

Escuela Matemática de América Latina y el Caribe
2010

*Solución Numérica de EDP, Introducción al
Método de Elemento Finito*

Dr. L. Héctor Juárez V.

Villahermosa, Tabasco
02 al 13 de agosto de 2010

Índice

1. Problemas elípticos unidimensionales	2
1.1. Problema modelo (Sturm–Liouville)	2
1.2. Existencia y unicidad de soluciones. Comportamiento asintótico	4
1.3. Ejercicios	8
2. Formulación variacional del problema unidimensional.	8
2.1. Problema de minimización	8
2.2. Derivación del problema variacional de la ecuación diferencial	9
2.3. Aproximaciones de Ritz–Galerkin	10
2.4. Ejercicios	12
3. El método de elemento finito en el problema unidimensional	12
3.1. Construcción de las funciones base	13
3.2. Aproximación del problema variacional	15
3.3. Propiedades de la matriz de rigidez	16
3.4. Algunos cálculos prácticos	17
3.5. Ejercicios	19
4. Problema multidimensional elíptico.	19
4.1. Problema modelo	19
4.2. Formulación variacional	21

4.3. Problema de minimización asociado	22
4.4. El problema de Neumann puro	23
4.5. Ejercicios	24
5. El método de elemento finito en el problema bidimensional	24
5.1. Discretización del dominio	25
5.2. Construcción de las funciones base	28
5.3. Aproximación del problema variacional	30
5.4. Propiedades de la matriz de rigidez	32
5.5. Aspectos prácticos	33
5.6. Ejercicios	37
6. Extensión a problemas parabólicos	37
6.1. Problema modelo	38
6.2. Formulación variacional	38
6.3. Aproximación por medio del método de elemento finito	39
6.4. Solución del sistema de EDO	40
6.5. Análisis de estabilidad de los métodos	42
6.6. Análisis de estabilidad por medio de métodos de energía	44
6.7. Ejercicios	44

1. Problemas elípticos unidimensionales

Objetivo: Introducir el método de elemento finito y las matemáticas asociadas al método. En una dimensión la construcción de los elementos es más simple y natural, además la manipulación matemática es directa y simple. Por lo tanto, comenzaremos con el caso de problemas unidimensionales y después extendemos los resultados a problemas en dos dimensiones.

1.1. Problema modelo (Sturm–Liouville)

Calcular $u = u(x)$ tal que

$$-\frac{d}{dx} \left(p(x) \frac{du}{dx} \right) + q(x) u = f(x), \quad 0 < x < L, \quad (1.1)$$

$$u(0) = 0, \quad (1.2)$$

$$\frac{du}{dx}(L) = 0. \quad (1.3)$$

A la condición (1.2) se le denomina condición de frontera de tipo Dirichlet ó esencial, mientras que a la condición (1.3) se le denomina condición de frontera de tipo Neumann ó natural. Podemos considerar $q(x) = 0$ en (1.1), sin pérdida de generalidad.

Notación. Con el objeto de simplificar la notación escribiremos indistintamente

$$u' = u_x \equiv \frac{du}{dx}.$$

A continuación presentamos dos ejemplos de aplicación en donde aparece un problema de Sturm–Liouville del tipo (1.1)–(1.3).

Ejemplo 1. Barra elástica sujeta a una carga. Consideremos una barra unidimensional de longitud L sujeta a una carga tangencial de intensidad f , como se ilustra en la Figura 1.

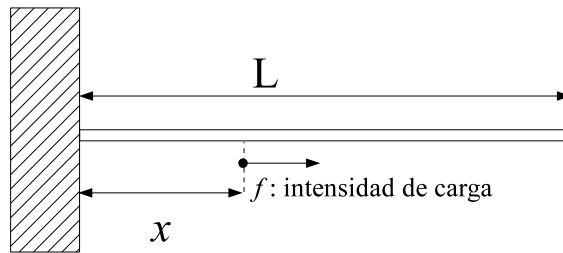


Figura 1: Barra elástica de longitud L sometida a una intensidad de carga f .

Denotamos por:

- $\sigma(x)$, la tracción en x debido a la intensidad de carga f .
- $u(x)$, el desplazamiento en x debido a la intensidad de carga f .
- $k(x)$, el módulo de elasticidad. Si $k(x) = \text{constante}$, el material es homogéneo.

Suponiendo que el material de la barra es elástico lineal y que los desplazamientos $u(x)$ son pequeños, se tiene

$$\begin{aligned} \sigma(x) &= k(x) u'(x), && \text{(Ley de Hooke)} \\ -\sigma'(x) &= f(x), && \text{(Ecuación de equilibrio)} \\ u(0) &= u'(L) = 0. && \text{(Condición de frontera)} \end{aligned}$$

Eliminando σ , obtenemos la ecuación diferencial

$$-\frac{d}{dx} \left(k(x) \frac{du}{dx} \right) = f(x), \quad 0 < x < L, \quad (1.4)$$

$$u(0) = u'(L) = 0. \quad (1.5)$$

Ejemplo 2. Conducción de calor en una barra. Consideremos el problema de la conducción de calor en una barra de longitud L , sujeta a una fuente de calor de intensidad f . Denotamos por:

- $u(x)$, la temperatura de la barra en la posición x debida a la fuente de calor f .
 - $q(x)$, la cantidad de calor en x debida a la fuente de calor f .
- $k(x)$, el coeficiente de conductividad térmica.

Suponiendo que el extremo derecho de la barra se aísla térmicamente y que el extremo izquierdo se mantiene a una temperatura constante cero, se tiene

$$\begin{aligned} -q(x) &= k(x) u'(x), && \text{(Ley de Fourier)} \\ q'(x) &= f(x), && \text{(Conservación de calor)} \\ u(0) &= u'(L) = 0. && \text{(Condiciones de frontera)} \end{aligned}$$

Eliminando q , obtenemos la ecuación diferencial

$$-\frac{d}{dx} \left(k(x) \frac{du}{dx} \right) = f(x), \quad 0 < x < L, \quad (1.6)$$

$$u(0) = u'(L) = 0. \quad (1.7)$$

1.2. Existencia y unicidad de soluciones. Comportamiento asintótico

Introducimos el operador

$$\mathcal{L} \doteq -\frac{d}{dx} \left(p(x) \frac{d}{dx} \right). \quad (1.8)$$

Estamos interesados en analizar el problema

$$\mathcal{L} u = f, \quad (1.9)$$

$$u(0) = u'(L) = 0. \quad (1.10)$$

Problema fundamental: ¿Para qué clase de funciones f hay una solución u ?

En primer término intentaremos obtener los espacios de funciones adecuados o *espacios de funciones admisibles*. El conocer los espacios de funciones admisibles es de importancia fundamental para estudiar los *principios variacionales* así como los *métodos de aproximación*. Una posible elección específica es admitir funciones f con “*energía finita*”, es decir que satisfacen

$$\int_0^L f(x)^2 dx < \infty. \quad (f, \text{ potencial de carga}) \quad (1.11)$$

Por ejemplo, cualquier función suave por tramos es admisible. La función δ de Dirac no se admite (descartamos las cargas puntuales de intensidad infinita). Denotamos por $L_2(\Omega)$ a este conjunto de funciones admisibles, donde $\Omega = (0, L)$. Para que la ecuación diferencial (1.9)–(1.10) tenga solución se requiere que

1. $p(x) \geq p_0 > 0$.
2. u, u' y u'' tengan "energía finita".

Denotamos por $H^2(\Omega)$ a este espacio de funciones. Entonces la solución de la ecuación diferencial se encuentra en el espacio

$$V = \{v \in H^2(\Omega) : v(0) = 0\}. \quad (1.12)$$

Teorema. El operador $\mathcal{L} : V \rightarrow L_2(\Omega)$ transforma biunívocamente V sobre $L_2(\Omega)$. Por lo tanto, para cada $f \in L_2(\Omega)$, existe una única solución $u \in V$. Además la solución depende continuamente de los datos, es decir pequeñas perturbaciones en el dato f producen pequeñas perturbaciones en la solución u .

Para explicar la dependencia continua entre f y u , observamos que estas funciones se deben medir bajo normas distintas:

$$\|f\|_0 = \left(\int_0^L f(x)^2 dx \right)^{1/2}, \quad (1.13)$$

y

$$\|u\|_2 = \left(\int_0^L [u(x)^2 + u'(x)^2 + u''(x)^2] dx \right)^{1/2}. \quad (1.14)$$

Entonces la **dependencia continua** de la solución respecto de los datos puede expresarse en forma cuantitativa de la siguiente manera

$$\|u\|_2 \leq C \|f\|_0. \quad (1.15)$$

Unicidad de la solución. La estimaciones del tipo (1.15) constituyen uno de los aspectos fundamentales de la **teoría moderna** de las ecuaciones diferenciales. Por ejemplo, utilizando esta estimación podemos probar *unicidad* de soluciones en forma inmediata, ya que si suponemos que u_1 y u_2 son soluciones de la ecuación diferencial, entonces $u = u_1 - u_2$ es solución de

$$\begin{aligned} \mathcal{L}u &= 0, \\ u(0) &= u'(1) = 0, \end{aligned}$$

y en consecuencia

$$\|u\| \leq C \|0\|_0 = 0.$$

Es decir $u_1 = u_2$.

Funciones propias del operador diferencial. Para continuar con el análisis de las soluciones de la ecuación, primero construimos una solución. La clave consiste en encontrar las funciones propias y los valores propios del operador \mathcal{L} . Para simplificar la exposición haremos la suposición de que $p(x) = p$ es constante. En este caso las funciones propias son

$$u_n(x) = \sqrt{\frac{2}{L}} \operatorname{sen} \left(\frac{\pi}{L} \left(n - \frac{1}{2} \right) x \right), \quad n = 1, 2, \dots, \infty. \quad (1.16)$$

Es fácil verificar que estas funciones satisfacen (1.10). Por otro lado, los valores propios λ_n , asociados a las funciones u_n , satisfacen

$$\mathcal{L}u_n = \lambda_n u_n,$$

o bien

$$-p u_n''(x) = \lambda_n u_n(x).$$

De (1.16) y de esta última condición es fácil obtener

$$\lambda_n = p \frac{\pi^2}{L^2} \left(n - \frac{1}{2} \right)^2. \quad (1.17)$$

Ortogonalidad de las funciones propias. Las funciones propias $\{u_n\}_{n=1}^{\infty}$ constituyen un conjunto ortonormal de funciones bajo el producto escalar usual en $L_2(\Omega)$. Se puede verificar que

$$\langle u_n, u_m \rangle = \int_0^L u_n(x) u_m(x) dx = 0, \quad \forall m, n \quad \text{con} \quad m \neq n,$$

y que

$$\langle u_n, u_n \rangle = 1, \quad \forall n.$$

Es decir

$$\langle u_n, u_m \rangle = \delta_{mn}. \quad (1.18)$$

Construcción de la solución. Para construir la solución de (1.9)–(1.10) expresamos el dato f como combinación lineal de las funciones propias. Es decir, escribimos

$$f(x) = \sum_{n=1}^{\infty} a_n u_n(x) = \sum_{n=1}^{\infty} a_n \sqrt{\frac{2}{L}} \operatorname{sen} \left(\frac{\pi}{L} \left(n - \frac{1}{2} \right) x \right), \quad (1.19)$$

para algunas constantes a_n , $n = 1, \dots, \infty$. Estas constantes deben satisfacer

$$\|f\|_0^2 = \int_0^L f(x)^2 dx = \sum_{n=1}^{\infty} a_n^2 < \infty, \quad (1.20)$$

para que $f \in L_2(\Omega)$.

Nota. Aparentemente hay una paradoja en la definición de f , puesto que uno podría pensar que dado que $u_n(0) = u_n'(L) = 0$, para toda $n = 1, 2, \dots$, entonces, en (1.19), $f(0) = f'(L) = 0$. Sin embargo, independientemente de que f satisfaga condiciones de frontera espurias o no, su expresión (1.19) es solamente válida en el sentido de mínimos cuadrados. Es decir

$$\int_0^L \left[f(x) - \sum_{n=1}^N a_n u_n(x) \right]^2 dx \rightarrow 0 \quad \text{cuando } N \rightarrow \infty. \quad (1.21)$$

Podemos pensar que las condiciones “a la frontera” de f , heredadas de las funciones propias u_n , son inestables y desaparecen tan pronto como $N \rightarrow \infty$.

Tomando en cuenta las anteriores consideraciones, podemos construir la solución del problema modelo (1.9)–(1.10) mediante la suma infinita

$$u(x) = \sum_{n=1}^{\infty} \frac{a_n}{\lambda_n} u_n(x) = \frac{L^2}{\pi^2} \sum_{n=1}^{\infty} \frac{a_n u_n(x)}{p(n-1/2)^2}, \quad (1.22)$$

en donde

$$a_n = \int_0^L f(x) u_n(x) dx. \quad (1.23)$$

Se puede verificar directamente (a partir de (1.22)) que

$$\|u\|_2^2 \leq \frac{\left[1 + \left(\frac{2L}{\pi}\right)^2 + \left(\frac{2L}{\pi}\right)^4 \right]}{p^2} \|f\|_0^2. \quad (1.24)$$

Por lo tanto, se satisface (1.15) con

$$c = \frac{\sqrt{1 + \left(\frac{2L}{\pi}\right)^2 + \left(\frac{2L}{\pi}\right)^4}}{p}.$$

Pregunta. ¿Porqué entonces $u(x)$ satisface las condiciones de frontera (1.10), mientras que $f(x)$ no las satisface, siendo que ambas tienen expansiones en términos de las funciones propias $u_n(x)$?

Respuesta. Porque

$$u_N(x) = \frac{L^2}{\pi^2} \sum_{n=1}^N \frac{a_n u_n(x)}{p(n-1/2)^2}$$

converge en sentido más fuerte. Es decir

$$\|u - u_N(x)\|_2 \rightarrow 0 \quad \text{cuando } N \rightarrow \infty,$$

mientras que $f_N(x) = \sum_{n=1}^N a_n u_n(x)$ no converge a $f(x)$ en la norma $\|\cdot\|_2$, sino sólo en la norma $\|\cdot\|_0$, como ya hemos comentado anteriormente.

1.3. Ejercicios

Ejercicio 1.1. Verificar que las funciones propias $\{u_n\}$, (1.16), y los valores propios $\{\lambda_n\}$, (1.17), satisfacen $\mathcal{L} u_n = \lambda_n u_n$.

Ejercicio 1.2. Verificar la ortonormalidad, (1.18), de las funciones propias $\{u_n\}$.

Ejercicio 1.3. Verificar que (1.22) es solución del problema (1.9)–(1.10).

Ejercicio 1.4. Verificar la desigualdad (1.24).

2. Formulación variacional del problema unidimensional.

El problema (1.9)–(1.10) está relacionado con la minimización del siguiente funcional cuadrático (también llamado *funcional de energía*).

$$J(v) = \frac{1}{2} \int_0^L p(v_x)^2 dx - \int_0^L f v dx. \quad (2.1)$$

El primer término del lado derecho describe la energía potencial total asociada al “desplazamiento” v . El segundo término está relacionado con el *potencial de carga* f .

2.1. Problema de minimización

$J(v)$ tiene un mínimo u sólo si la primera derivada (o primera variación) de J es cero en u . Es decir $J(v)$ tiene un mínimo u si

$$\frac{d}{d\epsilon} J(u + \epsilon v) |_{\epsilon=0} = \lim_{\epsilon \rightarrow 0} \frac{J(u + \epsilon v) - J(u)}{\epsilon} = 0, \quad \forall v \in V. \quad (2.2)$$

Obsérvese que

$$\begin{aligned} J(u + \epsilon v) &= \frac{1}{2} \int_0^L p(u + \epsilon v)_x^2 dx - \int_0^L f(u + \epsilon v) dx \\ &= \frac{1}{2} \int_0^L p u_x^2 dx + \epsilon \int_0^L p u_x v_x dx + \frac{\epsilon^2}{2} \int_0^L p v_x^2 dx - \int_0^L f u dx - \epsilon \int_0^L f v dx \\ &= J(u) + \epsilon \left\{ \int_0^L p u_x v_x dx - \int_0^L f v dx \right\} + \frac{\epsilon^2}{2} \int_0^L p v_x^2 dx. \end{aligned}$$

Entonces

$$\lim_{\epsilon \rightarrow 0} \frac{J(u + \epsilon v) - J(u)}{\epsilon} = \int_0^L p u_x v_x dx - \int_0^L f v dx, \quad \forall v \in V.$$

Por lo tanto, $u \in V$ es un mínimo de $J(v)$ sí

$$\int_0^L p u_x v_x dx = \int_0^L f v dx, \quad \forall v \in V. \quad (2.3)$$

A este último problema se le conoce como el **problema variacional** asociado a la ecuación diferencial, y su solución u es la función en la cual la variación del funcional J es mínima. El problema de minimización del funcional (2.1) corresponde al **principio de mínima energía potencial** en mecánica, mientras que el problema variacional (2.3) corresponde al **principio de trabajo virtual**.

2.2. Derivación del problema variacional de la ecuación diferencial

Es posible derivar también el problema variacional (2.3) directamente de la ecuación diferencial (1.9)–(1.10). El procedimiento consiste en escoger una *función de prueba* v , conocida también como *función de desplazamiento virtual*. Multiplicando la ecuación diferencial (1.9) por esta función e integrando, se obtiene

$$-\int_0^L (p u_x)_x v dx = \int_0^L f v dx, \quad \forall v \in V,$$

y al hacer integración por partes en el lado izquierdo, se obtiene

$$-p u_x v \Big|_{x=0}^{x=L} + \int_0^L p u_x v_x dx = \int_0^L f v dx, \quad \forall v \in V.$$

El término de frontera desaparece debido a que $u_x(L) = 0$ y $v(0) = 0$, obteniendo el problema variacional (2.3). A este problema variacional también se le conoce como *forma débil o de Galerkin* de la ecuación diferencial. La equivalencia de ecuaciones diferenciales con problemas variacionales es básica para la elección del esquema computacional. Por ejemplo,

- La ecuación diferencial (1.9) puede aproximarse por medio de diferencias finitas.
- El problema variacional (2.3) se puede aproximar por medio de un método de Galerkin, como veremos un poco más adelante.

En muchas aplicaciones el principio físico primario es el variacional. En tales aplicaciones el problema fundamental es de minimización y es el que debe ser aproximado. Obsérvese que en el problema de minimización del funcional (2.1) o en el problema variacional (2.3) no intervienen derivadas segundas de las funciones, como sucede en el caso de la ecuación diferencial (1.9). En tales casos solo se requiere que u_x tenga energía finita, y no es necesario que $u_{xx} \in L_2(\Omega)$.

Obsérvese que sí $p(x)$ es acotada en $0 \leq x \leq L$, y si además tanto u como v se escogen en el espacio

$$V = \{v \in H^1(\Omega) : v(0) = 0\}, \quad (2.4)$$

entonces

$$\left| \int_0^L p u_x v_x dx \right| \leq P \left| \int_0^L u_x v_x dx \right| \leq P \|u_x\|_0 \|v_x\|_0 < \infty,$$

donde $P = \max_{0 \leq x \leq L} |p(x)|$. Por lo tanto el problema variacional (2.3) se define en forma natural sobre el espacio de funciones (2.4). Es por esta razón que al problema variacional también se le conoce como el **problema débil** asociado a la ecuación diferencial, pues la función prueba y la solución se buscan en un espacio de funciones (2.4), el cual es más amplio pero con funciones menos suaves o regulares que el espacio de funciones (1.12).

Esta última observación tiene consecuencias prácticas (computacionales) importantes, ya que podemos escoger funciones que sean continuas, pero lineales por tramos (las cuales ciertamente pertenecen a $H_0^1(\Omega)$), con el objeto de buscar soluciones del problema variacional, en lugar de buscar funciones en $H^2(\Omega)$. Ciertos tipos de funciones continuas lineales por tramos, como las funciones polinomiales a trozos, son fáciles de construir, y su primera derivada tiene energía finita.

Para el espacio de Hilbert (2.4), el producto escalar y su norma correspondiente se definen por

$$\langle u, v \rangle_1 = \int_0^L [u(x)v(x) + u'(x)v'(x)] dx, \quad (2.5)$$

$$\|v\|_1 = \left(\int_0^L [v(x)^2 + v'(x)^2] dx \right)^{1/2}. \quad (2.6)$$

2.3. Aproximaciones de Ritz–Galerkin

El problema variacional consiste en encontrar

$$u \in V = \{v \in H^1(\Omega) : v(0) = 0\},$$

tal que

$$\int_0^L p u_x v_x dx = \int_0^L f v dx, \quad \forall v \in V. \quad (2.7)$$

Dos propiedades del espacio de funciones V que juegan un papel crucial en las aproximaciones del problema variacional (2.7) son:

1. V es un **espacio vectorial lineal** de funciones.
2. V es un espacio de **dimensión infinita**.

Por ejemplo, las funciones propias $u_n(x)$, $n = 1, 2, \dots$, definidas en (1.16), pertenecen a V y generan V . En realidad, no es necesario tener funciones tan suaves como las trigonométricas. Es suficiente con cualquier conjunto de funciones base $\{\varphi_n\}_{n=1}^\infty$ con una derivada generalizada en $\Omega = (0, L)$ y que se anulen en $x = 0$. Estas funciones, para que formen una base de V ,

deben ser linealmente independientes. Además, toda $v \in V$ se debe poder expresar como combinación lineal de la forma

$$v = \sum_{i=1}^{\infty} \alpha_i \varphi_i, \quad \alpha_i \in \mathbb{R}.$$

Si tomamos un número finito de términos en esta serie, digamos N , obtenemos una aproximación v_N a v , definida por

$$v_N(x) = \sum_{i=1}^N \alpha_i \varphi_i(x). \quad (2.8)$$

Las N funciones base $\{\varphi_1, \varphi_2, \dots, \varphi_N\}$ generan un subespacio N -dimensional de V que denotaremos por V_N . **El método de Galerkin** consiste en buscar una solución aproximada u_N del problema variacional:

Encontrar $u_N \in V_N$ tal que

$$\int_0^L p u_N' v' dx = \int_0^L f v dx, \quad \forall v \in V_N. \quad (2.9)$$

Se ha utilizado la notación $u_N' = du_N/dx$ y $v' = dv/dx$. Como V_N es un espacio de funciones de dimensión finita generado por $\{\varphi_i\}_{i=1}^N$, entonces el problema (2.9) es equivalente al siguiente problema:

Encontrar $u_N \in V_N$ tal que

$$\int_0^L p u_N' \varphi_i' dx = \int_0^L f \varphi_i dx, \quad i = 1, 2, \dots, N.$$

Si además $\alpha_1, \alpha_2, \dots, \alpha_N$ son las constantes tales que

$$u_N(x) = \sum_{j=1}^N \alpha_j \varphi_j(x), \quad (2.10)$$

entonces este problema es, a su vez, equivalente al problema:

Encontrar los coeficientes $\alpha_1, \alpha_2, \dots, \alpha_N$ tales que

$$\sum_{j=1}^N \left(\int_0^L p \varphi_i' \varphi_j' dx \right) \alpha_j = \int_0^L f \varphi_i dx, \quad i = 1, 2, \dots, N.$$

Este problema constituye un sistema de ecuaciones que podemos escribir en la forma

$$\mathbb{K} \alpha = \mathbf{F}, \quad (2.11)$$

en donde los coeficientes del vector α son las incógnitas $\{\alpha_i\}_{i=1}^N$, y los coeficientes de la matriz \mathbb{K} y el vector \mathbf{F} son

$$K_{ij} = \int_0^L p \varphi_i' \varphi_j' dx \quad \text{y} \quad F_i = \int_0^L f \varphi_i dx, \quad 1 \leq i, j \leq N,$$

respectivamente. A la matriz \mathbb{K} se le conoce como *matriz de rigidez* del problema, y al vector \mathbf{F} como el *vector de carga*. Estos nombres provienen de la mecánica de estructuras, en donde primero se utilizaron estos métodos de aproximación. La matriz \mathbb{K} es simétrica y, debido a que $p(x) \geq p_0 > 0$, la matriz también es definida positiva, lo cual implica que es invertible. Por lo tanto, el sistema de ecuaciones (2.11) tiene solución única.

La calidad de la aproximación está completamente determinada por la elección de las funciones base $\{\varphi_i\}_{i=1}^N$, y el número N de ellas. Una vez que estas funciones base se han escogido, la determinación de los coeficientes α_j (grados de libertad del problema aproximado) es un problema computacional, que consiste en la solución de un sistema de ecuaciones. Es posible verificar que la aproximación u_N es el elemento en V_N que mejor aproxima a u , pues

$$\langle u - u_N, v \rangle = 0, \quad \forall v \in V_N, \quad (2.12)$$

debido a la ortogonalidad de las funciones base φ_i .

2.4. Ejercicios

Ejercicio 2.1. Escribe el funcional de minimización $J(v)$ asociado con el problema (1.9)–(1.10). Escribe el correspondiente problema variacional.

Ejercicio 2.2. Demuestra que cualquier solución de la ecuación diferencial (1.9)–(1.10) es solución del problema variacional (2.7). Sin embargo, no toda solución del problema variacional es solución de la ecuación diferencial. Intenta explicar por qué e indica cuál es la condición adicional para que la solución del problema variacional también sea solución de la ecuación diferencial.

Ejercicio 2.3. Explica en que consiste el método de Galerkin para resolver el problema variacional (2.7). Demuestra que la matriz de rigidez \mathbb{K} es definida positiva si $p(x) \geq p_0 > 0$ para toda x .

Ejercicio 2.4. Demuestra la igualdad (2.12). Utiliza esa propiedad para demostrar que u_N es el elemento en V_N que mejor aproxima a u .

3. El método de elemento finito en el problema unidimensional

El *método de elemento finito* proporciona un técnica particular, además de sistemática, para construir funciones base $\{\varphi_i\}_{i=1}^N$ en la aproximación de Ritz–Galerkin de problemas con valores a la frontera. La idea principal detrás del método de elemento finito consiste en definir las funciones base φ_i como *funciones polinomiales por tramos*, sobre subregiones del dominio llamadas *elementos finitos*. En esta sección mostraremos la construcción del

elemento finito más simple para resolver problemas unidimensionales, el cual consiste de funciones polinomiales lineales por tramos (elemento de Lagrange de primer orden).

3.1. Construcción de las funciones base

Dividimos la región $\Omega = (0, L)$ en un número finito de elementos. En el caso unidimensional estos elementos son subintervalos $\Omega_i = (x_{i-1}, x_i)$, $i = 1, 2, \dots, N$. La Figura 2 ilustra el caso de cuatro elementos ($N=4$).

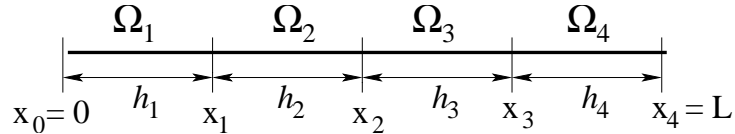


Figura 2: Malla unidimensional en el intervalo $[0, L]$.

Se definen $x_0 = 0$ y $x_4 = L$, y se introducen los puntos intermedios x_1, x_2, x_3 . Los puntos x_0, x_1, x_2, x_3, x_4 se denominan los **nodos** (o puntos nodales) de la malla, y los subintervalos $\Omega_1, \Omega_2, \Omega_3, \Omega_4$ se denominan los **elementos** de la malla. Sea $h_i = x_i - x_{i-1}$, $i = 1, \dots, 4$ la longitud del elemento Ω_i , entonces

nodos + elementos = malla de elemento finito.

El tamaño de la malla se define por $h = \max_{1 \leq i \leq 4} h_i$.

Las **funciones base** $\{\varphi_i\}_{i=1}^N$, que se utilizan en el método de elemento finito, usualmente se construyen tomando en cuenta los siguientes criterios:

1. Son funciones polinomiales simples definidas por tramos, es decir definidas elemento por elemento sobre la malla.
2. Son funciones que pertenecen a $H^1(\Omega)$, es decir son funciones cuadrado integrables con derivada generalizada también cuadrado integrable.
3. Son funciones que se escogen de tal manera que los coeficientes α_i en (2.10) sean iguales a $u_N(x_i)$ para $1 \leq i \leq N$.

Las **funciones sombrero**, (“*hat functions*” por su nombre en inglés, ó también conocidas como “*chapeau functions*”), satisfacen estos criterios. Estas funciones son funciones lineales (polinomios de grado 1) en cada elemento y se definen de la siguiente manera

$$\varphi_i(x) = \begin{cases} \frac{x - x_{i-1}}{h_i} & \text{si } x_{i-1} \leq x \leq x_i, \\ \frac{x_{i+1} - x}{h_{i+1}} & \text{si } x_i \leq x \leq x_{i+1}, \\ 0 & \text{en otro caso,} \end{cases} \quad (3.1)$$

para $i = 1, 2, \dots, N - 1$. Para $i = 0$ e $i = N$ se definen por

$$\varphi_0(x) = \begin{cases} \frac{x_1 - x}{h_1} & \text{si } x_0 \leq x \leq x_1, \\ 0 & \text{en otro caso.} \end{cases} \quad (3.2)$$

$$\varphi_N(x) = \begin{cases} \frac{x - x_{N-1}}{h_N} & \text{si } x_{N-1} \leq x \leq x_N, \\ 0 & \text{en otro caso.} \end{cases} \quad (3.3)$$

Estas funciones se ilustran en la Figura 3, para el caso $N = 4$. Claramente estas funciones

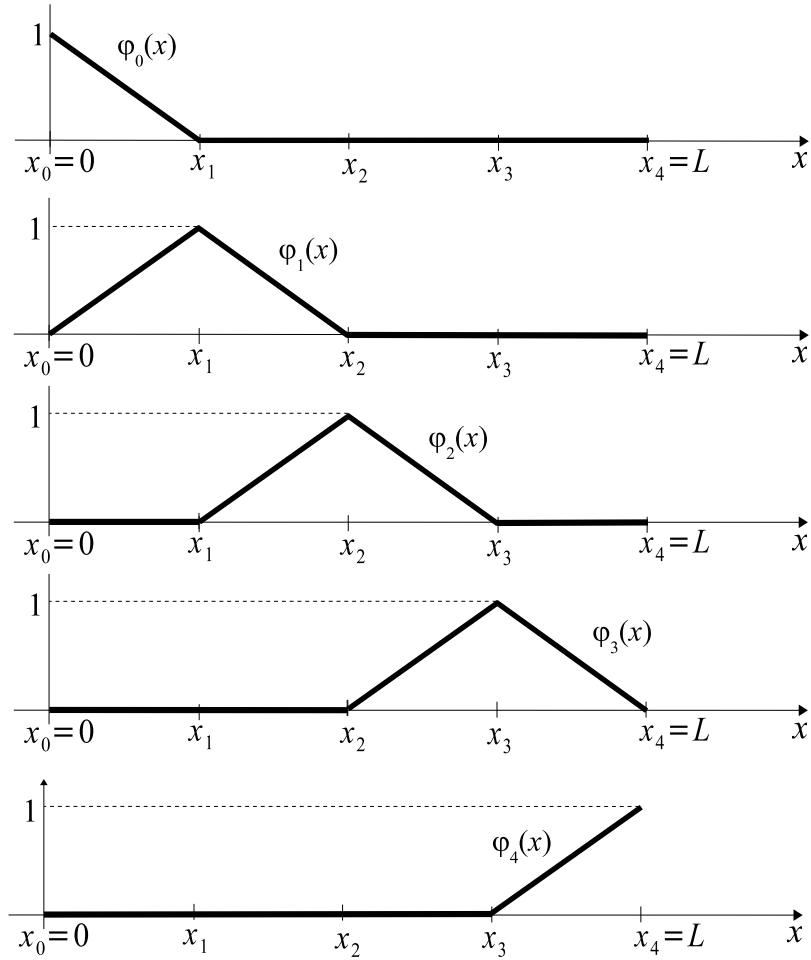


Figura 3: Funciones sombrero

satisfacen el primer criterio. Para verificar que satisfacen el segundo criterio debemos verificar que $\varphi_i(x)$ y $\varphi_i'(x)$ son cuadrado integrables. Esto se muestra a continuación

$$\int_0^L [\varphi_i(x)]^2 dx = \frac{1}{3} (h_i + h_{i+1}) < \infty, \quad i = 1, \dots, N - 1. \quad (3.4)$$

Para los casos $i = 0$ e $i = N$ se obtiene solamente $h_1/3$ y $h_N/3$, respectivamente. Por otro lado, para $1 \leq i \leq N - 1$, la derivada de φ_i es

$$\varphi_i'(x) = \begin{cases} \frac{1}{h_i} & \text{si } x_{i-1} < x < x_i, \\ -\frac{1}{h_{i+1}} & \text{si } x_i < x < x_{i+1}, \\ 0 & \text{si } x < x_{i-1} \text{ ó } x > x_{i+1}. \end{cases}$$

Obsérvese que para cada φ_i su derivada, en sentido clásico, no está definida en los nodos, pero si son cuadrado integrales, pues

$$\int_0^L [\varphi_i'(x)]^2 dx = \int_{x_{i-1}}^{x_i} \left(\frac{1}{h_i}\right)^2 dx + \int_{x_i}^{x_{i+1}} \left(\frac{1}{h_{i+1}}\right)^2 dx = \frac{1}{h_i} + \frac{1}{h_{i+1}} < \infty, \quad (3.5)$$

para $i = 1, \dots, N-1$. Para los casos $i = 0$ e $i = N$ se obtiene $1/h_1$ y $1/h_N$, respectivamente.

Se concluye que, dado que las integrales son finitas y que las funciones $\varphi_i(x)$, $i = 0, \dots, N$ son continuas, entonces ellas pertenecen al espacio de funciones $H^1(\Omega)$. Finalmente verificaremos el criterio 3. Por construcción de las funciones sombrero se tiene

$$\varphi_i(x_j) = \delta_{ij}. \quad (3.6)$$

Por lo tanto, en (2.10) se obtiene

$$u_N(x_j) = \sum_{i=1}^N \alpha_i \varphi_i(x_j) = \sum_{i=1}^N \alpha_i \delta_{ij} = \alpha_j.$$

3.2. Aproximación del problema variacional

Para aproximar el problema variacional (2.7), consideramos el espacio de dimensión finita V_h , que es subespacio de V , definido por

$$V_h = \{v_h \in \mathcal{C}(\Omega) : v_h|_{\Omega_e} \in P_1(\Omega_e), e = 1, \dots, N, v_h(0) = 0\}. \quad (3.7)$$

Éste es el conjunto de funciones continuas v_h definidas en $\Omega = (0, L)$ y que al restringirlas a cada elemento Ω_e , son polinomios de grado uno, además de que satisfacen la condición de tipo Dirichlet $v_h(0) = 0$. Este espacio es de dimensión finita y está generado por

$$V_h = \text{gen} \{\varphi_1, \varphi_2, \dots, \varphi_N\}.$$

Obsérvese que hemos excluido φ_0 del conjunto generador de V_h , para forzar la condición de tipo Dirichlet $v_h(0) = 0$. El problema variacional discreto asociado es:

Calcular $u_h \in V_h$ tal que

$$\int_0^L p \frac{du_h}{dx} \frac{dv_h}{dx} dx = \int_0^L f v_h dx, \quad \forall v_h \in V_h. \quad (3.8)$$

Este problema es equivalente al problema:

Calcular $u_h \in V_h$ tal que

$$\int_0^L p \frac{du_h}{dx} \frac{\varphi_i}{dx} dx = \int_0^L f \varphi_i dx, \quad \forall i = 1, 2, \dots, N.$$

Además u_h se puede escribir como

$$u_h(x) = \sum_{j=1}^N u_j \varphi_j(x), \quad (3.9)$$

donde u_j denota $u_h(x_j)$. Sustituyendo en la ecuación integral anterior, se obtiene que el problema variacional discreto es equivalente al siguiente problema aproximado:

Calcular u_1, u_2, \dots, u_N tales que

$$\sum_{j=1}^N \left(\int_0^L p \varphi_i' \varphi_j' dx \right) u_j = \int_0^L f \varphi_i(x) dx, \quad \forall i = 1, 2, \dots, N. \quad (3.10)$$

Éste es un sistema de ecuaciones

$$\mathbb{K} \mathbf{U} = \mathbf{F}, \quad (3.11)$$

en donde la matriz de rigidez \mathbb{K} y el vector de carga \mathbf{F} , del elemento lineal de Lagrange, se construyen con las funciones sombrero $\{\varphi_i\}_{i=1}^N$ (3.1)–(3.3). El vector solución \mathbf{U} del sistema (3.11) contiene los coeficientes u_j , $1 \leq j \leq N$, a partir de los cuales se obtiene la solución discreta u_h (3.9).

3.3. Propiedades de la matriz de rigidez

La matriz de rigidez \mathbb{K} en (3.11) posee algunas propiedades importantes desde el punto de vista computacional:

- 1. Sumabilidad.** Ésta es la propiedad más importante cuando se utiliza el método de elemento finito. La sumabilidad consiste en que los coeficientes K_{ij} pueden calcularse como una suma de contribuciones en cada elemento, es decir

$$K_{ij} = \int_0^L p \varphi_i' \varphi_j' dx = \sum_{e=1}^N \int_{\Omega_e} p \varphi_i' \varphi_j' dx = \sum_{e=1}^N K_{ij}^e, \quad (3.12)$$

donde

$$K_{ij}^e = \int_{\Omega_e} p \varphi_i' \varphi_j' dx.$$

En forma análoga se obtiene

$$F_i = \sum_{e=1}^N F_i^e, \quad \text{con} \quad F_i^e = \int_{\Omega_e} f \varphi_i dx. \quad (3.13)$$

La sumabilidad cobra una mayor importancia en problemas bidimensionales y tridimensionales, en donde realmente ayuda a simplificar el proceso de cálculo de \mathbb{K} y \mathbf{F} , así como el proceso de su almacenamiento en la memoria de la computadora.

- 2. Escasez de coeficientes no cero.** La mayoría de los coeficientes K_{ij} en la matriz de rigidez son cero, debido a que

$$\varphi_i'(x) \varphi_j'(x) \neq 0 \quad \text{solo si} \quad j = i - 1, i, i + 1.$$

Por lo tanto,

$$K_{ij} = \int_0^L p \varphi_i' \varphi_j' dx \neq 0,$$

sólo cuando $j = i - 1, i, i + 1$. Entonces \mathbb{K} es una matriz tridiagonal, de la forma

$$\mathbb{K} = \begin{pmatrix} K_{11} & K_{12} & & & 0 \\ K_{21} & K_{22} & K_{23} & & \\ & \ddots & \ddots & & \\ & & K_{N-1,N-2} & K_{N-1,N-1} & K_{N-1,N} \\ 0 & & & K_{N-1,N} & K_{N,N} \end{pmatrix}$$

El hecho de que la matriz de rigidez sea *rala* o “sparse”, es decir que la inmensa mayoría de sus coeficientes son cero, es una característica común de las matrices que aparecen en la discretización de problemas modelados por ecuaciones diferenciales cuando se utiliza el método de elemento finito. Otros métodos de discretización, como los métodos de diferencias finitas y volumen finito, también producen matrices ralas. También existen métodos de aproximación que producen matrices densas, como por ejemplo métodos de aproximación con funciones radiales (radial basis functions, por su nombre en inglés).

- 3. Simetría.** Como ya hemos visto anteriormente, la matriz \mathbb{K} es simétrica, es decir $K_{ij} = K_{ji}$. Esta propiedad no está asociada a las funciones base, ni con el método de aproximación utilizado, sino con el problema variacional de la ecuación diferencial. Es decir, el operador diferencial asociado al problema es autoadjunto.

3.4. Algunos cálculos prácticos

En esta sección calculamos los coeficientes (3.12) de la matriz de rigidez \mathbb{K} y los coeficientes (3.13) del vector de carga \mathbf{F} . Utilizando la propiedad de que $K_{ij} \neq 0$ sólo si $j = i - 1$, $j = i$ o $j = i + 1$, se obtiene

$$K_{i,i-1} = \int_0^L p(x) \varphi_i'(x) \varphi_{i-1}'(x) dx = -\frac{1}{h_i^2} \int_{x_{i-1}}^{x_i} p(x) dx, \quad (3.14)$$

para $i = 2, \dots, N$. Análogamente

$$K_{ii} = \int_0^L p(x) [\varphi_i'(x)]^2 dx = \frac{1}{h_i^2} \int_{x_{i-1}}^{x_i} p(x) dx + \frac{1}{h_{i+1}^2} \int_{x_i}^{x_{i+1}} p(x) dx, \quad (3.15)$$

para $i = 1, \dots, N$. Por simetría de \mathbb{K} y analogía con (3.14), se obtiene

$$K_{i,i+1} = -\frac{1}{h_{i+1}^2} \int_{x_i}^{x_{i+1}} p(x) dx, \quad (3.16)$$

para $i = 1, \dots, N-1$. Por otro lado

$$F_i = \int_0^L f(x) \varphi_i(x) dx = \int_{x_{i-1}}^{x_i} f(x) \varphi_i(x) dx + \int_{x_i}^{x_{i+1}} f(x) \varphi_i(x) dx, \quad (3.17)$$

para $i = 1, \dots, N$.

Las integrales resultantes en (3.14)–(3.17), cuando no se puedan calcular en forma exacta, se pueden aproximar mediante la regla del trapecio:

$$\int_a^b f(x) dx \approx \frac{b-a}{2} \{f(a) + f(b)\}.$$

Introduciendo las notaciones $p_i = p(x_i)$ y $f_i = f(x_i)$ y aplicando la regla del trapecio, obtenemos

$$K_{i,i-1} \approx -\frac{p_{i-1} + p_i}{2h_i} \quad i=2, \dots, N. \quad (3.18)$$

$$K_{ii} \approx \frac{p_{i-1} + p_i}{2h_i} + \frac{p_i + p_{i+1}}{2h_{i+1}} \quad i=1, \dots, N. \quad (3.19)$$

$$K_{i,i+1} \approx -\frac{p_i + p_{i+1}}{2h_{i+1}} \quad i=1, \dots, N-1. \quad (3.20)$$

$$F_i \approx \frac{h_i + h_{i+1}}{2} f_i \quad i=1, \dots, N. \quad (3.21)$$

El cálculo aproximado de las integrales no degrada el resultado final de la aproximación, debido a que el error de integración numérica es del mismo orden que el que resulta de la aproximación con el método de elemento finito. Se puede probar que este error de aproximación es $\mathcal{O}(h^2)$, en donde $h = \max_{1 \leq i \leq N} h_i$ es el tamaño de la malla.

Un caso muy frecuente se presenta cuando $p(x) = p_0 = \text{constante}$ y $h_i = h$, para toda $i = 1, \dots, N$ (malla uniforme). En este caso las fórmulas para el cálculo de los coeficientes de la matriz de rigidez y del vector de carga se simplifica, obteniendo

$$K_{i,i} = \frac{2p_0}{h}, \quad K_{i,i-1} = K_{i,i+1} = -\frac{p_0}{h}, \quad F_i = h f_i.$$

Por lo tanto, la matriz de rigidez y el vector de carga, en este caso especial, son de la forma

$$\mathbb{K} = \frac{p_0}{h} \begin{pmatrix} 2 & -1 & & & 0 \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & & \\ & & & -1 & 2 & -1 \\ 0 & & & -1 & 2 \end{pmatrix} \quad \text{y} \quad \mathbf{F} = h \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_{N-1}) \\ f(x_N) \end{bmatrix}$$

Las integrales también se pueden aproximar utilizando la regla del punto medio

$$\int_a^b f(x) dx \approx (b-a) f\left(\frac{a+b}{2}\right).$$

En este caso, los coeficientes son

$$\begin{aligned} K_{i,i-1} &= -\frac{1}{h_i} p(x_{i-1/2}), & K_{ii} &= \frac{1}{h_i} p(x_{i-1/2}) + \frac{1}{h_{i+1}} p(x_{i+1/2}), \\ K_{i,i+1} &= -\frac{1}{h_{i+1}} p(x_{i+1/2}), & F_i &= \frac{h_i}{2} f(x_{i-1/2}) + \frac{h_{i+1}}{2} f(x_{i+1/2}), \end{aligned}$$

en donde

$$x_{i-1/2} = \frac{x_{i-1} + x_i}{2}, \quad x_{i+1/2} = \frac{x_i + x_{i+1}}{2}.$$

3.5. Ejercicios

Ejercicio 3.1. El método de elemento finito es un método particular de tipo Galerkin para resolver ecuaciones diferenciales. Explica cuál es la idea principal detrás del método de elemento finito.

Ejercicio 3.2. Escribir las propiedades principales de las funciones sombrero.

Ejercicio 3.3. Verificar las propiedades (3.4), (3.5), (3.6).

Ejercicio 3.4. Indicar cuáles son las propiedades de la matriz de rigidez obtenida con el elemento lineal de Lagrange.

Ejercicio 3.5. Verificar las aproximaciones (3.18) a (3.21).

4. Problema multidimensional elíptico.

4.1. Problema modelo

Considérese una región Ω en \mathbb{R}^d , con $d = 2$ o 3 , la cual es abierta y conexa, y cuya frontera es la unión de dos porciones $\partial\Omega = \Gamma_0 \cup \Gamma_1$. La porción Γ_0 es la parte de la frontera

donde se imponen condiciones de frontera de tipo Dirichlet (también llamadas esenciales), y la porción Γ_1 es la parte de la frontera donde se imponen condiciones de frontera de tipo Neumann (también llamadas naturales), como se ilustra en la Figura 4. Denotamos por \mathbf{n} la normal unitaria exterior a la frontera $\partial\Omega$, y por $\mathbf{x} = (x_1, \dots, x_d)$ a cualquier punto en $\bar{\Omega} = \Omega \cup \partial\Omega$.

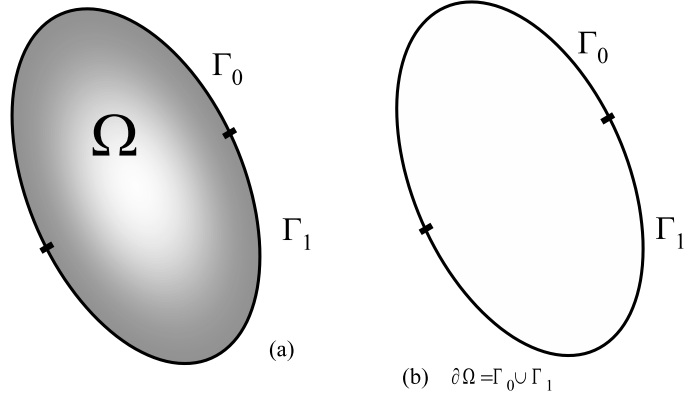


Figura 4: a) Dominio del problema, b) frontera que lo delimita

Consideremos el siguiente problema elíptico:

Encontrar $u = u(\mathbf{x})$ tal que

$$-\nabla \cdot (a \nabla u) = f \quad \text{en } \Omega, \quad (4.1)$$

sujeta a las condiciones de frontera

$$\text{Dirichlet:} \quad u = g_0 \quad \text{sobre } \Gamma_0, \quad (4.2)$$

$$\text{Neumann:} \quad a \frac{\partial u}{\partial \mathbf{n}} = g_1 \quad \text{sobre } \Gamma_1, \quad (4.3)$$

en donde a y f son funciones que dependen de \mathbf{x} y están definidas sobre Ω , mientras que g_0, g_1 están definidas sobre Γ_0 y Γ_1 , respectivamente. En (4.3), $\partial u / \partial \mathbf{n}$ indica la derivada direccional de u en la dirección de la normal exterior \mathbf{n} , es decir

$$\frac{\partial u}{\partial \mathbf{n}} = \nabla u \cdot \mathbf{n}. \quad (4.4)$$

En el problema elíptico (4.1)–(4.3) el *coeficiente de difusión* $a(\mathbf{x})$ se entenderá como una función escalar, pero en algunos problemas sobre medios anisotrópicos puede manejarse como una matriz $A(\mathbf{x}) = \{a_{ij}(\mathbf{x})\}_{1 \leq i, j \leq d}$. Supondremos que existen constantes a_0 y M tales que

$$0 < a_0 \leq a(\mathbf{x}) \leq M, \quad \forall \mathbf{x} \in \bar{\Omega}. \quad (4.5)$$

Obsérvese que para el caso en que el coeficiente de difusión sea una constante a (medio homogéneo), la ecuación (4.1) se reduce a la ecuación de Laplace

$$-a \nabla^2 u = f \quad \text{en } \Omega.$$

4.2. Formulación variacional

Al igual que el caso unidimensional, es conveniente introducir una función de prueba v suficientemente suave y que, de preferencia, pertenezca a un espacio de Hilbert, además de anularse en la frontera Dirichlet Γ_0 . Multiplicando la ecuación (4.1) por una función de este tipo e integrando sobre Ω , se obtiene

$$-\int_{\Omega} \nabla \cdot (a \nabla u) v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x}, \quad (4.6)$$

en donde $d\mathbf{x} = dx_1 \cdots dx_d$ (la integración es múltiple). Utilizando la identidad vectorial $\nabla \cdot (a \nabla u v) = \nabla \cdot (a \nabla u) v + a \nabla u \cdot \nabla v$, la ecuación se puede reescribir en la forma

$$\int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} + \int_{\Omega} \nabla \cdot (a \nabla u v) \, d\mathbf{x}.$$

Utilizando el teorema de la divergencia (Gauss)

$$\int_{\Omega} \nabla \cdot \mathbf{F} \, d\mathbf{x} = \int_{\partial\Omega} \mathbf{F} \cdot \mathbf{n} \, d\Gamma,$$

con $\mathbf{F} = a \nabla u v$, la ecuación anterior resulta en

$$\int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} + \int_{\partial\Omega} (a \nabla u v) \cdot \mathbf{n} \, d\Gamma.$$

La integral de frontera en esta última expresión se puede simplificar debido a que $\partial\Omega = \Gamma_0 \cup \Gamma_1$, la función de prueba satisface $v = 0$ sobre Γ_0 , y a que la solución u debe satisfacer $\partial u / \partial \mathbf{n} = g_1$ sobre Γ_1 . Por lo tanto,

$$\int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} + \int_{\Gamma_1} g_1 v \, d\Gamma.$$

Se observa que tanto u como v deben tener al menos una derivada generalizada para que las integral del lado izquierdo tenga sentido. Por lo tanto, basta con escoger como espacio de funciones prueba al conjunto

$$V_0 = \{v \in H^1(\Omega) : v = 0 \text{ sobre } \Gamma_0\}, \quad (4.7)$$

mientras que la solución la buscaremos en el espacio afín

$$V_g = \{v \in H^1(\Omega) : v = g_0 \text{ sobre } \Gamma_0\}. \quad (4.8)$$

Luego, el **problema variacional** asociado a (4.1)–(4.3) se puede formular de la siguiente manera:

Encontrar $v \in V_g$, tal que

$$\int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} + \int_{\Gamma_1} g_1 v \, d\Gamma, \quad \forall v \in V_0. \quad (4.9)$$

En torno a este problema se tiene el siguiente resultado:

Sí los problemas (4.1)–(4.3) y (4.9) tienen solución, entonces ésta tiene que ser única.

Demostración: Si u_1 y u_2 son dos soluciones de (4.9), entonces $\bar{u} = u_1 - u_2$ satisface

$$\int_{\Omega} a \nabla \bar{u} \cdot \nabla v \, d\mathbf{x} = 0, \quad \forall v \in V_0.$$

Puesto que, en particular, $\bar{u} \in V_0$, entonces tomando $v = \bar{u}$ en la última ecuación, se obtiene:

$$\int_{\Omega} a |\nabla \bar{u}|^2 \, d\mathbf{x} = 0.$$

Esta última igualdad y (4.5) implican que $|\nabla \bar{u}| = 0$, es decir \bar{u} es constante. Sin embargo, $\bar{u} = 0$ sobre Γ_0 por lo que se debe satisfacer $u_1 = u_2$.

4.3. Problema de minimización asociado

Es bien conocido que la solución del problema variacional (4.9) minimiza el funcional de energía potencial del problema, definido por

$$J(v) = \frac{1}{2} \int_{\Omega} a |\nabla v|^2 \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x} - \int_{\Gamma_1} g_1 v \, d\Gamma, \quad v \in V_0. \quad (4.10)$$

Este funcional cuadrático es estrictamente convexo si $a > 0$, por lo que una función u minimiza $J(v)$ sí, y sólo sí, la primera variación de J en u , en cualquier dirección v , es cero. Es decir, si la derivada de de Gâteaux,

$$J^{(1)}(u; v) = \lim_{t \rightarrow 0} \frac{J(u + tv) - J(u)}{t}, \quad (4.11)$$

es igual a cero para toda $v \in V_0$.

Afirmación. La primera derivada de $J(v)$ es

$$J^{(1)}(u; v) = \int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x} - \int_{\Gamma_1} g_1 v \, d\Gamma. \quad (4.12)$$

Demostración: De acuerdo a (4.10) se tiene

$$\begin{aligned} J(u + tv) &= \frac{1}{2} \int_{\Omega} \nabla(u + tv) \cdot \nabla(u + tv) \, d\mathbf{x} - \int_{\Omega} f(u + tv) \, d\mathbf{x} - \int_{\Gamma_1} g_1(u + tv) \, d\Gamma \\ &= J(u) + t \left\{ \int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x} - \int_{\Gamma_1} g_1 v \, d\Gamma \right\} + \frac{t^2}{2} \int_{\Omega} a |\nabla v|^2 \, d\mathbf{x}. \end{aligned}$$

De esta última expresión se obtiene

$$\frac{J(u + tv) - J(u)}{t} = \int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} - \int_{\Omega} f v \, d\mathbf{x} - \int_{\Gamma_1} g_1 v \, d\Gamma + \frac{t}{2} \int_{\Omega} a |\nabla v|^2 \, d\mathbf{x}.$$

Tomando el límite cuando $t \rightarrow 0$, obtenemos la expresión (4.12). Esto demuestra que la solución del problema variacional (4.9) minimiza el funcional cuadrático (4.10).

4.4. El problema de Neumann puro

Cerraremos la discusión del problema elíptico considerando el caso $\Gamma_0 = \phi$, es decir cuando no se tienen condiciones tipo Dirichlet en el problema. El problema resultante se le denomina problema de Neumann puro, y consiste de las ecuaciones

$$-\nabla \cdot (a \nabla u) = f \quad \text{en } \Omega, \quad (4.13)$$

$$a \frac{\partial u}{\partial n} = g_1 \quad \text{sobre } \Gamma = \partial\Omega. \quad (4.14)$$

Este problema no tiene solución para funciones arbitrarias f y g_1 . Para que exista solución, las funciones f y g_1 deben estar relacionadas de manera especial mediante una **condición de compatibilidad**. Esta condición de compatibilidad (de los datos) puede derivarse de la formulación variacional

$$\int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} + \int_{\Gamma} g_1 v \, d\Gamma, \quad \forall v \in H^1(\Omega).$$

Tomando $v = 1$ sobre Ω , la ecuación anterior se reduce a

$$\int_{\Omega} f \, d\mathbf{x} + \int_{\Gamma} g_1 \, d\Gamma = 0. \quad (4.15)$$

Esta última ecuación es precisamente la condición de compatibilidad que los datos f y g_1 deben satisfacer para que el problema Neumann puro (4.13)–(4.14) tenga solución. Esta condición significa físicamente que el flujo g_1 , a través de la frontera Γ , debe estar blanceado por la fuente f , definida en Ω . Por lo tanto, suponiendo que se satisface (4.15), entonces el problema de Neumann puro tiene solución. Pero también se observa que la solución no es única, debido a que si u es una solución del problema, entonces $u + c$ con $c \in \mathbb{R}$ también es solución del mismo problema. Sin embargo, podemos pensar que la solución es única salvo constantes, introduciendo el espacio cociente $H^1(\Omega)/\mathbb{R}$. En resumen, tenemos que

Si f es una función en $L_2(\Omega)$, y g_1 es una función en $L_2(\Gamma)$, y ambas satisfacen la condición de compatibilidad (4.15), entonces el problema de Neumann (4.13)–(4.14) tiene solución y esta solución es única salvo por una constante. Es decir, la solución es única en el espacio cociente $H^1(\Omega)/\mathbb{R}$.

El hecho de que la solución sea única solo en el espacio cociente, indica implícitamente que la cantidad física relevante no es la función escalar (potencial) $u(\mathbf{x})$, sino la función vectorial $\boldsymbol{\sigma}(\mathbf{x}) = -a(\mathbf{x}) \nabla u(\mathbf{x})$, la cual si es única cuando el problema tiene solución. La función $\boldsymbol{\sigma}$ representa el vector flujo de energía térmica en un problema de calor y la tracción en un problema de elasticidad. Tomando en cuenta estas consideraciones, es posible derivar la condición de compatibilidad (4.15) utilizando el teorema de la divergencia:

$$\int_{\Omega} f \, d\mathbf{x} = \int_{\Omega} \nabla \cdot \boldsymbol{\sigma} \, d\mathbf{x} = \int_{\Gamma} \boldsymbol{\sigma} \cdot \mathbf{n} \, d\Gamma = \int_{\Gamma} -a \nabla u \cdot \mathbf{n} \, d\Gamma = - \int_{\Gamma} g_1 \, d\Gamma.$$

4.5. Ejercicios

Ejercicio 4.1. Escribe el problema modelo elíptico multidimensional, incluyendo la ecuación diferencial y las condiciones de frontera. Indica las condiciones que se pide que satisfaga el coeficiente de difusión.

Ejercicio 4.2. Repite el procedimiento de la Sección 4.2 para encontrar la formulación variacional del problema elíptico, y verifica que esta minimiza el funcional definido en (4.10).

Ejercicio 4.3. Indica bajo que condiciones el problema de Neumann puro (4.13)–(4.14) tiene solución y ésta es única.

5. El método de elemento finito en el problema bidimensional

Para facilitar el estudio, escribimos de nuevo el problema elíptico:

$$-\nabla \cdot (a \nabla u) = f \quad \text{en } \Omega, \quad (5.1)$$

$$u = g_0 \quad \text{sobre } \Gamma_0, \quad (5.2)$$

$$a \frac{\partial u}{\partial \mathbf{n}} = g_1 \quad \text{sobre } \Gamma_1, \quad (5.3)$$

cuya formulación variacional es:

Encontrar $u \in V_g$ tal que

$$\int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} + \int_{\Gamma_1} g_1 v \, d\Gamma, \quad \forall v \in V_0, \quad (5.4)$$

en donde

$$V_0 = \{v \in H^1(\Omega) : v = 0 \quad \text{sobre } \Gamma_0\}, \quad (5.5)$$

$$V_g = \{v \in H^1(\Omega) : v = g_0 \quad \text{sobre } \Gamma_0\}. \quad (5.6)$$

Al igual que en el caso unidimensional, para encontrar soluciones numéricas del problema (5.4), se deben aproximar los espacios de funciones de prueba V_0 y el espacio afín V_g por medio de espacios de funciones de dimensión finita. Para realizar la aproximación por medio del método de elemento finito, los espacios de funciones de aproximación deberán consistir de funciones continuas y que son polinomiales por tramos. Es decir, sobre cada elemento de una triangulación de la región Ω , la solución se aproxima por medio de un polinomio. Cuando se utilizan polinomios lineales, se obtienen las **funciones piramidales**, las cuales son el análogo bidimensional de las funciones sombrero, como veremos en esta sección.

5.1. Discretización del dominio

Dado el dominio Ω , existen muchas formas de subdividirlo en subdominios o elementos. Para simplificar la exposición, supondremos que el dominio es poligonal y escogemos triángulos como elementos de discretización. La Figura 5 ilustra el dominio y una triangulación del mismo.

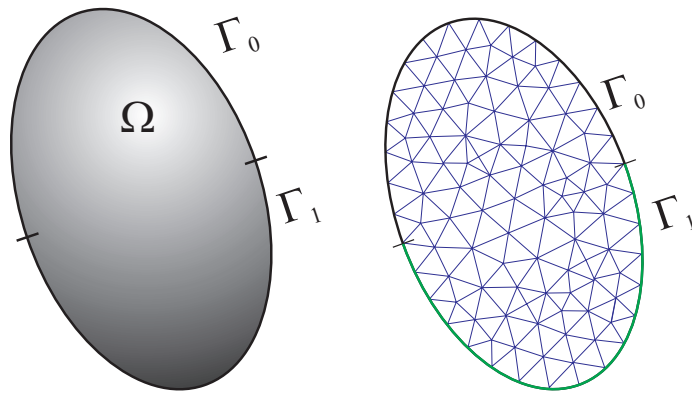


Figura 5: Dominio Ω con frontera $\partial\Omega = \Gamma_0 \cup \Gamma_1$ y su triangulación.

Denotaremos a la triangulación de Ω por

$$\mathcal{T}_h = \text{conjunto de triángulos en la discretización de } \Omega.$$

El parámetro de discretización h , proporciona una medida de la malla o triangulación, y se refiere al diámetro máximo de los triángulos en \mathcal{T}_h . A diferencia del caso unidimensional, en donde la subdivisión de Ω consiste de subintervalos adyacentes y no requiere de restricciones especiales, en el caso bidimensional se requiere que la triangulación satisfaga ciertas propiedades adicionales con el objeto de asegurar una aproximación correcta de las funciones en el espacio de funciones V_0 y en el espacio afín V_g . Estas propiedades (restricciones) se introducen a continuación.

La triangulación \mathcal{T}_h debe satisfacer las siguientes propiedades:

- a) El conjunto de todos los triángulos forma el dominio y sus fronteras, es decir $\bigcup_{T \in \mathcal{T}_h} T = \overline{\Omega}$.

- b) Dos triángulos T y T' en \mathcal{T}_h satisfacen una, y sólo una, de las siguientes propiedades
1. No se intersectan, es decir $T \cap T' = \phi$.
 2. Si se intersectan, lo hacen en un vértice común ó en una arista común.
- c) En la interfase entre la frontera Dirichlet Γ_0 y la frontera Neumann Γ_1 debe encontrarse un vértice de la triangulación.

La Figura 6 ilustra estas propiedades. Ahí se indican con una cruz (roja) las situaciones que no son permitidas en la triangulación, mientras que, encerradas (en verde), se indican algunas situaciones permitidas.

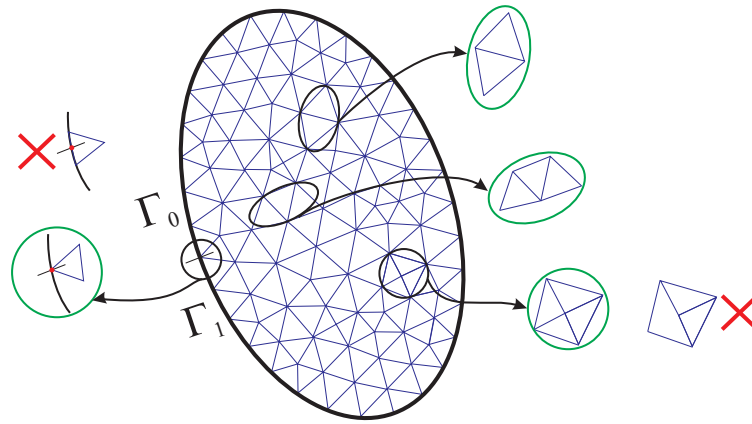


Figura 6: Propiedades de la triangulación

Una vez obtenida la triangulación \mathcal{T}_h , se construye la aproximación al espacio $H^1(\Omega)$. Los elementos en este espacio, como ya hemos mencionado anteriormente, se aproximan mediante funciones continuas que al restringirlas a cada triángulo resultan ser polinomios de primer grado. Es decir, las funciones se aproximan mediante polinomios lineales por tramos y que coinciden a lo largo de cada arista de dos triángulos adyacentes. A este espacio discreto lo denotamos por V_h , y formalmente se define como

$$V_h = \{v \in \mathcal{C}^0(\Omega) : v|_T \in \mathbb{P}_1, \quad \forall T \in \mathcal{T}_h\}, \quad (5.7)$$

en donde $\mathcal{C}^0(\Omega)$ denota al conjunto de funciones continuas en Ω , mientras que \mathbb{P}_1 denota el conjunto de polinomios de primer grado (lineales) en \mathbb{R}^2 . Los polinomios lineales en \mathbb{R}^2 son de la forma

$$p(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2, \quad \text{con } a_0, a_1, a_2 \text{ constantes.} \quad (5.8)$$

Por lo tanto, si queremos aproximar una función $u(x_1, x_2) \in H^1(\Omega)$ por medio de una función $u_h \in V_h$, es necesario que en cada elemento triangular calculemos los coeficientes a_i . Estos

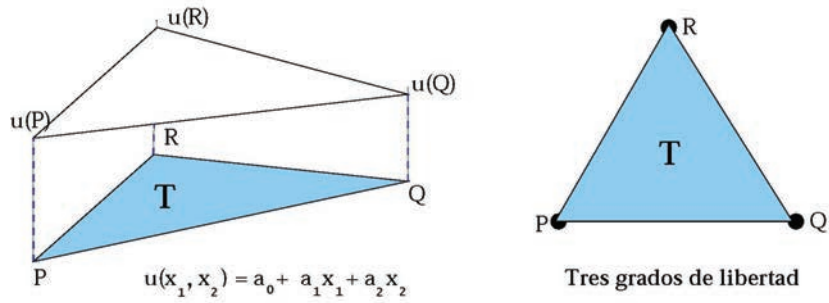


Figura 7: Polinomio lineal que aproxima a u sobre T .

coeficientes se pueden determinar de manera única, sobre cada triángulo T , con los valores de u en los tres vértices de dicho triángulo, como se muestra en la Figura 7.

Por otro lado, si se requiere un mayor orden de aproximación, el método de elemento finito permite aproximar con polinomios de orden mayor. Así por ejemplo, los polinomios de segundo grado en \mathbb{R}^2 son de la forma

$$p(x_1, x_2) = a_0 + a_1 x_1 + a_2 x_2 + a_3 x_1 x_2 + a_4 x_1^2 + a_5 x_2^2,$$

por lo que se tienen seis grados de libertad, y hay que determinar seis coeficientes. En este caso, para determinar los coeficientes de manera única, es posible utilizar los valores de la función sobre los vértices y sobre los puntos medios de cada triángulo $T \in \mathcal{T}$. En forma análoga, se construyen elementos de grados mayores. En la Figura 8 se muestran los elementos de segundo y tercer orden y sus grados de libertad.

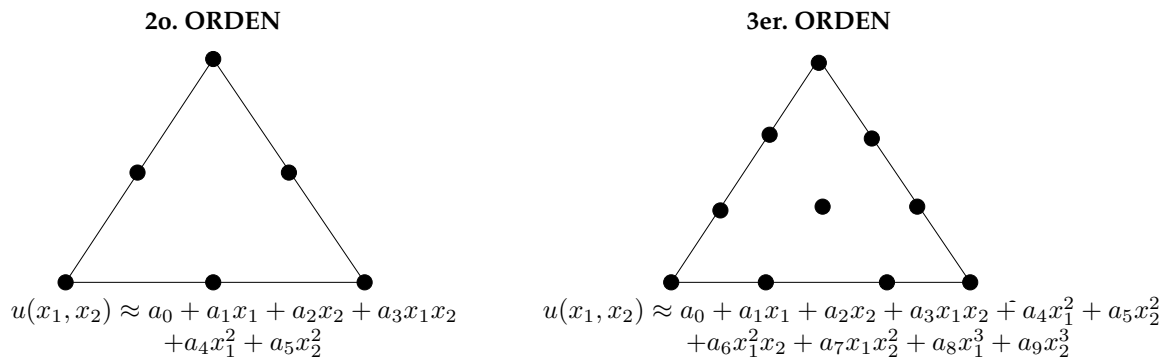


Figura 8: Elementos de 2o. y 3er. orden.

Malla y nodos de elemento finito

Al conjunto de vértices en la triangulación lo denotaremos por ϑ_h . Es decir

$$\vartheta_h = \{P = (x_1, x_2) \in \mathbb{R}^2 : P \text{ es vértice de algún } T \in \mathcal{T}_h\}. \quad (5.9)$$

De esta manera el conjunto de vértices de la triangulación más el conjunto de triángulos mismos constituyen la **mall**a triangular del elemento finito:

$$\vartheta_h + \mathcal{T}_h = \text{Malla triangular de elemento finito.} \quad (5.10)$$

Por otro lado, los **nod**os del elemento finito son los puntos que se utilizan para realizar la aproximación sobre cada triángulo. Así por ejemplo, para el elemento de primer orden el conjunto de nodos coincide con los vértices de la triangulación, mientras que para el elemento de segundo orden, el conjunto de nodos son los vértices más lo puntos medios de cada triángulo.

5.2. Construcción de las funciones base

El análogo bidimensional de las funciones sombrero (3.4) son las **funciones piramidales**. Estas funciones pertenecen al espacio de aproximación V_h , (5.7), y cada nodo P tiene asociada una de estas funciones. La función piramidal asociada al nodo P se define como uno en dicho nodo y cero en todos lo demás nodos de la mall

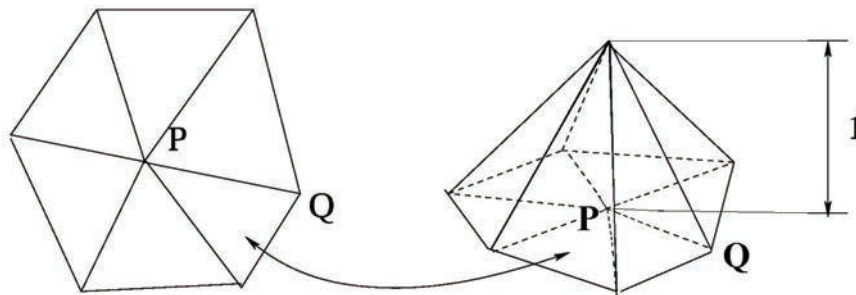
, como se ilustra en la Figura 9. En dicha figura no se muestra todo el dominio, solo se muestra la gráfica de la función en el soporte de la misma, pero se entiende que en todos los demás triángulos de la mall la función es igual a cero. Obsérvese que, restringida a cada triángulo T , la función piramidal es un polinomio lineal.


Figura 9: Soporte y gráfica de la función piramidal φ_P .

Las funciones piramidales son las funciones básicas que sirven para para generar las funciones de aproximación en el elemento finito lineal de Lagrange. Claramente, estas funciones pertenecen a V_h y su propiedad básica es que, para todo P, Q en ϑ_h

$$\varphi_P(Q) = \delta_{PQ} = \begin{cases} 1 & \text{si } P = Q, \\ 0 & \text{si } P \neq Q. \end{cases} \quad (5.11)$$

Propiedades de las funciones piramidales. Las principales propiedades de las funciones piramidales son:

1. **Son funciones en V_h con soporte compacto.** El soporte de cada función piramidal φ_P , es el conjunto de triángulos en \mathcal{T}_h que tienen como vértice común al punto P . Utilizaremos la siguiente notación para indicar al soporte de φ_P

$$\mathcal{T}_h(P) = \text{supp } \varphi_P.$$

2. **Son funciones linealmente independientes.** Con el objeto de verificarlo, sean α_P , con $P \in \vartheta_h$ un conjunto de coeficientes constantes tales que

$$\sum_{P \in \vartheta_h} \alpha_P \varphi_P(\mathbf{x}) = 0, \quad \forall \mathbf{x} = (x_1, x_2) \in \Omega.$$

Para cada $Q \neq P$ en ϑ_h , se tiene

$$0 = \sum_{P \in \vartheta_h} \alpha_P \varphi_P(Q) = \sum_{P \in \vartheta_h} \alpha_P \delta_{PQ} = \alpha_Q,$$

lo cual comprueba la independencia lineal.

3. **Son funciones que generan V_h .** Para probarlo, sea cualquier $v_h \in V_h$ y defínase la función $v_h^* \in V_h$ por medio de la combinación lineal

$$v_h^*(\mathbf{x}) = \sum_{P \in \vartheta_h} v_h(P) \varphi_P(\mathbf{x}).$$

Para cada $Q \in \vartheta_h$ se tiene

$$v_h^*(Q) = \sum_{P \in \vartheta_h} v_h(P) \varphi_P(Q) = \sum_{P \in \vartheta_h} v_h(P) \delta_{PQ} = v_h(Q),$$

por lo que $v_h^* = v_h$. Es decir, v_h es una combinación lineal de las φ_P con coeficientes constantes $\alpha_P = v_h(P)$.

Resumiendo, toda $v_h \in V_h$, es combinación lineal de funciones piramidales y además

$$v_h(\mathbf{x}) = \sum_{P \in \vartheta_h} v_h(P) \varphi_P(\mathbf{x}). \quad (5.12)$$

Esta propiedad es el análogo a la propiedad 3 de las funciones sombrero.

Conclusión. El conjunto de las funciones piramidales φ_P , con $P \in \vartheta_h$, forma una base del espacio de funciones V_h que aproxima $H^1(\Omega)$. A este conjunto base lo denotaremos por

$$\beta_h = \{\varphi_P : P \in \vartheta_h\}. \quad (5.13)$$

Por lo tanto, V_h es un espacio de funciones de dimensión finita y su dimensión es

$$\dim V_h = \#(\vartheta_h) = \text{número de vértices de la triangulación.}$$

Denotamos por N este número, es decir

$$N \equiv \dim V_h \equiv \text{número de vértices de la triangulación.} \quad (5.14)$$

5.3. Aproximación del problema variacional

Con el objeto de encontrar una solución aproximada del problema variacional (5.4), debemos encontrar una aproximación del espacio de funciones prueba V_0 (5.5). Para ello introducimos la siguiente notación

$$N_I \equiv \text{número de vértices "interiores"}. \quad (5.15)$$

Por **"vértices interiores"** entendemos aquellos vértices donde se debe calcular el valor de la solución u . Más precisamente, un vértice es "interior" si este no es un vértice sobre la frontera Dirichlet Γ_0 . Por lo tanto, los vértices interiores son los que se encuentran en el interior de Ω y sobre el interior de la porción de frontera Neumann Γ_1 . Numerando los vértices de la triangulación, y suponiendo que primero se numeran los vértices interiores, entonces escribimos

$$\vartheta_{Ih} = \{P_i\}_{i=1}^{N_I}, \quad (\text{Nodos interiores}) \quad (5.16)$$

$$\vartheta_{Dh} = \{P_i\}_{i=N_I+1}^N. \quad (\text{Nodos Dirichlet}) \quad (5.17)$$

Obsérvese que

$$\vartheta_{Ih} = \vartheta_h - \vartheta_{Dh}.$$

Por lo tanto, el conjunto de funciones que aproxima el espacio V_0 es

$$\beta_{0h} = \{\varphi_P : P \in \vartheta_{Ih}\}, \quad (5.18)$$

por lo que podemos escribir

$$V_{0h} = \text{gen} \{\beta_{0h}\}. \quad (5.19)$$

Con los anteriores elementos podemos formular el problema variacional discreto, asociado a (5.4) de la siguiente manera:

Encontrar $u_h \in V_h$ con $u_h(P) = g_0(P)$, $\forall P \in \vartheta_{Dh}$, tal que

$$\int_{\Omega} a \nabla u_h \cdot \nabla v_h \, d\mathbf{x} = \int_{\Omega} f v_h \, d\mathbf{x} + \int_{\Gamma_1} g_1 v_h \, d\Gamma, \quad \forall v_h \in V_{0h}. \quad (5.20)$$

Este problema discreto puede formularse en términos de funciones piramidales, notando que

- Para que se satisfaga (5.20), basta con que se cumpla para cada una las funciones base $\varphi_Q \in \beta_{0h}$, debido a (5.19).
- La incógnita $u_h \in V_h$ se puede expandir como combinación lineal de las funciones base piramidales

$$u_h(\mathbf{x}) = \sum_{P \in \vartheta_h} u_h(P) \varphi_P(\mathbf{x}) = \underbrace{\sum_{P \in \vartheta_{Ih}} u_h(P) \varphi_P(\mathbf{x})}_{\text{interior}} + \underbrace{\sum_{P \in \vartheta_{Dh}} g_0(P) \varphi_P(\mathbf{x})}_{\text{Frontera Dirichlet}}.$$

En la última sumatoria se ha hecho uso de que $u_h(P) = g_0(P)$ si P es un vértice sobre la frontera Dirichlet Γ_0 .

Tomando en cuenta estas consideraciones, el problema variacional discreto (5.20) también se puede escribir de la siguiente manera:

Para cada $P \in \vartheta_{Ih}$, encontrar $u_h(P)$, tal que:

$$\begin{aligned} \sum_{P \in \vartheta_{Ih}} \left(\int_{\Omega} a \nabla \varphi_P \cdot \nabla \varphi_Q \, d\mathbf{x} \right) u_h(P) &= \int_{\Omega} f \varphi_Q \, d\mathbf{x} + \int_{\Gamma_1} g_1 \varphi_Q \, d\Gamma \\ &- \sum_{P \in \vartheta_{Dh}} \left(\int_{\Omega} a \nabla \varphi_P \cdot \nabla \varphi_Q \, d\mathbf{x} \right) g_0(P), \quad \forall Q \in \vartheta_{Ih}. \end{aligned}$$

Tomando en cuenta la numeración de los nodos Dirichlet y de los nodos interiores, (5.16) y (5.17), también se puede escribir el problema variacional discreto de la siguiente manera:

Encontrar los valores $u_j \equiv u_h(P_j)$, $j = 1, \dots, N_I$, tales que

$$\begin{aligned} \sum_{j=1}^{N_I} \left(\int_{\Omega} a \nabla \varphi_j \cdot \nabla \varphi_i \, d\mathbf{x} \right) u_j &= \int_{\Omega} f \varphi_i \, d\mathbf{x} + \int_{\Gamma_1} g_1 \varphi_i \, d\Gamma \\ &- \sum_{j=N_I+1}^N \left(\int_{\Omega} a \nabla \varphi_j \cdot \nabla \varphi_i \, d\mathbf{x} \right) g_0(P_j), \quad \forall i = 1, \dots, N_I. \end{aligned} \quad (5.21)$$

en donde se ha hecho la identificación $\varphi_j = \varphi_{P_j}$ para todo $1 \leq j \leq N$. Nótese que la expresión anterior constituye un sistema de N_I ecuaciones con N_I incógnitas de la forma

$$\mathbb{A}_h \mathbf{U}_h = \mathbf{b}_h, \quad (5.22)$$

en donde \mathbf{U}_h es el vector de incógnitas $\{u_j\}$. Los coeficientes de la matriz \mathbb{A}_h y el vector \mathbf{b}_h son, respectivamente,

$$a_{ij} = \int_{\Omega} a \nabla \varphi_i \cdot \nabla \varphi_j \, d\mathbf{x}, \quad (5.23)$$

$$b_i = \int_{\Omega} f \varphi_i \, d\mathbf{x} + \int_{\Gamma_1} g_1 \varphi_i \, d\Gamma - \sum_{j=N_I+1}^N \left(\int_{\Omega} a \nabla \varphi_j \cdot \nabla \varphi_i \, d\mathbf{x} \right) g_0(P_j). \quad (5.24)$$

NOTA. En el desarrollo anterior se pueden tomar las siguientes aproximaciones

$$a(\mathbf{x}) \approx a_h(\mathbf{x}) = \sum_{j=1}^N a(P_j) \varphi_j(\mathbf{x}). \quad (5.25)$$

$$g_1(\mathbf{x}) \approx g_{1h}(\mathbf{x}) = \sum_{j=1}^{N_I} g_1(P_j) \varphi_j(\mathbf{x}). \quad (5.26)$$

$$g_0(\mathbf{x}) \approx g_{0h}(\mathbf{x}) = \sum_{j=N_I+1}^N g_0(P_j) \varphi_j(\mathbf{x}), \quad (5.27)$$

o bien trabajar con las funciones originales y después aproximar las integrales resultantes en (5.23), (5.24). Obsérvese que se puede sustituir (5.27) en (5.24) para obtener la fórmula más corta

$$b_i = \int_{\Omega} f \varphi_i \, d\mathbf{x} + \int_{\Gamma_1} g_1 \varphi_i \, d\Gamma - \int_{\Omega} a \nabla \varphi_i \cdot \nabla g_0 \, d\mathbf{x}. \quad (5.28)$$

5.4. Propiedades de la matriz de rigidez

Las propiedades más importantes de la matriz de rigidez \mathbb{A}_h son que ésta es simétrica y definida positiva, por lo que el sistema de ecuaciones (5.22) tiene **solución única** $\mathbf{U}_h = \{u_j\}_{j=1}^{N_I}$. Esta propiedad inmediatamente implica que el problema discreto (5.20) asociado al problema variacional (5.4) tiene la solución única

$$u_h(\mathbf{x}) = \sum_{j=1}^{N_I} u_j \varphi_j(\mathbf{x}) + \sum_{j=N_I+1}^N g_0(P_j) \varphi_j(\mathbf{x}). \quad (5.29)$$

A continuación demostramos las principales características de la matriz \mathbb{A}_h :

Simetría. Es obvio que la matriz \mathbb{A}_h es simétrica dado que

$$a_{ij} = \int_{\Omega} a(\mathbf{x}) \nabla \varphi_i(\mathbf{x}) \cdot \nabla \varphi_j(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} a(\mathbf{x}) \nabla \varphi_j(\mathbf{x}) \cdot \nabla \varphi_i(\mathbf{x}) \, d\mathbf{x} = a_{ji}.$$

Positividad. No es tan obvio, sin embargo, que \mathbb{A}_h sea definida positiva. Para demostrarlo, considérese la i -ésima componente de $\mathbb{A}_h \mathbf{U}_h$

$$(\mathbb{A}_h \mathbf{U}_h)_i = \sum_{j=1}^{N_I} a_{ij} u_j.$$

Entonces

$$\begin{aligned} \mathbf{U}_h^T \mathbb{A}_h \mathbf{U}_h &= \sum_{i=1}^{N_I} u_i (\mathbb{A}_h \mathbf{U}_h)_i = \sum_{i=1}^{N_I} \sum_{j=1}^{N_I} a_{ij} u_i u_j \\ &= \sum_{i=1}^{N_I} \sum_{j=1}^{N_I} \left(\int_{\Omega} a(\mathbf{x}) \nabla \varphi_i(\mathbf{x}) \cdot \nabla \varphi_j(\mathbf{x}) \, d\mathbf{x} \right) u_i u_j \\ &= \int_{\Omega} a(\mathbf{x}) \nabla \bar{u}_h \cdot \nabla \bar{u}_h \, d\mathbf{x} \quad , \quad \text{donde } \bar{u}_h(\mathbf{x}) = \sum_{j=1}^{N_I} u_j \varphi_j(\mathbf{x}), \\ &\geq a_0 \int_{\Omega} \|\nabla \bar{u}_h\|^2 \, d\mathbf{x} \geq 0, \quad \text{por (4.5),} \end{aligned}$$

lo cual demuestra que \mathbb{A}_h es positiva semidefinida. Para completar la demostración basta con verificar que

$$\int_{\Omega} \|\nabla \bar{u}_h\|^2 d\mathbf{x} = 0 \quad \text{solo si } u_h = 0.$$

Suponiendo que la integral es cero, se sigue que u_h es constante y, dado que $u_h = 0$ sobre Γ_0 , entonces $u_h = 0$ sobre Ω . Se concluye que la matriz \mathbb{A}_h es definida positiva.

Sumabilidad. Al igual que en el caso unidimensional la matriz de rigidez \mathbb{A}_h posee la propiedad de sumabilidad

$$a_{ij} = \int_{\Omega} a(\mathbf{x}) \nabla \varphi_i \cdot \nabla \varphi_j d\mathbf{x} = \sum_{T \in \mathcal{T}_h} \int_T a(\mathbf{x}) \nabla \varphi_i \cdot \nabla \varphi_j d\mathbf{x}. \quad (5.30)$$

Esta propiedad es clave en el método de elemento finito y se utiliza para ensamblar la matriz de rigidez en forma eficiente, así como para simplificar los cálculos. Análogamente, se obtiene la propiedad de sumabilidad para el vector de carga (fuente) \mathbf{b}_h :

$$\int_{\Omega} f \varphi_i d\mathbf{x} = \sum_{T \in \mathcal{T}_h} \int_T f \varphi_i d\mathbf{x}. \quad (5.31)$$

Escasez de coeficientes no cero. La matriz de rigidez \mathbb{A}_h , al igual que en el caso unidimensional, también tiene escasez de coeficientes no cero, es decir es una matriz rala (“sparse”). Sin embargo, a diferencia del caso unidimensional, la matriz \mathbb{A}_h no es tridiagonal, debido a que $a_{ij} \neq 0$ solo si P_i y P_j son vértices adyacentes en la malla triangular. Por lo tanto los coeficientes no cero de la matriz dependerán de la triangulación y de la forma en como se numeren los vértices de la malla.

5.5. Aspectos prácticos

Con el objeto de simplificar la exposición, hacemos un resumen de la notación utilizada hasta el momento:

\mathcal{T}_h , triangulación del dominio Ω .

ϑ_h , conjunto de vértices de la triangulación.

ϑ_{Ih} , conjunto de vértices “interiores” (sobre $\Omega \cup \Gamma_1$).

ϑ_{Dh} , conjunto de vértices Dirichlet (sobre Γ_0).

N , número de vértices en la triangulación = $\#(\vartheta_h)$.

N_I , Número de “vértices interiores” = $\#(\vartheta_{Ih})$.

$\beta_h = \{\varphi_P\}_{P \in \mathcal{P}_h} = \{\varphi_i\}_{i=1}^N$, funciones base piramidales.

$\beta_{0h} = \{\varphi_P\}_{P \in \mathcal{P}_{0h}} = \{\varphi_i\}_{i=1}^{N_I}$, funciones base interiores.

$V_h = \{v_h \in \mathcal{C}^0(\Omega) : v_h|_T \in P_1, \forall T \in \mathcal{T}_h\} = \text{gen}\{\beta_h\}$, espacio discreto de funciones que aproxima $H^1(\Omega)$.

$V_{0h} = \{v_h \in V_h : v_h = 0 \text{ sobre } \Gamma_0\} = \text{gen}\{\beta_{0h}\}$, espacio discreto de funciones que aproxima V_0 .

La solución aproximada u_h del problema variacional (5.4) satisface el problema discreto (5.20). Esta solución discreta es de la forma

$$u_h(\mathbf{x}) = \sum_{j=1}^{N_I} u_j \varphi_j(\mathbf{x}) + \sum_{j=N_I+1}^N g_0(P_j) \varphi_j(\mathbf{x}), \quad \forall \mathbf{x} \in \Omega.$$

en donde los coeficientes $\{u_j\}_{j=1}^{N_I}$ resuelven el sistema de ecuaciones

$$\sum_{j=1}^{N_I} a_{ij} u_j = b_i, \quad i = 1, \dots, N_I,$$

con

$$a_{ij} = \int_{\Omega} a \nabla \varphi_i \cdot \nabla \varphi_j d\mathbf{x}, \quad i = 1, \dots, N_I,$$

y

$$b_i = \int_{\Omega} f \varphi_i d\mathbf{x} + \int_{\Gamma_1} g_1 \varphi_i d\Gamma - \int_{\Omega} a \nabla \varphi_i \cdot \nabla g_0 d\mathbf{x}.$$

Entonces, para tener completamente determinada la solución discreta u_h , solo hay que calcular las integrales que definen $a_{i,j}$ y b_i , y después resolver el sistema de ecuaciones resultante.

Cuando las funciones $a(x)$, $f(x)$ y $g_1(x)$ no sean simples será necesario utilizar métodos de integración numérica para calcular, con cierto nivel de precisión, los coeficientes a_{ij} y b_i . Estos métodos de integración numérica no deben introducir un error mayor al que produce la aproximación del problema por medio del método de elemento finito. El siguiente resultado, indica el orden de aproximación de la solución cuando se utiliza el elemento lineal de Lagrange para resolver el problema variacional.

Teorema. Sea u_h la solución del problema variacional discreto (5.20) obtenida mediante el método de elemento finito lineal, y sea u la solución exacta del problema variacional (5.4). Supóngase que a medida que refinamos la malla, los ángulos de cada triángulo T de \mathcal{T}_h están acotados inferiormente por un ángulo $\theta_0 > 0$, entonces

$$\lim_{h \rightarrow 0} \int_{\Omega} |u_h - u|^2 d\mathbf{x} = 0 \quad \text{y} \quad \lim_{h \rightarrow 0} \int_{\Omega} |\nabla u_h - \nabla u|^2 d\mathbf{x} = 0.$$

Además el error e_h está dado por

$$e_h = \|u_h - u\|_{L^2(\Omega)} \leq c h^2, \quad (5.32)$$

en donde c es una constante que no depende de h .

Entonces, el error en el cálculo de la solución por medio del método del elemento finito es de orden $\mathcal{O}(h^2)$. Por lo tanto, basta utilizar métodos de integración numérica que tengan al menos este orden de aproximación.

Cálculo de los coeficientes a_{ij} : Utilizando la propiedad de sumabilidad se tiene

$$a_{ij} = \int_{\Omega} a \nabla \varphi_i \cdot \nabla \varphi_j \, d\mathbf{x} = \sum_{T \in \mathcal{T}_h} \int_T a \nabla \varphi_i \cdot \nabla \varphi_j \, d\mathbf{x}.$$

Debido a que $\nabla \varphi_i$ es una función vectorial constante sobre cada triángulo T , entonces

$$a_{ij} = \sum_{T \in \mathcal{T}_h} \nabla \varphi_i \cdot \nabla \varphi_j \Big|_T \int_T a(\mathbf{x}) \, d\mathbf{x}, \quad (5.33)$$

en donde $\nabla \varphi_i \cdot \nabla \varphi_j \Big|_T$ denota el valor (constante) de $\nabla \varphi_i \cdot \nabla \varphi_j$ sobre el triángulo T . La integral en (5.33), cuando $a(\mathbf{x}) = a = \text{constante}$ es $a |T|$, en donde $|T|$ denota el área del triángulo T . Por lo tanto, en este caso

$$a_{ij} = a \sum_{T \in \mathcal{T}_h} \nabla \varphi_i \cdot \nabla \varphi_j \Big|_T |T|.$$

Por otro lado, cuando a depende de \mathbf{x} , y no es posible calcular la integral en (5.33) de manera exacta, se puede utilizar cualquiera de las siguientes aproximaciones:

1. Regla del trapecio:

$$\int_T a(\mathbf{x}) \, d\mathbf{x} \approx \frac{|T|}{3} \sum_{k=1}^3 a(P_{kT}),$$

en donde P_{1T}, P_{2T}, P_{3T} son los vértices del triángulo T .

2. Regla de Simpson:

$$\int_T a(\mathbf{x}) \, d\mathbf{x} \approx \frac{|T|}{3} \sum_{k=1}^3 a(M_{kT}),$$

en donde M_{1T}, M_{2T}, M_{3T} son los puntos medios de las aristas del triángulo T .

3. Regla del centro gravitacional

$$\int_T a(\mathbf{x}) \, d\mathbf{x} \approx |T| a(G_T),$$

en donde G_T es el centro gravitacional del triángulo T .

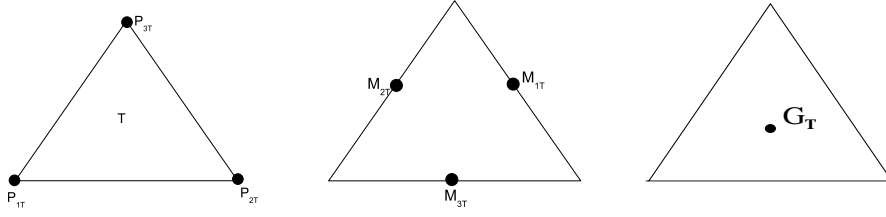


Figura 10: Vértices, puntos medios y centro gravitacional.

La Figura 10 muestra los puntos de cuadratura para cada una de las reglas anteriores.

El orden de aproximación de la integral es, $\mathcal{O}(h^2)$ para la regla del trapecio, $\mathcal{O}(h^4)$ para la regla de Simpson y $\mathcal{O}(h^2)$ para la regla del centro gravitacional. Además la regla del trapecio proporciona la integral exacta si $a(x)$ es un polinomio de grado 1 sobre cada triángulo T ; la regla de Simpson es exacta si $a(x)$ es un polinomio de grado 3 sobre cada triángulo T ; la regla del centro gravitacional es exacta si $a(x)$ es un polinomio de grado 1 sobre cada triángulo T .

Por otro lado, las integrales en (5.28) se pueden calcular en forma aproximada como se indica a continuación.

1. Las integrales de la forma

$$\int_{\Omega} a(\mathbf{x}) \nabla \varphi_i \cdot \nabla g_0 \, d\mathbf{x},$$

pueden aproximarse en forma análoga a como se calcularon los coeficientes a_{ij}

2. Las integrales de la forma

$$\int_{\Omega} f \varphi_i \, d\mathbf{x},$$

se expresan como:

$$\int_{\Omega} f \varphi_i \, d\mathbf{x} = \sum_{T \in \mathcal{T}_h} \int_T f \varphi_i \, d\mathbf{x},$$

y sobre cada triángulo $T \in \mathcal{T}_h$ se utiliza ya sea la regla del trapecio o bien la regla de Simpson. Estas integrales son no cero sólo cuando el nodo P_i es un vértice del triángulo T , y en este caso

$$\text{Regla del trapecio : } \int_T f \varphi_i \, d\mathbf{x} \approx \frac{|T|}{3} f(P_i).$$

$$\text{Regla de Simpson : } \int_T f \varphi_i \, d\mathbf{x} \approx \frac{|T|}{3} \sum_{k=1}^3 a(M_{kT}) \varphi_i(M_{kT}).$$

Obsérvese que $\varphi_i(M_{kT}) = 0$ para el punto medio sobre la arista opuesta al vértice P_i , y vale $1/2$ sobre los puntos medios de las aristas adyacentes a dicho vértice.

3. Las integrales de línea

$$\int_{\Gamma_1} g_1 \varphi_i d\Gamma,$$

se pueden calcular con la regla del trapecio o la regla de Simpson a lo largo de cada arista de la triangulación que se encuentre sobre Γ_1 . Esta integral es no cero sólo en los segmentos sobre Γ_1 que contenga a el nodo P_i .

5.6. Ejercicios

Ejercicio 5.1. Dada una región $\Omega \in \mathbb{R}^2$, indica como se construye una triangulación de elemento finito y las propiedades que debe satisfacer. ¿Cuáles son los ingredientes de una malla de elemento finito?

Ejercicio 5.2. Dada la triangulación \mathcal{T}_h del dominio Ω describe al conjunto de funciones que aproximan el espacio $H^1(\Omega)$. Las funciones piramidales forman una base de este conjunto, indica cuáles son sus propiedades principales.

Ejercicio 5.3. Dado el dominio $\Omega \in \mathbb{R}^2$ y una triangulación de elemento finito, indica lo que significa cada uno de los siguientes símbolos: $h, \mathcal{T}_h, N, N_I, \vartheta_h, \vartheta_{Dh}, \vartheta_{Ih}, V_h, V_{0h}, \beta_h, \beta_{0h}$.

Ejercicio 5.4. Escribe el procedimiento para aproximar el problema variacional mediante las funciones piramidales para llevarlo a un sistema de ecuaciones con matriz \mathbb{A} y lado derecho \mathbf{b}_h .

Ejercicio 5.5. ¿Cuáles son las propiedades de la matriz de rigidez del problema discreto (5.21)? Demuestra que la matriz de rigidez es definida positiva.

Ejercicio 5.6. Escribe la regla del trapecio, la regla de Simpson y la regla del centro gravitacional para aproximar la integral de una función f sobre un triángulo $T \in \mathcal{T}_h$. Ahora escribe las reglas de aproximación para integrar sobre f sobre todo $\Omega = \bigcup_{T \in \mathcal{T}_h} T$.

6. Extensión a problemas parabólicos

En esta sección consideraremos un *problema parabólico*. Este tipo de problemas son una extensión del problema elíptico al agregar un término dependiente del tiempo de la forma $\rho(\mathbf{x}, t) \partial u / \partial t$. Este tipo de ecuación diferencial aparece en problemas de difusión dependientes del tiempo. El nuevo término nos obliga a discretizar la variable en el tiempo además de las variables espaciales. Para la integración en el tiempo consideraremos el estudio de dos de los métodos más sencillos, a saber, el *método explícito* y el *método implícito* de Euler.

6.1. Problema modelo

Considérese el siguiente problema parabólico:

Encontrar $u = u(\mathbf{x}, t)$ que satisfice

$$\rho \frac{\partial u}{\partial t} - \nabla \cdot (a \nabla u) = f \quad \text{en} \quad \Omega \times (0, T), \quad (6.1)$$

$$u = g_0 \quad \text{sobre} \quad \Gamma_0 \times (0, T), \quad (6.2)$$

$$a \frac{\partial u}{\partial n} = g_1 \quad \text{sobre} \quad \Gamma_1 \times (0, T), \quad (6.3)$$

$$u(\mathbf{x}, 0) = u_0(\mathbf{x}) \quad \forall \mathbf{x} \in \Omega, \quad (6.4)$$

en donde las funciones ρ , a , g_0 , g_1 , y f , así como la incógnita u , dependen de $\mathbf{x} \in \Omega$ y $t \geq 0$. La frontera está dividida en dos porciones como se ilustra en la Figura 11. En una porción de la misma, Γ_0 , se imponen condiciones de tipo Dirichlet (6.2), y en la porción restante, Γ_1 , se imponen condiciones de tipo Neumann (6.3).

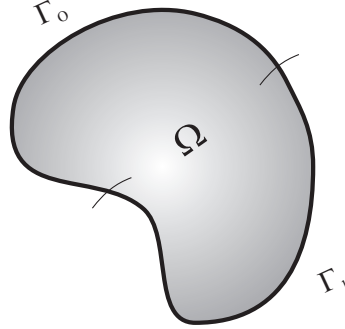


Figura 11: Dominio de interés, $\Omega \subset \mathbb{R}^d$, $d = 2$ ó 3 .

Además, suponemos que las funciones a y ρ satisfacen las siguientes restricciones:

$$0 \leq a(\mathbf{x}, t) \leq M, \quad (6.5)$$

$$0 < \rho_0 \leq \rho(\mathbf{x}, t) \leq Q, \quad (6.6)$$

en donde M y Q son constantes reales positivas.

6.2. Formulación variacional

Al igual que para el problema elíptico, introducimos los siguientes espacios de funciones

$$V = H^1(\Omega), \quad (6.7)$$

$$V_0 = \{v \in V : v = 0 \quad \text{sobre} \quad \Gamma_0\}. \quad (6.8)$$

El espacio V_0 es el espacio de funciones de prueba para derivar el *problema parabólico variacional*. Para simplificar la notación denotamos $\partial u / \partial t$ por \dot{u} . La formulación variacional del

problema parabólico es:

Dado $u(0) = u_0(\mathbf{x})$, para $t > 0$ calcular $u(t) \in V$ tal que $u(t) = g_0(t)$ sobre Γ_0 , y que satisfice la ecuación integral

$$\int_{\Omega} \rho \dot{u} v \, d\mathbf{x} + \int_{\Omega} a \nabla u \cdot \nabla v \, d\mathbf{x} = \int_{\Omega} f v \, d\mathbf{x} + \int_{\Gamma_1} g_1 v \, d\Gamma, \quad \forall v \in V_0. \quad (6.9)$$

En esta formulación hemos utilizado la notación $u(t) = u(\mathbf{x}, t)$ y $g_0(t) = g_0(\mathbf{x}, t)$, con el objeto de simplificar.

6.3. Aproximación por medio del método de elemento finito

El procedimiento para discretizar el problema variacional (6.9) es el mismo que utilizamos para discretizar la formulación variacional del problema elíptico. Dada una triangulación \mathcal{T}_h de Ω , como la mostrada en la Figura 5, se definen las funciones piramidales $\{\varphi_i\}_{i=1}^N$, cuya construcción y propiedades se establecen en la Sección 5.2. Los espacios de funciones V y V_0 , definidos en (6.7) y (6.8), se aproximan mediante los espacios discretos

$$V_h = \{v \in \mathcal{C}(\Omega) : v|_T \in P_1, \quad \forall T \in \mathcal{T}_h\} = \text{gen}\{\varphi_1, \dots, \varphi_N\}, \quad (6.10)$$

$$V_{0h} = \{v \in V_h : v = 0 \text{ sobre } \Gamma_0\} = \text{gen}\{\varphi_1, \dots, \varphi_{N_I}\}. \quad (6.11)$$

Por lo tanto, el problema variacional discreto asociado a (6.9) se puede escribir de la siguiente manera:

Dado el estado inicial $u_h(\mathbf{x}, 0) = u_0(\mathbf{x})$, para $t > 0$ encontrar $u_h(\mathbf{x}, t) \in V_h$ con $u_h(P, t) = g_0(P, t)$ en cada vértice $P \in \vartheta_{0h}$, de tal manera que:

$$\int_{\Omega} [\rho \dot{u}_h v_h + a \nabla u_h \cdot \nabla v_h] \, d\mathbf{x} = \int_{\Omega} f v_h \, d\mathbf{x} + \int_{\Gamma_1} g_1 v_h \, d\Gamma, \quad \forall v_h \in V_{0h}. \quad (6.12)$$

Utilizando la notación $u_j(t) = u_h(P_j, t)$ y $g_{0j}(t) = g_0(P_j, t)$, en donde $P_j \in \vartheta_h$, son los vértices de la triangulación, podemos escribir a $u_h \in V_h$ como

$$u_h(\mathbf{x}, t) = \sum_{j=1}^{N_I} u_j(t) \varphi_j(\mathbf{x}) + \sum_{j=N_I+1}^N g_{0j}(t) \varphi_j(\mathbf{x}). \quad (6.13)$$

Además, si tomamos $v_h = \varphi_i$, con $i = 1, \dots, N_I$, en (6.12), obtenemos

$$\begin{aligned} \sum_{j=1}^{N_I} \left[\left(\int_{\Omega} \rho \varphi_j \varphi_i \, d\mathbf{x} \right) \dot{u}_j + \left(\int_{\Omega} a \nabla \varphi_j \cdot \nabla \varphi_i \, d\mathbf{x} \right) u_j \right] &= \int_{\Omega} f \varphi_i \, d\mathbf{x} + \int_{\Gamma_1} g_1 \varphi_i \, d\Gamma \\ - \sum_{j=N_I+1}^N \left[\left(\int_{\Omega} \rho \varphi_j \varphi_i \, d\mathbf{x} \right) \dot{g}_{0j} + \left(\int_{\Omega} a \nabla \varphi_j \cdot \nabla \varphi_i \, d\mathbf{x} \right) g_{0j} \right], & \end{aligned} \quad (6.14)$$

con

$$u_j(0) = u_0(P_j), \quad j = 1, \dots, N_I. \quad (6.15)$$

Observamos que (6.14) y (6.15) constituyen un **sistema de N_I ecuaciones diferenciales ordinarias (EDO) con N_I incógnitas**, que podemos representar en la forma

$$\mathbb{S}_h \dot{\mathbf{U}}_h(t) + \mathbb{A}_h \mathbf{U}_h(t) = \mathbf{b}_h(t), \quad t > 0, \quad (6.16)$$

$$\mathbf{U}_h(0) = \mathbf{U}_0. \quad (6.17)$$

en donde las matrices \mathbb{S}_h y \mathbb{A}_h tienen los coeficientes

$$s_{ij} = \int_{\Omega} \rho(t) \varphi_i \varphi_j dx, \quad a_{ij} = \int_{\Omega} a(t) \nabla \varphi_i \cdot \nabla \varphi_j dx, \quad 1 \leq i, j \leq N_I, \quad (6.18)$$

respectivamente. Los vectores $\mathbf{U}_h(t)$, \mathbf{U}_0 , son

$$\mathbf{U}_h(t) = \{u_j(t)\}_{j=1}^{N_I}, \quad \mathbf{U}_0 = \{u_0(P_j)\}_{j=1}^{N_I}, \quad (6.19)$$

respectivamente. Finalmente el vector $\mathbf{b}_h(t)$ tiene los coeficientes:

$$\begin{aligned} b_i(t) = & \int_{\Omega} f(t) \varphi_i dx + \int_{\Gamma_1} g_1(t) \varphi_i d\Gamma - \sum_{j=N_I+1}^N \left(\int_{\Omega} \rho \varphi_i \varphi_j dx \right) \dot{g}_{0j}(t) \\ & - \sum_{j=N_I+1}^N \left(\int_{\Omega} a \nabla \varphi_i \cdot \nabla \varphi_j dx \right) g_{0j}(t). \end{aligned} \quad (6.20)$$

6.4. Solución del sistema de EDO

La discretización del problema parabólico (6.1)–(6.3), mediante el método de elemento finito, produce un sistema de ecuaciones diferenciales ordinarias (EDO) lineales de primer orden con matrices simétricas \mathbb{S} y \mathbb{A} . El sistema de EDO será un sistema con coeficientes constantes si las funciones ρ y a no dependen de la variable t , de otra manera el sistema será de coeficientes variables. Al igual que en el caso del problema elíptico, es fácil demostrar que la matriz \mathbb{A}_h es definida semipositiva si $a(\mathbf{x}, t) \geq 0$, $\forall \mathbf{x} \in \Omega$ and $t > 0$. En forma análoga, se puede demostrar que la matriz \mathbb{S}_h es definida positiva (y por lo tanto invertible) si existe $\rho_0 > 0$ tal que $\rho(\mathbf{x}, t) > \rho_0$, $\forall \mathbf{x} \in \Omega$ and $t > 0$.

Existencia y unicidad del sistema lineal de EDO. De acuerdo a lo anterior, si suponemos que las condiciones (6.5)–(6.6) se satisfacen para las funciones $\rho(\mathbf{x}, t)$ y $a(\mathbf{x}, t)$, entonces podemos asegurar que el sistema lineal de EDO (6.16)–(6.17) tiene solución y que esta solución es única. Esto debido a que de (6.16) se obtiene

$$\dot{\mathbf{U}}(t) = \mathbb{S}_h^{-1} \mathbf{b} + \mathbb{S}_h^{-1} \mathbb{A}_h \mathbf{U}(t),$$

en donde el lado derecho es una función continua respecto de t y Lipschitz continua respecto de \mathbf{U} .

Por otro lado, dado que el número de ecuaciones diferenciales N_I puede llegar a ser de decenas de miles, o incluso millones en el caso tridimensional, entonces es conveniente aplicar algún método numérico para resolver el sistema lineal de EDO. En este trabajo solo consideraremos los métodos más simples: el método explícito de Euler y el método implícito de Euler. Estos métodos, aunque son métodos de primer orden (orden bajo), son muy robustos y permiten resolver el sistema lineal de EDO en forma sencilla y económica, sobre todo cuando el número de ecuaciones es grande.

Método explícito de Euler. Sea Δt el parámetro de discretización en el tiempo, denotamos por \mathbf{U}^n a $\mathbf{U}(n\Delta t)$ para cada $n \in \mathbb{N}$. El **método explícito de Euler** consiste en reemplazar la derivada temporal de \mathbf{U} en $t = n \Delta t$ por medio del cociente de diferencias hacia adelante, así como reemplazar los términos restantes por sus valores en el tiempo $t = n \Delta t$. Se obtiene el siguiente esquema:

Dado $\mathbf{U}^0 = \mathbf{U}_0$, para $n \geq 1$ suponiendo conocido \mathbf{U}^n , calcular \mathbf{U}^{n+1} por medio del siguiente esquema:

$$\mathbb{S}_h \frac{\mathbf{U}^{n+1} - \mathbf{U}^n}{\Delta t} + \mathbb{A}_h \mathbf{U}^n = \mathbf{b}^n. \quad (6.21)$$

Entonces, si se conoce \mathbf{U}^n , se calcula \mathbf{U}^{n+1} resolviendo el sistema de ecuaciones algebraicas:

$$\mathbb{S}_h \mathbf{U}^{n+1} = \Delta t \mathbf{b}^n + (\mathbb{S}_h - \Delta t \mathbb{A}_h) \mathbf{U}^n. \quad (6.22)$$

Sin embargo, es posible simplificar el proceso mediante una técnica, muy utilizada por los usuarios del método de elemento finito, denominada *mass lumping*. Esto es posible cuando la función ρ no depende de \mathbf{x} , ya que es posible aproximar las integrales *de masa* en (6.18) mediante la regla del trapecio para obtener los coeficientes s_{ij} sin degradar el orden del método de elemento finito. El resultado es que la matriz \mathbb{S} será un múltiplo de la matriz identidad, lo cual es crucial, ya que para calcular \mathbf{U}^{n+1} en (6.22) no será necesario resolver un sistemas de ecuaciones. En su lugar solo se necesita evaluar la multiplicación de una matriz por un vector en cada paso de tiempo, lo cual es mucho más económico, justificando la denominación del método como esquema explícito.

Método implícito de Euler. El método implícito de Euler consiste en reemplazar la derivada temporal de \mathbf{U} en $t = (n + 1) \Delta t$ por medio del cociente de diferencias hacia atrás, así como reemplazar los términos restantes por sus valores en el tiempo $t = (n + 1) \Delta t$. Se obtiene el esquema:

Dado $\mathbf{U}^0 = \mathbf{U}_0$, para $n \geq 1$ suponiendo conocido \mathbf{U}^n , calcular \mathbf{U}^{n+1} por medio del siguiente esquema:

$$\mathbb{S}_h \frac{\mathbf{U}^{n+1} - \mathbf{U}^n}{\Delta t} + \mathbb{A}_h \mathbf{U}^{n+1} = \mathbf{b}^{n+1}. \quad (6.23)$$

Entonces, si se conoce \mathbf{U}^n , se calcula \mathbf{U}^{n+1} implícitamente de la solución del sistema de ecuaciones algebraicas

$$(\mathbb{S}_h + \Delta t \mathbb{A}_h) \mathbf{U}^{n+1} = \Delta t \mathbf{b}^{n+1} + \mathbb{S}_h \mathbf{U}^n. \quad (6.24)$$

Por lo tanto, con el método implícito es necesario resolver un sistema de ecuaciones de tamaño $N_I \times N_I$ en cada paso del tiempo, independientemente de si se realiza *mass lumping* o no, y por eso el esquema se denomina implícito. Por esta razón se podría pensar que los métodos implícitos son mucho más caros, desde el punto de vista computacional, que los métodos explícitos. Sin embargo, esto no es del todo cierto debido a que los esquemas implícitos usualmente no necesitan restricciones sobre el parámetro de discretización Δt para garantizar estabilidad numérica, a diferencia de los esquemas explícitos, como veremos a continuación. Además, los métodos implícitos son mucho más efectivos en problemas rígidos (problemas de mutiescala).

6.5. Análisis de estabilidad de los métodos

Existen diferentes conceptos de estabilidad cuando se estudian ecuaciones diferenciales. En el caso que nos ocupa estamos interesados en la *estabilidad numérica* del esquema utilizado. La estabilidad de los esquemas numéricos está muy relacionada con el error numérico. *Un esquema es estable si los errores cometidos en un paso de tiempo no crecen conforme se avanza en los cálculos en pasos posteriores.* Para los problemas dependientes del tiempo, la estabilidad garantiza que el método numérico proporciona una solución acotada cuando la solución exacta es acotada. En muchos casos se investiga la estabilidad por medio del análisis de Von Neumann.

Estabilidad del método explícito de Euler. Para el análisis de estabilidad de los esquemas numéricos, usualmente se suprime el efecto de las fuentes, para descartar los efectos externos, y solo se toman en cuenta las propiedades fundamentales del sistema discretizado. Por lo tanto, tomamos $\mathbf{b} = \mathbf{0}$ en (6.22), y obtenemos la siguiente solución de la ecuación homogénea:

$$\mathbf{U}^{n+1} = (I - \Delta t \mathbb{S}_h^{-1} \mathbb{A}_h) \mathbf{U}^n. \quad (6.25)$$

Con el objeto de garantizar que \mathbf{U}^n permanezca acotado para toda n , es necesario que los modos fundamentales de la solución numérica permanezcan acotados conforme se avanza en el tiempo. Los modos fundamentales de (6.25) están dados por los vectores propios de la matriz $I - \Delta t \mathbb{S}_h^{-1} \mathbb{A}_h$. Debido a que \mathbb{S} es simétrica definida positiva y \mathbb{A} es simétrica definida semipositiva, entonces $\mathbb{S}_h^{-1} \mathbb{A}_h$ es simétrica definida semipositiva, y sus valores propios son reales y no negativos.

Sea $\lambda_i \geq 0$ un valor propio de $\mathbb{S}_h^{-1} \mathbb{A}_h$, y sea $\mathbf{\Lambda}_i$ su correspondiente vector propio. Entonces por (6.25), se obtiene

$$\mathbf{\Lambda}_i^{n+1} = (1 - \Delta t \lambda_i) \mathbf{\Lambda}_i^n, \quad \forall n.$$

Procediendo recursivamente, se obtiene

$$\mathbf{\Lambda}_i^n = (1 - \Delta t \lambda_i)^n \mathbf{\Lambda}_i^0, \quad \forall n.$$

Por lo tanto, el modo permanecerá acotado, para toda n , si

$$|1 - \Delta t \lambda_i| \leq 1, \quad \text{o equivalentemente, si } \Delta t \leq 2/\lambda_i.$$

Este resultado lo podemos resumir en el siguiente enunciado:

El método explícito de Euler es estable si el paso de discretización en el tiempo satisface la condición

$$\Delta t \leq \frac{2}{\lambda_{\max}(\mathbb{S}_h^{-1} \mathbb{A}_h)}. \quad (6.26)$$

Debido a esto, se dice que el método explícito es **condicionalmente estable**.

Por supuesto que esta condición podría ser demasiado restrictiva para ciertos problemas, en los cuales el paso de discretización se escogería muy pequeño con el fin de mantener la estabilidad. Esta es la principal razón por la cual los métodos explícitos no son los más populares para resolver numéricamente ecuaciones diferenciales.

Estabilidad del método implícito de Euler. Haciendo $\mathbf{b} = 0$ en (6.24), obtenemos

$$(I + \Delta t \mathbb{S}_h^{-1} \mathbb{A}_h) \mathbf{U}^{n+1} = \mathbf{U}^n. \quad (6.27)$$

Si disponemos de los valores propios de la matriz $\mathbb{S}_h^{-1} \mathbb{A}_h$, y estos son $\lambda_1, \lambda_2, \dots, \lambda_q$, con sus correspondientes vectores propios $\mathbf{\Lambda}_1, \mathbf{\Lambda}_2, \dots, \mathbf{\Lambda}_q$, entonces para cada uno de ellos se satisface

$$(1 + \Delta t \lambda_i) \mathbf{\Lambda}_i^{n+1} = \mathbf{\Lambda}_i^n, \quad \forall n.$$

Procediendo recursivamente y despejando, se obtiene

$$\mathbf{\Lambda}_i^n = \left(\frac{1}{1 + \Delta t \lambda_i} \right)^n \mathbf{\Lambda}_i^0, \quad \forall n. \quad (6.28)$$

Como $\lambda_i \geq 0$, se obtienen los siguientes casos

$$\begin{aligned} \mathbf{\Lambda}_i^n &= \mathbf{\Lambda}_i^0 & \text{si } \lambda_i &= 0, \\ \lim_{n \rightarrow \infty} \mathbf{\Lambda}_i^n &\rightarrow \mathbf{0}, & \text{si } \lambda_i &> 0. \end{aligned}$$

En consecuencia, la solución \mathbf{U}^n del sistema homogéneo en (6.24) es acotada para toda n , independientemente del valor del parámetro de discretización $\Delta t > 0$. Por esta razón se dice que el esquema es **incondicionalmente estable**. Por lo tanto, la elección del parámetro de discretización en el tiempo esta determinada únicamente por el grado de precisión deseado en la solución numérica y no por restricciones de estabilidad. Los métodos implícitos generalmente se prefieren sobre los implícitos principalmente en problemas rígidos ó cuando se buscan soluciones estacionarias.

6.6. Análisis de estabilidad por medio de métodos de energía

Una forma alternativa de verificar la estabilidad de los esquemas numéricos es mediante técnicas de estimaciones de energía. La idea central se basa en la observación de que la matriz \mathbb{S}_h , al ser simétrica y definida positiva, induce una norma:

$$\|\mathbf{U}\|_S^2 = (\mathbb{S}_h \mathbf{U}) \cdot \mathbf{U}. \quad (6.29)$$

Para ilustrar como podemos utilizar esta propiedad en el análisis de estabilidad, consideremos el método implícito de Euler (6.23) con $\mathbf{b} = \mathbf{0}$.

Multiplicando la ecuación (6.23) escalarmente por \mathbf{U}^{n+1} , obtenemos

$$\left(\mathbb{S}_h \frac{\mathbf{U}^{n+1} - \mathbf{U}^n}{\Delta t} \right) \cdot \mathbf{U}^{n+1} + (\mathbb{A} \mathbf{U}^{n+1}) \cdot \mathbf{U}^{n+1} = 0.$$

De esta ecuación y de (6.29), obtenemos

$$\frac{1}{\Delta t} \|\mathbf{U}^{n+1}\|_S^2 + (\mathbb{A} \mathbf{U}^{n+1}) \cdot \mathbf{U}^{n+1} = \frac{1}{\Delta t} (\mathbb{S} \mathbf{U}^n) \cdot \mathbf{U}^{n+1}.$$

Puesto que $(\mathbb{A} \mathbf{U}^{n+1}) \cdot \mathbf{U}^{n+1} \geq 0$, entonces podemos escribir la siguiente desigualdad

$$\frac{1}{\Delta t} \|\mathbf{U}^{n+1}\|_S^2 \leq \frac{1}{\Delta t} (\mathbb{S} \mathbf{U}^n) \cdot \mathbf{U}^{n+1} \leq \frac{1}{\Delta t} \|\mathbf{U}^n\|_S \|\mathbf{U}^{n+1}\|_S,$$

que es una forma de la desigualdad de Cauchy-Schwartz. De esta última desigualdad se obtiene

$$\|\mathbf{U}^{n+1}\|_S \leq \|\mathbf{U}^n\|_S.$$

Procediendo recursivamente, obtenemos

$$\|\mathbf{U}^n\|_S \leq \|\mathbf{U}^0\|_S. \quad \forall n,$$

lo cual demuestra que la solución numérica permanece acotada para toda n , independientemente del tamaño del paso de discretización Δt .

6.7. Ejercicios

Ejercicio 6.1. Escribe el problema parabólico incluyendo la ecuación diferencial, las condiciones de frontera y las condiciones iniciales. Indica las condiciones que deben satisfacer el coeficientes de difusión a y de densidad ρ .

Ejercicio 6.2. Encuentra el problema variacional del problema parabólico y la aproximación de elemento finito para obtener el problema de valores iniciales (6.14)–(6.15).

Ejercicio 6.3. Describe los métodos explícito e implícito de Euler para integrar el sistema de ecuaciones (6.14)–(6.15) e indica cuales son sus diferencias.

Ejercicio 6.4. ¿Qué se entiende por estabilidad de un método numérico para resolver un problema dependiente del tiempo? ¿Cuáles son la propiedades de estabilidad de los métodos explícito e implícito de Euler?