura (refs 9 a 11), llamados principios de Hamilton generalizados. La formulación aquí presentada,puede derivarse de los principios anteriores cuando la funcional tiene una forma especial. Finalmente, los principios de Lagrange generalizados que también han sido discutidos ampliamente (refs 9 a 11) pueden derivarse de los de Hamilton por medio de la transformada de Legendre.

Casi todos los principios variacionales conocidos para ecuaciones de la forma 3.1.1 son de alguno de estos tipos. Por tanto, la formulación presentada en este capítulo es de considerable generalidad.

Se supone que el operador N definido en esta sección posee derivada en D y que la misma es continua.

3.2 Condiciones suficientes para principios variacionales

El teorema siguiente, que establece condiciones suficientes para la existencia de principios variacionales, es una consecuencia inmediata del teorema 2.6.2.

3.2.1 Teorema

Supóngase que para cada $x \epsilon D$, la funcional bilineal $\frac{dN}{dx}(x)$ es simétrica. Entonces, $x_o \epsilon D$ es solución de la ec 3.1.1, si y solo si

$$\frac{d\Omega}{dx}(x) = 0 \qquad (3.2.1)$$

donde

$$\Omega(x) = \psi(x) - \langle f, x \rangle \qquad (3.2.2)$$

para cada $x \epsilon D$. Aquí

$$\psi : D \rightarrow R^1 \qquad (3.2.3)$$

es tal que

$$\frac{d\psi}{dx}(x) = N(x) \;\; ; \; x \epsilon D \qquad (3.2.4)$$

Demostración. La existencia de la funcional ψ está garantizada por el teorema 2.6.2. Con la definición 3.2.2 de Ω, se tiene

$$\frac{d\Omega}{dx}(x) = \frac{d\psi}{dx}(x) - f = N(x) - f \qquad (3.2.5)$$

para toda $x \varepsilon D$. Consecuentemente, 3.1.1 se satisface si y solo si 3.2.1 se satisface.

Si el operador N es lineal, el teorema anterior toma una forma especialmente sencilla. En este caso escríbase

$$N = L \qquad (3.2.6)$$

Aplicando la definición de la derivada a L lineal, se obtiene

$$\left\langle \frac{dL}{dx}(x), y, z \right\rangle = \langle Ly, z \rangle \qquad (3.2.7)$$

Defínase la funcional bilineal $\mathcal{L}$ en D por la ecuación

$$\langle \mathcal{L}, y, z \rangle = \langle Ly, z \rangle \qquad (3.2.8)$$

para toda $y, z \varepsilon D$. Se observa entonces que la ecuación 3.2.7 implica

$$\frac{dL}{dx}(x) = \mathcal{L} \qquad (3.2.9)$$

para toda $x \varepsilon D$. Es decir, la funcional bilineal $\frac{dL}{dx}(x)$ es independiente de x (una funcional bilineal constante).

3.2.2 Definición

Se dice que el operador

$$L : D \rightarrow D^* \tag{3.2.10}$$

es simétrico si

$$\langle Ly, z\rangle = \langle Lz, y\rangle \tag{3.2.11}$$

para toda $y, z \varepsilon D$.

3.2.3 Corolario

Supóngase que

$$L : D \rightarrow D^* \tag{3.2.12}$$

es lineal y simétrico, y que $f \varepsilon D^*$. Entonces $x_0 \varepsilon D$ satisface la ecuación

$$Lx = f \tag{3.2.13}$$

si y solo si

$$\frac{d\Omega}{dx}(x_0) = 0 \tag{3.2.14}$$

Aquí

$$\Omega(x) = \frac{1}{2} \langle Lx,\ x\rangle - \langle f,\ x\rangle \tag{3.2.15}$$

Demostración. Es consecuencia del teorema 3.2.1. En efecto, escriba

$$\psi(x) = \frac{1}{2} \langle Lx,\ x\rangle \tag{3.2.16}$$

Entonces

$$\frac{d\Omega}{dx}(x) = Lx - f \tag{3.2.17}$$

Ejemplo 1. Sea D el espacio de las funciones con segunda derivada continua en $\left[0,\ 1\right]$. Defina L por

$$\langle Lu,\ v\rangle = \int_0^1 u''(\tau)v(\tau)d\tau + u(1)v'(1) - u(0)v'(0) \tag{3.2.18}$$

para cada u, $v \varepsilon D$ y sea $f \varepsilon D^*$, definido por

$$\langle f,\ v\rangle = \int_0^1 f_R(\tau)v(\tau)d\tau + u_1 v'(1) - u_0 v'(0) \tag{3.2.19}$$

donde u_1 y u_0 son dos números reales y f_R una función continua.

Integrando por partes la ec 3.2.18, es fácil probar que L es simétrico. Consecuentemente, el corolario 3.2.3 proporciona un principio variacional para la ecuación

$$Lu = f \tag{3.2.20}$$

en que la funcional Ω está dada por

$$\Omega(u) = -\frac{1}{2}\int_0^1 \left[u'(\tau)u'(\tau) - 2f_R(\tau)u(\tau)\right]d\tau + \left[u(1) - u_1\right]u'(1)$$

$$- \left[u(0) - u_0\right]u'(0) \qquad (3.2.21)$$

También obsérvese que la ec 3.2.20 se satisface si y solo si $\langle Lu - f, v\rangle$ se anula para toda $v \varepsilon D$.

Pero

$$\langle Lu - f, v\rangle = \int_0^1 \left[u''(\tau) - f_R(\tau)\right]v(\tau)d\tau + \left[u(1) - u_1\right]v'(1) -$$

$$- \left[u(0) - u_0\right]v'(0) \qquad (3.2.22)$$

por tanto, la ec 3.2.20 se satisface si y solo si

$$u'' = f_R \qquad (3.2.23a)$$

$$u(0) = u_0 \qquad (3.2.23b)$$

$$u(1) = u_1 \qquad (3.2.23c)$$

El sistema de ecs 3.2.23 puede pensarse como la ecuación de Laplace en una dimensión, con valores en la frontera (problema de Dirichlet).

Habitualmente, el problema de interés se plantea en términos de ecuaciones diferenciales y condiciones de frontera* (en el ejemplo sería el sistema 3.2.23). Para aplicar la teoría desarrollada aquí, es necesario asociar con

---

* Para lograr mayor brevedad en el lenguaje se considerarán sistemáticamente las condiciones iniciales como caso particular de condiciones de frontera.

dicho sistema de ecuaciones un operador con valores en funcionales para el cual la ec 3.1.1 sea equivalente al sistema de ecuaciones diferenciales.

También conviene observar que si en la ec 3.1.1 el operador N no es potencial, en principio es posible aplicarle una transformación para obtener una ecuación equivalente. En el caso lineal considérese una transformación lineal

$$\Sigma : D^* \to D^*$$

no singular. Aplicando esta transformación a la 3.2.20, se obtiene

$$\Sigma Lu = \Sigma f$$

que es equivalente por ser Σ no singular. Si ΣL es simétrico, el corolario 3.2.3 es aplicable a esta ecuación transformada. Esta forma de proceder ha sido usada con éxito en problemas de condiciones iniciales (ref 5).

Ejemplo 2. Sea D el espacio de las funciones con primera derivada continua en $\left[0,\ 1\right]$. Defina L por

$$<Lu,\ v> = \int_0^1 u'(\tau)v(\tau)d\tau + u(0)v(0) \qquad (3.2.24)$$

para cada u, $v\varepsilon D$ y sea $f\varepsilon D^*$, definida por

$$<f,\ v> = \int_0^1 f_R(\tau)v(\tau)d\tau + u_0v(0) \qquad (3.2.25)$$

donde u_0 es un número real y f_R una función continua. Es fácil ver que en este caso la ec 3.2.20 es equivalente al problema de condiciones iniciales

$$u' = f_R \tag{3.2.26a}$$

$$u(0) = u_0 \tag{3.2.26b}$$

Sin embargo, la relación

$$\int_0^1 u'(\tau)v(\tau)d\tau + u(0)v(0) = - \int_0^1 v'(\tau)u(\tau) + u(1)v(1) \tag{3.2.27}$$

muestra que el operador L no es simétrico. Conviene, por lo mismo, definir una transformación del tipo $\Sigma : D^* \to D^*$. Para esto, se define primero una transformación $\sigma : D \to D$ dada para cada $v \varepsilon D$ por la ecuación

$$(\sigma v)\ (t) = v(1 - t) \tag{3.2.28}$$

La transformación $\Sigma : D^* \to D^*$ se define ahora para cada $\alpha \varepsilon D^*$ mediante la ecuación

$$\langle \Sigma\alpha, v \rangle = \langle \alpha, \sigma v \rangle \tag{3.2.29}$$

Consecuentemente

$$\begin{aligned} \langle \Sigma Lu, v \rangle &= \langle Lu, \sigma v \rangle \\ &= \int_0^1 u'(\tau)v(1 - \tau)d\tau + u(0)v(1) \\ &= \int_0^1 v'(\tau)u(1 - \tau)d\tau + v(0)u(1) \end{aligned} \tag{3.2.30}$$

y ΣL es un operador simétrico. Por otra parte

$$\langle \Sigma f, v \rangle = \int_0^1 f_R(\tau)v(1 - \tau)d\tau + u_0 v(1) \tag{3.2.31}$$

Puede comprobarse fácilmente que el operador Σ es no singular y por lo mismo son aplicables las observaciones hechas anteriormente. El principio que se deriva de esta discusión está asociado a la funcional $\Omega : D \to R^1$ dada para cada $u \varepsilon D$ por

$$\Omega(u) = \frac{1}{2} <\Sigma L u, u> - <\Sigma f, u> \tag{3.2.32}$$

que es

$$\Omega(u) = \frac{1}{2} \int_0^1 \left[u'(\tau) - 2f_R(\tau)\right] u(1-\tau)d\tau + \frac{1}{2}\left[u(0) - 2u_0\right]u(1) \tag{3.2.33}$$

En resumen, el problema de condiciones iniciales definido por las ecs 3.2.26 es equivalente a la condición de que la derivada de la funcional Ω dada por 3.2.33 se anule.

3.3 *Principios extremales*

Para discutir los principios extremales es conveniente introducir algunos conceptos auxiliares. Se considerará aquí un subespacio afín $\hat{E} \subset D$, como se explicó al principio de este capítulo.

3.3.1 Definición

Se dice que $x_0 \varepsilon \hat{E}$ es un máximo de Ω si

$$\Omega(x_0) \geq \Omega(x) \tag{3.3.1}$$

para toda $x \varepsilon D$. En tal caso se dice que $\Omega(x_0)$ es el valor máximo de Ω. En caso de que en la ec 3.3.1, la igualdad funcione solo cuando $x = x_0$, al máximo se le llama absoluto.

Un mínimo y un mínimo absoluto de una funcional, se definen en forma análoga invirtiendo el signo en la desigualdad 3.3.1. Además, se dice que x_0 es un punto extremo de Ω, si es un máximo o un mínimo de Ω.

3.3.2 Definición

Se llama *principio de máximo* de la ec 3.1.1 a una afirmación que establece que una funcional alcanza un máximo si y solo si la ec 3.1.1 se satisface. Análogamente se define un *principio de mínimo*. La clase formada por los principios de máximo y por los principios de mínimo, constituye la de los *principios extremales*.

El siguiente teorema aclara la relación entre los principios extremales y los principios variacionales.

3.3.3 Teorema

Sea la funcional

$$\Omega : D \to R^1 \tag{3.3.2}$$

diferenciable. Entonces, si Ω alcanza un valor extremo en $x_0 \varepsilon D$, se tiene

$$\delta\Omega(x_0) = 0 \tag{3.3.3}$$

Demostración. Para el caso en que x_0 es un máximo, las desigualdades

$$<\frac{d\Omega}{dx}(x_0),\ y> = \lim_{t\to 0} \frac{\Omega(x_0 + ty) - \Omega(x_0)}{t} \leq 0 \tag{3.3.4a}$$

y

$$\left\langle \frac{d\Omega}{dx}(x_0), -y \right\rangle = \lim_{t\to 0} \frac{\Omega(x_0 - ty) - \Omega(x_0)}{t} \leq 0 \qquad (3.3.4b)$$

implican que

$$\left\langle \frac{d\Omega}{dx}(x_0), y \right\rangle = 0 \qquad (3.3.5)$$

para toda $y \varepsilon E$, y consecuentemente la relación 3.3.3. Si x_0 es un mínimo la demostración es similar.

3.4 *Principios extremales para operadores lineales*

El teorema 3.2.1 permite construir un principio variacional para cada operador potencial. Es interesante establecer bajo qué condiciones dichos principios son extremales. En el caso de operadores lineales es fácil establecer una condición suficiente para que esto suceda.

3.4.1 Teorema

Bajo las hipótesis del corolario 3.2.3, suponga además:

i) Para cada $x \varepsilon \hat{E}$

$\langle Lx - f, y \rangle = 0$ para toda $y \varepsilon E$, implica $Lx = f$

ii) L es no negativo en E

Entonces, la funcional Ω dada por 3.2.15 alcanza un mínimo en $x_0 \varepsilon \hat{E}$ si y solo si x_0 satisface 3.2.13

Demostración. Conviene introducir una notación adecuada. Para x_+, $x_- \varepsilon D$, escríbase

$$\Delta x = x_+ - x_- \qquad (3.4.1a)$$

$$\Delta\Omega = \Omega(x_+) - \Omega(x_-) \qquad (3.4.1b)$$

$$\frac{d\Omega}{dx_+} = \Omega'(x_+) \qquad (3.4.1c)$$

$$\frac{d\Omega}{dx_-} = \Omega'(x_-) \qquad (3.4.1d)$$

Póngase $x_- = x_0 \varepsilon \hat{E}$ y x_+ cualquier otro elemento de $\hat{E}$.

Entonces $\Delta x \varepsilon E$, y consecuentemente

$$0 \leq \frac{1}{2} <L\Delta x,\ \Delta x> = \psi(\Delta x) = \Delta\Omega - <\frac{d\Omega}{dx_-}\ ,\ \Delta x> \qquad (3.4.2)$$

para toda $\Delta x \varepsilon E$, donde

$$\frac{d\Omega}{dx_-} = Lx_0 - f \qquad (3.4.3)$$

en vista de 3.2.17.

Luego, si x_0 satisface 3.2.13, la relación 3.4.2 implica

$$\Omega(x_0) \leq \Omega(x_+) \qquad (3.4.4)$$

para toda $x_+ \varepsilon \hat{E}$. Es decir, x_0 es un mínimo.

Inversamente, si x_0 es un mínimo, entonces por el teorema 3.3.4 la variación $\delta\Omega(x_0)$ se anula, lo que implica que la derivada $\Omega'(x_0)$ también se anula por el lema 3.1.1. Consecuentemente, se satisface 3.2.13.

3.4.2 Corolario

Si en el teorema 3.4.1 se supone que L es positivo en E, entonces Ω alcanza un mínimo absoluto en x_0, si y solo si x_0 satisface 3.2.13.

Demostración. Porque las desigualdades 3.4.3 y 3.4.4 son en este caso estrictas siempre que $x_+ \neq x_- = x_0$.

3.4.3 Corolario

Cuando L es positivo, la solución de 3.2.13, si existe, es única.

Demostración. En vista del corolario 3.4.2, es inmediata, ya que el máximo absoluto, si existe, es único.

Obsérvese que los teoremas y corolarios de esta sección se conservan si en todas partes *positivo* se intercambia con *negativo* y simultáneamente *mínimo* con *máximo*.

Ejemplo 3. Tómense D y L como en el ejemplo 1. Se observa que L no es positivo en D. Por ello, conviene definir $\hat{E}$ como el subconjunto de D, cuyos elementos u satisfacen las condiciones 3.2.23b y c. En este caso, cada uno de los elementos $v \in E \subset D$ satisface las condiciones homogéneas

$$v(0) = v(1) = 0 \qquad (3.4.5)$$

Es fácil observar que las hipótesis del teorema 3.4.1 (sustituyendo positivo por negativo) se satisfacen. En efecto, para cada $u \in \hat{E}$ y $v \in E$, se tiene

$$<Lu, - f, v> = \int_0^1 \left[u''(\tau) - f_R(\tau)\right] v(\tau)d\tau \qquad (3.4.6)$$

Luego, la hipótesis *i* del teorema mencionado se satisface. Además, L es negativo en E, porque si $v \varepsilon E$

$$<Lv, v> = - \int_0^1 v'(\tau)v'(\tau)d\tau \qquad (3.4.7)$$

Por lo mismo, el teorema 3.4.1 y los corolarios 3.4.2 y 3.4.3 son aplicables. En resumen, se puede decir que si se toman las funciones admisibles de $\hat{E}$ como las funciones con segunda derivada continua que satisfacen las condiciones de frontera 3.2.23b y c, y para cada $u\varepsilon\hat{E}$, la funcional Ω está dada por

$$\Omega(u) = - \frac{1}{2} \int_0^1 \left[u'(\tau)u'(\tau) + 2f_R(\tau)u(\tau)\right] d\tau \qquad (3.4.8)$$

Entonces, una función $u\varepsilon\hat{E}$ es solución del sistema 3.2.23 si y solo si u es el máximo absoluto de Ω. Además, tal solución es única.

3.5 *Funcionales convexas y cóncavas*

Los resultados del subcap 3.4 pueden extenderse fácilmente a operadores no lineales. Para ello conviene introducir las nociones de funcionales cóncavas y convexas. La propiedad que servirá de base para estas definiciones es la relación 3.4.2.

3.5.1 Definición

Se dice que la funcional Ω es convexa si para toda pareja x_+, $x_- \varepsilon \hat{E}$ se tiene

$$\Delta\Omega - <\frac{d\Omega}{dx_-}, \Delta x> \geq 0 \qquad (3.5.1)$$

Es estrictamente convexa si la igualdad solo funciona cuando $x_+ = x_-$.

3.5.2 Definición

$\Omega(x)$ es cóncava si $-\Omega(x)$ es convexa. En forma análoga, estrictamente cóncava si $-\Omega(x)$ es estrictamente convexa.

Intercambiando x_+ con x_- es fácil establecer definiciones alternativas de funcionales convexas y cóncavas.

3.5.3 Teorema

La funcional Ω es convexa si y solo si para toda pareja $x_+, x_- \varepsilon \hat{E}$, se tiene

$$\Delta\Omega - \left\langle \frac{d\Omega}{dx_+}, \Delta x \right\rangle \leq 0 \qquad (3.5.2)$$

Por otra parte, es cóncava si y solo si para toda pareja $x_+, x_- \varepsilon D$, se tiene

$$\Delta\Omega - \left\langle \frac{d\Omega}{dx_+}, \Delta x \right\rangle \geq 0 \qquad (3.5.3)$$

Las propiedades estrictas se definen correspondiendo al caso en que las igualdades de 3.5.2 y 3.5.3 solo funcionan cuando $x_+ = x_-$.

3.6 *Principios extremales para operadores no lineales*

Utilizando las nociones anteriores, se establecen con facilidad los resultados que a continuación se enuncian.

3.6.1 Teorema

Sea

$$N : D \to D^* \qquad (3.6.1)$$

un operador potencial y la funcional ψ tal que

$$\frac{d\psi}{dx} = N \tag{3.6.2}$$

Defínase la funcional Ω por la ecuación

$$\Omega(x) = \psi(x) - \langle f, x\rangle \tag{3.6.3}$$

Supóngase que $\Omega(x)$ es convexa. Entonces $x_0 \varepsilon \hat{E}$ satisface la ec 3.1.1, si y solo si x_0 es un mínimo de Ω.

Demostración. Póngase $x_- = x_0$ y sea x_+ cualquier elemento de $\hat{E}$. Por definición Ω satisface a la relación 3.5.1.

Supóngase que x_0 es solución de la ec 3.1.1, entonces, en vista del teorema 3.2.3, la relación 3.5.1 implica

$$\Omega(x_+) \geq \Omega(x_-) = \Omega(x_0) \tag{3.6.4}$$

Como x_+ es cualquier elemento de D, x_0 es un mínimo de Ω.

Inversamente, si x_0 es un mínimo de Ω, el teorema 3.3.4 y el hecho de que

$$\frac{d\Omega}{dx}(x) = \frac{d\psi}{dx}(x) - f \tag{3.6.5}$$

implican 3.1.1.

3.6.2 Corolario

Si $\Omega(x)$ es estrictamente convexa en el teorema anterior, entonces Ω alcanza un mínimo absoluto, si y solo si x_0 satisface la ec 3.1.1.

Demostración. Porque en 3.5.1 la desigualdad es estricta siempre que $x_+ \neq x_- = x_0$.

3.6.3 Corolario

Si Ω es estrictamente convexa, la solución de la ec 3.1.1, si existe, es única.

Demostración. Se sigue del corolario 3.6.2, ya que el mínimo absoluto, si existe, es único.

Es importante hacer notar que la hipótesis de convexidad de Ω puede ser sustituida por la convexidad de ψ. El resultado preciso está contenido en el lema siguiente.

3.6.4 Lema

La funcional Ω definida por la ec 3.6.3 es convexa si y solo si ψ lo es. Además, Ω es estrictamente convexa si y solo si ψ lo es.

Demostración. Porque

$$\Delta\Omega - \left\langle \frac{\partial\Omega}{\partial x_-}, \Delta x \right\rangle = \Delta\psi - \left\langle \frac{\partial\psi}{\partial x_-}, \Delta x \right\rangle \qquad (3.6.6)$$

Es importante observar que todos los teoremas y corolarios de este subcapítulo se conservan si en todas partes *convexo* se sustituye por *cóncavo* y *mínimo* por *máximo*.

3.7 *Condiciones locales para convexidad y concavidad*

Para funcionales que poseen derivadas de orden superior, pueden establecerse condiciones suficientes para la convexidad.

3.7.1 Teorema

Sea

$$\Omega : D \to R^1 \tag{3.7.1}$$

Supóngase

i) Ω tiene segunda derivada en D

ii) El operador

$$\Omega'' : D \to D^{2*}$$

es bidimensionalmente continuo en D.

Entonces, una condición necesaria y suficiente para que Ω sea convexa en D, es que para toda $x \varepsilon D$, $\Omega''(x) \varepsilon D^{2*}$ sea positiva. Esta afirmación sigue siendo válida si *convexa* se remplaza por *estrictamente convexa* y *positiva* por *positiva definida*.

Demostración. Con la notación del cap 2 y usando el desarrollo de Taylor con residuo (teorema 2.5.4), se tiene

$$\Delta\Omega - <\frac{d\Omega}{dx_+}, \Delta x> = \frac{1}{2} <\Omega''(\xi), \Delta x, \Delta x> \tag{3.7.2}$$

donde ξ es un elemento del segmento que une a x_- con x_+. El teorema es una consecuencia inmediata de 3.7.2.

En vista de la definición 3.5.2 de concavidad, es claro que en este teorema *convexidad* se puede remplazar por *concavidad* si al mismo tiempo *positivo* se remplaza por *negativo*.

3.8 *Principios duales para operadores lineales*

En esta sección se desarrolla una clase de principios variacionales duales que tiene relación con la desarrollada por Noble y Sewell (ref 11). La atención se restringe a operadores lineales, dejándose para el subcap 3.10 su extensión a operadores no lineales.

El corolario 3.2.3 proporciona un principio variacional para operadores lineales y simétricos. Además, el teorema 3.4.1 establece que dicho principio es extremal cuando L es positivo o alternativamente cuando L es negativo. En el caso general L no es positivo ni negativo y los resultados del subcap 3.4 no son aplicables a él. Sin embargo, los principios variacionales duales que se desarrollarán en esta sección no adolecen de estas limitaciones y son por ello de aplicabilidad más general.

Sea Ω la funcional del corolario 3.2.3, definida por la ec 3.2.15. Se supone además que el subespacio afín $\hat{E}\subset D$ de los estados admisibles es tal que la hipótesis *ii* del lema 3.1.1 se satisface. Entonces, bajo las hipótesis del corolario 3.2.3, $x_0 \varepsilon \hat{E}$ satisface la ec 3.2.13 si y solo si

$$\delta\Omega(x_0) = 0 \tag{3.8.1}$$

Por otra parte, supóngase que E_1 y E_2 son una descomposición de E. En tal caso, los subespacios lineales E_1, E_2 tienen la propiedad de que para cada $y \varepsilon E$ existen $y_1 \varepsilon E_1$ y $y_2 \varepsilon E_2$ únicos, tales que

$$y = y_1 + y_2 \tag{3.8.2}$$

Entonces, para cada $x \varepsilon \hat{E}$:

$$\delta\Omega(x) = \delta_1\Omega(x) + \delta_2\Omega(x) \tag{3.8.3}$$

donde $\delta_1\Omega(x)\varepsilon E^*$ y $\delta_2\Omega(x)\varepsilon E^*$ están definidas por

$$<\delta_1\Omega(x),\ y> = <\delta\Omega(x),\ y_1> \tag{3.8.4}$$

$$<\delta_2\Omega(x),\ y> = <\delta\Omega(x),\ y_2> \tag{3.8.5}$$

válida para cada $y\varepsilon E$ que cumple la ec 3.8.2.

En vista de la relación 3.8.3, la ec 3.8.1 es equivalente al sistema de ecuaciones

$$\delta_1\Omega(x_0) = 0 \tag{3.8.6a}$$

y

$$\delta_2\Omega(x_0) = 0 \tag{3.8.6b}$$

Se supondrá, además, que L es no negativo en E_1 y no positivo en E_2; es decir, para toda $y_1\varepsilon E_1$

$$\psi(y_1) = \frac{1}{2}<Ly_1,\ y_1> \geq 0 \tag{3.8.7a}$$

y para toda $y_2\varepsilon E_2$

$$\psi(y_2) = \frac{1}{2}<Ly_2,\ y_2) \leq 0 \tag{3.8.7b}$$

Con esta notación pueden formularse los principios duales a que se ha hecho referencia antes.

3.8.1 Teorema

Supóngase que

a) $L : D \to D^*$ es lineal y simétrico

b) Los subespacios lineales E_1, E_2 constituyen una descomposición de E, con la propiedad de que L es no negativo en E_1 y no positivo en E_2.

Defina dos subconjuntos A, B de $\hat{E}$:

$$A = \{x_a = x_{1a} + x_{2a} \mid x_a \text{ satisface } 3.8.6a\} \qquad (3.8.8a)$$

$$B = \{x_b = x_{1b} + x_{2b} \mid x_b \text{ satisface } 3.8.6b\} \qquad (3.8.8b)$$

Entonces:

i) Para cada $x_a \varepsilon A$ y $x_b \varepsilon B$

$$\Omega(x_a) \leq \Omega(x_b) \qquad (3.8.9)$$

ii) Si existe un elemento $x_0 \varepsilon \hat{E}$ que es solución de la ec 3.8.1, se tiene:

α) el valor máximo de la funcional Ω se alcanza en x_0

β) El valor mínimo de la funcional Ω se alcanza en x_0

γ) El valor máximo de Ω en A y el valor mínimo de Ω en B coinciden.

Demostración. Se introduce primero una notación conveniente para facilitar la demostración de este teorema: Dados dos elementos cualesquiera

$$x_+ = (x_{1+}\ ,\ x_{2+})\varepsilon\hat{E} \text{ y } x_- = (x_{1-},\ x_{2-})\varepsilon\hat{E}$$

se define

$$\Delta x_1 = x_{1+} - x_{1-} \tag{3.8.10a}$$

$$\Delta x_2 = x_{2+} - x_{2-} \tag{3.8.10b}$$

$$\Delta\Omega = \Omega(x_{1+}\ ,\ x_{2+}) - \Omega(x_{1-}\ ,\ x_{2-}) \tag{3.8.10c}$$

$$\delta_{1-}\Omega = \delta_1\Omega(x_{1-}\ ,\ x_{2-}) \tag{3.8.11a}$$

$$\delta_{2+}\Omega = \delta_2\Omega(x_{1+}\ ,\ x_{2-}) \tag{3.8.11b}$$

En vista de las relaciones 3.8.7 y de que $\Delta x_1\varepsilon E_1$ y $\Delta x_2\varepsilon E_2$, se tiene

$$\psi(\Delta x_1) - \psi(\Delta x_2) = \Delta\Omega - \langle\delta_{1-}\Omega,\ \Delta x_1\rangle - \langle\delta_{2+}\Omega\ ,\ \Delta x_2\rangle \geq 0 \tag{3.8.12}$$

Dados $x_a\varepsilon A$ y $x_b\varepsilon B$ póngase $x_- = x_a$ y $x_+ = x_b$. Entonces, 3.8.12 implica 3.8.9. Finalmente, si x_0 satisface 3.8.1, entonces $x_0\varepsilon A\ B$, por lo que 3.8.9 implica

$$\Omega(x_a) \leq \Omega(x_0) \leq \Omega(x_b) \tag{3.8.13}$$

para toda $x_a\varepsilon A$ y $x_b\varepsilon B$. En vista de 3.8.13, la segunda parte del teorema es ahora obvia.

3.8.2 Corolario

Con las hipótesis del teorema 3.8.1, supóngase además que L es positivo en E_1 y negativo en E_2. Entonces

i) Para cada $x_a \varepsilon A$ y $x_b \varepsilon B$ se tiene

$$\Omega(x_a) < \Omega(x_b) \tag{3.8.14}$$

siempre que $x_a \neq x_b$.

ii) Si la solución del sistema de ecs 3.8.6 existe, es única

iii) Si el valor máximo de Ω en A y el valor mínimo de Ω en B se alcanzan y ambos son iguales, entonces el máximo en A y el mínimo en B son la solución del sistema de ecs 3.8.6.

Demostración. La afirmación *i* se sigue de la demostración del teorema tomando en cuenta que en 3.8.12 la igualdad solo funciona cuando

$$\Delta x = \Delta x_1 + \Delta x_2 = 0 \tag{3.8.15}$$

Para probar *ii*, obsérvese que si x_0 y x_0' son soluciones del sistema, entonces x_0, $x_0' \varepsilon A \cap B$. Supóngase $x_0' \neq x_0$, en tal caso

$$\Omega(x_0') < \Omega(x_0)$$

y

$$\Omega(x_0) < \Omega(x_0')$$

en vista de 3.8.14. Estas desigualdades constituyen una constradicción.

Finalmente, bajo las hipótesis del inciso *iii*, existen $x_a \varepsilon A$ y $x_b \varepsilon B$, tales que

$$\Omega(x_a) = \Omega(x_b)$$

En vista del inciso *i* ya probado, esta igualdad implica que

$$x_a = x_b$$

Consecuentemente $x_a \varepsilon A \cap B$, y x_a es solución del sistema de ecs 3.8.6.

Ejemplo 4. El principio variacional que se formuló en el ejemplo 2 puede extenderse a un principio dual. Tome $\hat{E} = E = D$ y E_1, E_2 como los subespacios de funciones simétricas y antisimétricas. Obsérvese que E_1, E_2 constituyen una descomposición de E y que, además, si εE_1

$$\langle \Sigma Lu,\ u \rangle = u(0)u(0) \geq 0 \qquad (3.8.16a)$$

y si $u \varepsilon E_2$

$$\langle \Sigma Lu,\ u \rangle = -u(0)u(0) \leq 0 \qquad (3.8.16b)$$

Por lo mismo, el teorema 3.8.1 es aplicable al operador ΣL y la funcional Ω definidos por las ecs 3.2.30 y 3.2.33, respectivamente. Se ha establecido así un principio dual para el sistema de ecuaciones diferenciales 3.2.26.

El interés del principio variacional de este ejemplo, se restringe por el hecho de que el corolario 3.8.2 no es aplicable, ya que el operador ΣL no es positivo en E_1 ni negativo en E_2. A continuación se presenta un ejemplo que carece de esta limitación.

Ejemplo 5. Considérense las ecuaciones

$$\frac{\partial u}{\partial t} - \frac{\partial^2 u}{\partial x^2} = f_R(x,t) \tag{3.8.17a}$$

$$u(0,t) = u_{B1}(t) \tag{3.8.17b}$$

$$u(1,t) = u_{B2}(t) \tag{3.8.17c}$$

$$u(x,0) = u_0(x) \tag{3.8.17d}$$

válidas en la región $0 \leq x \leq 1$, $0 \leq t \leq 1$. Las funciones f_R, u_{B_1}, u_{B_2} y u_0 son datos del problema y están definidas en el intervalo correspondiente.

Considérese el espacio D de funciones con segundas derivadas continuas en la región $0 \leq x \leq 1$, $0 \leq t \leq 1$ y supóngase que f_R, u_{B_1}, u_{B_2} y u_0 satisfacen la ec 3.8.17 para alguna $u \varepsilon D$. Defínase $L : D \rightarrow D^*$ por la ecuación

$$\begin{aligned} \langle Lu, v\rangle = & \int_0^1\int_0^1 \left[\frac{\partial u}{\partial t}(x,t) - \frac{\partial^2 u}{\partial x^2}(x,t)\right] v(x,1-t)dxdt \\ & + \int_0^1 u(x,0)v(x,1)dx + \int_0^1 \left[u(0,t)\frac{\partial v}{\partial x}(0,1-t)\right. \\ & \left. - u(1,t)\frac{\partial v}{\partial x}(1,1-t)\right]dt \end{aligned} \tag{3.8.18}$$

y, consecuentemente, $f \varepsilon D^*$ por

$$\begin{aligned} \langle f, v\rangle = & \int_0^1\int_0^1 f_R(x,t)v(x,1-t)dxdt \\ & + \int_0^1 u_0(x)v(x,1)dx + \int_0^1 \left[u_{B_1}(t)\frac{\partial v}{\partial x}(0,1-t)\right. \\ & \left. - u_{B2}(t)\frac{\partial v}{\partial x}(1,1-t)\right]dt \end{aligned} \tag{3.8.19}$$

Entonces

$$Lu = f \qquad (3.8.20)$$

es equivalente al sistema 3.8.17.

Ya que L definido por la ec 3.8.18 es un operador simétrico, un principio variacional para el sistema 3.8.17, se puede definir mediante la funcional

$$\Omega(u) = \frac{1}{2}\int_1^0 \left[\frac{\partial u}{\partial x} * \frac{\partial u}{\partial x} + u * \frac{\partial u}{\partial t} - 2f_R * u\right] dx$$

$$+ \frac{1}{2}\int_1^0 \left[u(x,0) - 2u_0(x)\right] u(x,1)dx$$

$$+\left\{\left[u(0) - u_{B1} * \frac{\partial u}{\partial x}\right](0) - \left[u(1) - u_{B2}\right] * \frac{\partial u}{\partial x}(1)\right\} \qquad (3.8.21)$$

Tómese $\hat{E}$ como el subconjunto de D que satisface las ecs 3.8.17b y c. Entonces, toda función $v \varepsilon E$ satisface las ecuaciones

$$u(0,t) = 0 \qquad (3.8.22a)$$

y

$$u(1,t) = 0 \qquad (3.8.22b)$$

Sean E_1 y E_2 los subespacios de E, en que los elementos de E_1 son funciones simétricas, mientras que los de E_2 son antisimétricas. Entonces, para toda $v \varepsilon E_1$

$$\langle Lv, v\rangle = \frac{1}{2}\int_0^1\int_0^1 \left[\frac{\partial v}{\partial x}(x,t)\right]^2 dx\, dt$$

$$+ \frac{1}{2}\int_0^1 \left[v(x,0)\right]^2 dx \geq 0 \qquad (3.8.23a)$$

y para toda $v \varepsilon E_2$

$$\langle Lv,v\rangle = -\frac{1}{2}\int_0^1\int_0^1 \left[\frac{\partial v}{\partial x}(x,t)\right]^2 dx\, dt$$

$$-\frac{1}{2}\int_0^1 \left[v(x,0)\right]^2 dx \leq 0 \tag{3.8.23b}$$

Las relaciones 3.8.23 permiten observar que L es positivo definido en E_1 y negativo definido en E_2. Consecuentemente, son aplicables el teorema 3.8.1 y el corolario 3.8.2. Se ha obtenido así un principio variacional dual para el problema de condiciones iniciales para la ecuación del calor definido por las ecuaciones 3.8.17.

Tomando las derivadas parciales de la funcional Ω dada por la ec 3.8.21 se observa que las ecuaciones duales son

$$\left.\begin{aligned} \left[\frac{\partial u}{\partial t}\right]_s - \left[\frac{\partial^2 u}{\partial x^2}\right]_s &= \left[f_R\right]_s \\ u(x,0) &= u_0(x) \end{aligned}\right\} \tag{3.8.24a}$$

y

$$\left.\begin{aligned} \left[\frac{\partial u}{\partial t}\right]_a - \left[\frac{\partial^2 u}{\partial x^2}\right]_a &= \left[f_R\right]_a \\ u(x,0) &= u_0(x) \end{aligned}\right\} \tag{3.8.24b}$$

y que entre todas las funciones que satisfacen la ec 3.8.24a, la funcional Ω alcanza un máximo cuando se satisface el sistema 3.8.17. Inversamente, entre las funciones que satisfacen 3.8.24b, Ω alcanza un mínimo cuando se satisface el sistema 3.8.17.

3.9 *Funcionales silla*

Puede efectuarse con gran sencillez la generalización de los resultados del subcap 3.8 a operadores no lineales. Para ello basta observar que las propiedades demostradas en la sección anterior se derivaron de la desigualdad 3.8.13. Consecuentemente, se introducirá el concepto de funcional silla como a continuación se indica:

Sea

$$X : D \to R^1 \tag{3.9.1}$$

una funcional, $\hat{E}$ el conjunto de los estados admisibles y E_1, E_2 una descomposición de E.

3.9.1 Definición

Se dice que la funcional X es silla con convexidad en E_1 y concavidad en E_2, si para cualquier pareja de puntos x_+, $x_- \varepsilon \hat{E}$, se tiene

$$\Delta X - \langle \delta_{1-} X, \Delta x_1 \rangle - \langle \delta_{2+} X, \Delta x_2 \rangle \geq 0 \tag{3.9.2}$$

Si la desigualdad estricta funciona en 3.9.2 siempre que $x_+ \neq x_-$, se dice que X es silla estricta.

Existe una estrecha relación entre los conceptos de funcional silla y de funcionales convexas y cóncavas.

3.9.2 Teorema

Una condición necesaria y suficiente para que la funcional X sea silla con convexidad en E_1 y concavidad en E_2 es que para cada $x_2 \varepsilon E_2$, fija, X sea convexa en E_1 y para cada $x_1 \varepsilon E_1$, fija, X sea cóncava en E_2.

La afirmación anterior conserva su validez si los conceptos de *silla, convexidad* y *concavidad* se sustituyen por los de *estrictamente silla, convexidad estricta* y *concavidad estricta.*

Demostración. Supóngase que X es silla. Entonces, si x_2 se mantiene fija, empleando una notación conveniente, la desigualdad 3.9.2 se reduce a

$$X(x_{1+}, x_{2-}) - X(x_{1-}, x_{2-}) - <\delta_{1-}X, \Delta x_1> \geq 0 \qquad (3.9.3a)$$

y, en forma análoga, si x_1 se mantiene fija

$$X(x_{1+}, x_{2+}) - X(x_{1+}, x_{2-}) - <\delta_{2+}X, \Delta x_2> \geq 0 \qquad (3.9.3b)$$

Pero la desigualdad 3.9.3a es la definición de convexidad, mientras que la 3.9.3b es la definición alternativa de concavidad contenida en el teorema 3.5.3

Por otra parte, supóngase que 3.9.3a se cumple siempre que se mantenga x_2-fija y que 3.9.3b se cumple siempre que permanezca x_{1+} fija. Entonces, de la suma de ambas relaciones se obtiene 3.9.2.

Finalmente, es claro que los mismos argumentos pueden desarrollarse si la igualdad se elimina de todas esas relaciones.

3.10 *Principios duales*

En esta parte del trabajo se generalizan los resultados del subcap 3.8 a operadores no lineales.

Con la notación empleada en el subcap 3.8 y bajo la hipótesis de que la funcional X posee derivadas que son bidimensionalmente continuas, se considera el sistema de ecs 3.10.1:

$$\delta_1 X(x_1, x_2) = 0 \qquad (3.10.1a)$$

$$\delta_2 X(x_1, x_2) = 0 \qquad (3.10.1b)$$

3.10.1 Teorema

Sean A,B dos subconjuntos de $\hat{E}$:

$$A = x_a = (x_{1a}, x_{2a}) \mid x_a \text{ satisface } 3.10.1a \qquad (3.10.2a)$$

$$B = x_b = (x_{1b}, x_{2b}) \mid x_b \text{ satisface } 3.10.1b \qquad (3.10.2b)$$

Supóngase que la funcional $X(x_1, x_2)$ es silla convexa con respecto a E_1 y cóncava con respecto a E_2, entonces

i) Para cada $x_a \varepsilon A$ y $x_b \varepsilon B$, se tiene

$$X(x_a) \leq X(x_b) \qquad (3.10.3)$$

ii) Si se supone además que existe $x_0 = (x_{10}, x_{20}) \varepsilon \hat{E}$ que satisface el sistema 3.10.1, se tiene

α) El valor máximo en el conjunto A de la funcional X, se alcanza en x_0

β) El valor mínimo en el conjunto B de la funcional X se alcanza en x_0

γ) El valor máximo de X en A y su valor mínimo en B coinciden.

Demostración. Es la misma del teorema 3.8.1

3.10.2 Corolario

Si en el teorema anterior se supone que además X es estrictamente silla, entonces

i) Para cada $x_a \varepsilon A$ y $x_b \varepsilon B$, se tiene

$$X(x_a) < X(x_b) \tag{3.10.4}$$

siempre que $x_a \neq x_b$.

ii) Si la solución del sistema de ecuaciones 3.10.1 existe, es única.

iii) Si el valor máximo de X en A y el valor mínimo de X en B se alcanzan y ambos son iguales, entonces el máximo en A y el mínimo en B son la solución del sistema de ecs 3.10.1.

Demostración. Es la misma que la del corolario 3.8.2.

3.11 *Condiciones locales para ser funcional silla*

Las condiciones locales para convexidad y concavidad presentadas en el subcap 3.7 implican, en vista del teorema 3.9.2, las condiciones locales correspondientes para que una funcional sea silla. Por sencillez se tomará

$$\hat{E} = D = D_1 + D_2$$

3.11.1 Teorema

Sea

$$X : D \to R^1 \tag{3.11.1}$$

tal que

i) X es diferenciable

ii) El operador

$$X'' : D \to D^{2*}$$

es bidimensionalmente continuo en D.

Entonces una condición necesaria y suficiente para que X sea silla convexa con respecto a D_1 y cóncava con respecto a D_2, es que para toda $x\varepsilon D$, $\frac{\partial^2 X}{\partial x_1^2}(x)\varepsilon D^{2*}$ sea no negativa y al mismo tiempo $\frac{\partial^2 X}{\partial x_2^2}(x)\varepsilon D^{2*}$ sea no positiva. Esta afirmación sigue siendo cierta si *silla* se remplaza por *estrictamente silla*, *no negativa* por *positiva* en D_1 y *no positiva* por *negativa* en D_2.

Demostración. En vista del teorema 3.9.2, es una consecuencia inmediata del teorema 3.7.1 y de la observación al final del subcap 3.7.

3.12 Principios de Hamilton

En esta sección se derivan, a partir de los principios duales del subcap 3.10, los principios de Hamilton generalizados que han sintetizado Noble y Sewell (ref 11). Por sencillez se tomará nuevamente $\hat{E} = D = D_1 + D_2$.

Considérese un operador lineal

$$\hat{T} : D \to D^* \tag{3.12.1}$$

Defínase la funcional

$$\Psi : D \to R^1 \tag{3.12.2}$$

para cada $x = (x_1, x_2) \varepsilon D$ por

$$\Psi(x) = \langle \hat{T}x_1 , x_2 \rangle = \langle \hat{T}^* x_2 , x_1 \rangle \tag{3.12.3}$$

Dada una funcional X, defínase

$$\hat{X} : D \to R^1 \tag{3.12.4}$$

para cada $x = (x_1, x_2) \varepsilon D$, por

$$\hat{X}(x_1, x_2) = X(x_1, x_2) - \Psi(x) \tag{3.12.5}$$

Los principios duales del subcap 3.10 son aplicables a la funcional X, y cuando se procede de esta manera se obtienen los principios de Hamilton generalizados presentados por Noble y Sewell (ref 11). Los resultados correspondientes están contenidos en los siguientes dos teoremas.

3.12.1 Teorema

Existe un operador lineal

$$T : D \to D^* \qquad (3.12.6)$$

tal que

$$\frac{\partial \hat{X}}{\partial x_1}(x) = \frac{\partial X}{\partial x_1}(x) - T^* x_2 \qquad (3.12.7a)$$

$$\frac{\partial \hat{X}}{\partial x_2}(x) = \frac{\partial X}{\partial x_2}(x) - T x_1 \qquad (3.12.7b)$$

Demostración. Defínase

$$T : D \to D^* \qquad (3.12.8)$$

para cada $x = (x_1, x_2) \varepsilon D$ y $y = (y_1, y_2) \varepsilon D$ por la ecuación

$$\langle T(x), y\rangle = \langle \hat{T} x_1, y_2\rangle \qquad (3.12.9)$$

Evidentemente, T es lineal y su adjunto T^* está dado por

$$\langle T^* x, y\rangle = \langle \hat{T} y_1, x_2\rangle \qquad (3.12.10)$$

Es fácil probar que

a) $\frac{\partial\Psi}{\partial x}(x)\varepsilon D^{*}$ es lineal en x

b) Para toda $x_2 \varepsilon D_2$

$$\frac{\partial\Psi}{\partial x_1}(x_2) = 0 \qquad (3.12.11)$$

c) Para toda $y_1 \varepsilon D_1$ y toda $x \varepsilon D$, se tiene

$$<\frac{\partial\Psi}{\partial x_1}(x), y_1> = 0 \qquad (3.12.12)$$

d) Finalmente

$$\frac{\partial\Psi}{\partial x_1}(x) = T^{*}x = T^{*}x_2 \qquad (3.12.13a)$$

$$\frac{\partial\Psi}{\partial x_2}(x) = Tx = Tx_1 \qquad (3.12.13b)$$

En vista de estos resultados, el teorema es evidente.

Obsérvese que

$$Tx_2 = 0 \qquad (3.12.14)$$

siempre que $x_2 \varepsilon D_2$ y que

$$<Tx, y_1> = 0$$

siempre que $y_1 \varepsilon D_1$. Por lo mismo, el operador T puede pensarse como un operador

$$T : D_1 \to D_2^* \tag{3.12.15a}$$

Similarmente

$$T^* : D_2 \to D_1^* \tag{3.12.15b}$$

que es el punto de vista adoptado por Rall (ref 14).

Al aplicar los resultados del subcap 3.10 a la funcional $\hat{X}$, es conveniente tener presente la siguiente propiedad.

3.12.2 Teorema

La funcional $\hat{X}$ dada por 3.12.5 es silla con convexidad con respecto a D_1 y concavidad con respecto a D_2 si y solo si X lo es. La validez de esta afirmación se mantiene si en ella *silla* se remplaza por *estrictamente silla*.

Demostración. Porque para cualquier $x_+ = (x_{1+} , x_{2+})$, $x_- = (x_{1-} , x_{2-}) \varepsilon D$, se tiene

$$\Delta\Psi - <\frac{\partial\Psi}{\partial x_{1-}} , \Delta x_1> - <\frac{\partial\Psi}{\partial x_{2+}} , \Delta x_2> = 0 \tag{3.12.16}$$

4. RECONOCIMIENTO

Se agradece la colaboración de los investigadores del Instituto de Ingeniería y del CIMAS que intervinieron en diversos aspectos del desarrollo de este trabajo, en particular a Jacobo Bielak, quien participó en varias investigaciones colaterales que se publican por separado, así como a Hugo Martínez, cuyas ideas fueron un valioso auxiliar en diversas etapas de la formalización de los resultados que se presentan y a Martha Cerrilla de Canudas por la elaboración mecanográfica del borrador.

5. REFERENCIAS

1. O. C. Zienkiewicz, *The finite element method in engineering science*, McGraw-Hill Book Co., Londres (1971)

2. J. T. Oden, *Finite elements of non-linear continua*, McGraw-Hill Book Co., Nueva York (1972)

3. M. Z. Nashed, *Differentiability and related properties of nonlinear operators: some aspects of the role of differentials in nonlinear functional analysis*, en L. B. Rall (ed.), Non-linear functional analysis and applications, Academic Press, Nueva York (1971)

4. E. Tonti, *On the variational formulation for linear initial value problems* Annali di Matematica pura ed applicata, Vol 95 (1973), pp 331-360

5. I. Herrera y J. Bielak, *A simplified version of Gurtin's variational principles*, Arch. Rational Mech. Anal. (1974)

6. B. Noble, *Complementary variational principles for boundary-value problems: I. Basic Principles*, Informe No 473, Mathematical Research Center, Universidad de Wisconsin (1964)

7. B. Noble, *Complementary variational principles: II. Nonlinear networks*, Informe No 643, Mathematical Research Center, Universidad de Wisconsin (1966)

8. M. J. Sewell, *On dual approximation principles and optimization in continuum mechanics*, Phil. Trans. Roy. So., Vol 265, Londres (1969), pp 319-351

9. A. M. Arthurs, *Complementary variational principles*, Oxford University Press (1970)

10. P. D. Robinson, *Complementary variational principles*, en L. B. Rall (ed.), Nonlinear functional analysis and applications, Academic Press, Nueva York (1971), pp 507-576

11. B. Noble y M. J. Sewell, *On dual extremum principles in Applied Mathematics*, Journal of Mathematical Analysis and Applications, Vol 9 (1972), pp 123-193

12. I. Herrera, *A general formulation of variational principles*, Instituto de Ingeniería, UNAM, Informe E-10, México, D. F. (1974)

13. M. M. Vainberg, *Variational methods for the study of nonlinear operator equations*, Holden-Day, San Francisco (1964)

14. L. B. Rall, *On complementary variational principles*, Journal of Mathematical Analysis and Applications, Vol 14 (1966), pp 174-184