

# Algunos Aspectos Básicos del Método de Elementos Finitos

GABRIEL N. GATICA

*Departamento de Ingeniería Matemática, Universidad de Concepción,  
Casilla 160-C, Concepción, Chile, e-mail: ggatica@ing-mat.udec.cl*

**Resumen.** El propósito de este artículo de divulgación es proporcionar algunas ideas fundamentales sobre el método de elementos finitos (MEF) a través de su aplicación a problemas de valores de contorno en una y dos dimensiones. Se usan conceptos elementales de la teoría de distribuciones para describir las formulaciones débiles correspondientes. La existencia y unicidad de solución es deducida mediante el Lema de Lax-Milgram, el cual constituye una de las herramientas básicas de Análisis Funcional para el estudio y aproximación numérica de formulaciones variacionales. Siguiendo el procedimiento usual, el MEF se introduce como un caso particular del clásico método de Galerkin, y se proporcionan las estimaciones de error en diferentes normas mediante el Lema de Cea y argumentos de dualidad. También, se incluyen algunos aspectos algorítmicos del método para los problemas modelos considerados.

## Contenidos

1. Introducción.
2. La formulación variacional.
3. Existencia y unicidad de solución.
4. El método de Galerkin.
5. Estimaciones del error.
  - 5.1 El Lema de Cea.
  - 5.2 Cotas de error en normas más débiles.
  - 5.3 Una propiedad de aproximación.
6. Algunos elementos de análisis funcional.
  - 6.1 Preliminares.
  - 6.2 Lema de Lax-Milgram.
7. La ecuación de Poisson.
  - 7.1 Identidades de Green.
  - 7.2 Un problema de Dirichlet.
  - 7.3 Un método de elementos finitos.
  - 7.4 El caso no-homogéneo.

# 1 Introducción

El Método de Elementos Finitos (MEF) es un procedimiento numérico que permite obtener una aproximación de la solución de una ecuación diferencial, ordinaria o parcial, bajo ciertas condiciones iniciales y de frontera. El MEF es un caso particular del clásico método de Galerkin, el cual consiste, básicamente, en aproximar la solución de una ecuación de operadores en un espacio de dimensión infinita, por la solución de la misma ecuación, pero ahora planteada sobre un subespacio de dimensión finita del espacio original. Lo que caracteriza e identifica al MEF de cualquier otro esquema de Galerkin es el hecho de elegir funciones seccionalmente polinomiales para definir el subespacio respectivo. Originalmente, el método fue propuesto por ingenieros en la década del 60 para resolver el problema de elasticidad lineal, y durante sus inicios se utilizó prácticamente sólo para problemas elípticos, es decir, en términos simples, para ecuaciones diferenciales dependientes de la variable espacial e independientes de la variable temporal. Posteriormente, los avances han sido tan numerosos e importantes que hoy en día se puede aplicar a una amplia gama de situaciones, incluyendo comportamientos lineales y no-lineales, y tanto para problemas estacionarios como para aquellos de carácter evolutivo (ver [ 1], [ 2], [ 3], [ 5], [ 6], [ 7] y las referencias indicadas ahí). Además de lo anterior, el MEF ha mostrado una gran capacidad para combinarse con otros procedimientos similares, por ejemplo el Método de Elementos de Frontera, lo cual permite resolver una clase más amplia de problemas diferenciales en física y en ciencias de la ingeniería (ver [ 4]).

El objetivo del presente trabajo es proporcionar las ideas esenciales del MEF mediante la resolución de algunos problemas simples de valores de contorno. Como primer modelo físico se considera una barra de longitud 1, fija en sus extremos, y sujeta a una carga tangencial  $f$ . Dado  $x \in [0, 1]$ , denotamos por  $\sigma(x)$  y  $u(x)$ , la tracción y desplazamiento tangencial en  $x$ , respectivamente, bajo la carga  $f$  (ver Fig. 1.1).

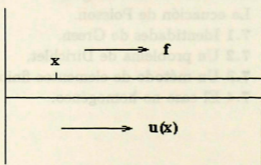


Figura 1.1

Suponiendo que ocurren pequeños desplazamientos y que el material del cual está hecha la barra es lineal-elástico, se tiene:

$$\begin{aligned} \sigma &= E u' \quad \text{en } (0,1) \quad (\text{ley de Hooke}), \\ -\sigma' &= f \quad \text{en } (0,1) \quad (\text{ecuación de equilibrio}), \\ u(0) &= u(1) = 0 \quad (\text{condiciones de frontera}) \end{aligned} \quad (1.1)$$

donde  $E$  es el módulo de elasticidad. Al eliminar la incógnita  $\sigma$ , el sistema (1.1) se reduce a:

$$\begin{aligned} -(E u')' &= f \quad \text{en } (0,1), \\ u(0) &= u(1) = 0. \end{aligned}$$

En particular, si  $E$  es constante e igual a 1, obtenemos el siguiente problema de valores de contorno: *Hallar  $u$  tal que*

$$\begin{aligned} -u'' &= f \quad \text{en } \Omega := (0,1), \\ u(0) &= u(1) = 0. \end{aligned} \quad (1.2)$$

Otra situación que da origen al problema (1.2) se describe a continuación. Sea  $u$  la temperatura y  $q$  el flujo de calor en una barra de longitud 1 sujeta a una fuente de intensidad  $f$ . Si la temperatura es cero en los extremos de la barra, entonces en el caso estacionario se obtiene lo siguiente:

$$\begin{aligned} -q &= k u' \quad (\text{ley de Fourier}), \\ q' &= f \quad (\text{conservación de la energía}), \\ u(0) &= u(1) = 0, \end{aligned} \quad (1.3)$$

donde  $k$  es la conductividad. Similarmente al caso anterior, es fácil ver que si  $k = 1$  entonces la temperatura  $u$  también satisface (1.2).

El resto del artículo se presenta como sigue. En la Sección 2 se deduce la formulación variacional del problema modelo unidimensional (1.2). La existencia y unicidad de solución de esta formulación es estudiada en la Sección 3. El método de Galerkin y un esquema de elementos finitos usando funciones seccionalmente polinomiales de grado  $\leq 1$  constituyen el tema de la Sección 4. En la Sección 5 se proporcionan estimaciones del error en dos normas diferentes y se prueba un resultado de aproximación para el espacio de elementos finitos utilizado en la Sección anterior. En la Sección 6 se recopilan algunos resultados elementales de Análisis Funcional y se demuestra el Lema de Lax-Milgram. Finalmente, la Sección 7 estudia la aplicación del método de elementos finitos al problema de Dirichlet para la ecuación de Poisson. Las identidades de Green se utilizan para obtener la formulación variacional respectiva, y se ilustran en detalle los aspectos algorítmicos del método para un caso bi-dimensional.

## 2 La formulación variacional

Con el objeto de deducir una formulación variacional de (1.2), necesitamos introducir algunos conceptos previos. Dado un abierto  $\Omega$  de  $\mathbf{R}$ ,  $C^\infty(\Omega)$  denota el espacio de funciones infinitamente diferenciables definidas sobre  $\Omega$ , a valores reales. Para cada  $\varphi \in C^\infty(\Omega)$  se define su soporte, y se denota  $\text{sop } \varphi$ , por

$$\text{sop } \varphi := \overline{\{x \in \Omega : \varphi(x) \neq 0\}}.$$

Además, se introduce el espacio de funciones

$$C_0^\infty(\Omega) := \{\varphi \in C^\infty(\Omega) : \text{sop } \varphi \text{ es compacto y } \text{sop } \varphi \subseteq \Omega\}.$$

Luego, multiplicando la ecuación diferencial de (1.2) por una función  $\varphi \in C_0^\infty(\Omega)$  e integrando por partes en  $\Omega$ , resulta

$$\int_{\Omega} f \varphi \, dx = - \int_{\Omega} u'' \varphi \, dx = \int_{\Omega} u' \varphi' \, dx - u' \varphi \Big|_0^1,$$

de donde, usando que  $\varphi(0) = \varphi(1) = 0$ , se obtiene la identidad

$$\int_{\Omega} u' \varphi' \, dx = \int_{\Omega} f \varphi \, dx \quad \forall \varphi \in C_0^\infty(\Omega). \quad (2.1)$$

**DEFINICION 2.1.** Dada  $v \in L^2(\Omega)$ , se dice que  $v' \in L^2(\Omega)$ , en el sentido distribucional, si existe  $z \in L^2(\Omega)$  tal que

$$- \int_{\Omega} v \varphi' \, dx = \int_{\Omega} z \varphi \, dx \quad \forall \varphi \in C_0^\infty(\Omega),$$

y se escribe  $v' := z$ .

**EJEMPLO 2.2.** Considere la función:

$$v(x) = \begin{cases} x, & 0 \leq x \leq 1, \\ 1, & 1 \leq x \leq 2. \end{cases}$$

Es fácil probar que  $v' \in L^2(\Omega)$ , donde

$$v'(x) = \begin{cases} 1, & 0 \leq x < 1, \\ 0, & 1 < x \leq 2. \end{cases}$$

En virtud de la definición anterior, podemos introducir el *espacio de Sobolev*,

$$H^1(\Omega) := \{v \in L^2(\Omega) : v' \in L^2(\Omega)\}, \quad (2.2)$$

el cual es un espacio de Hilbert (ver [7, Teorema 1.2-1]) provisto del siguiente producto escalar

$$\langle v, w \rangle_{H^1(\Omega)} := \int_{\Omega} (v' w' + v w) dx \quad \forall v, w \in H^1(\Omega). \quad (2.3)$$

Se sigue que la norma de  $H^1(\Omega)$  está dada por

$$\begin{aligned} \|\cdot\|_{H^1(\Omega)} : H^1(\Omega) &\rightarrow \mathbf{R} \\ \|v\|_{H^1(\Omega)} &:= \left\{ \int_{\Omega} ((v')^2 + v^2) dx \right\}^{1/2} \quad \forall v \in H^1(\Omega). \end{aligned}$$

Además, denotamos por  $H_0^1(\Omega)$  la adherencia de  $C_0^\infty(\Omega)$  en  $H^1(\Omega)$ . Puede mostrarse que  $H_0^1(\Omega)$  también se caracteriza como el subespacio de elementos de  $H^1(\Omega)$  que se anulan en la frontera  $\partial\Omega$  en el sentido de trazas (ver [7, Sección 1.3]). Es decir, formalmente,

$$H_0^1(\Omega) := \{v \in H^1(\Omega) : v(0) = v(1) = 0\}. \quad (2.4)$$

De acuerdo a la identidad (2.1) y a la densidad de  $C_0^\infty(\Omega)$  en  $H_0^1(\Omega)$ , podemos pensar en un método de solución para (1.2), distinto del esquema clásico, y en el cual en vez de pedir que la ecuación diferencial se satisfaga en cada punto  $x \in \Omega$ , se requiera solamente que se verifique la identidad integral (2.1) para cada  $\varphi \in H_0^1(\Omega)$ . Esto da origen al llamado *método débil*, el cual se reduce a la siguiente formulación: Hallar  $u \in H_0^1(\Omega)$  tal que

$$\int_{\Omega} u' v' dx = \int_{\Omega} f v dx \quad \forall v \in H_0^1(\Omega). \quad (2.5)$$

### 3 Existencia y unicidad de solución

Para estudiar la existencia y unicidad de solución de (2.5) se introducen primero algunas notaciones. Para este efecto, sea  $V := H_0^1(\Omega)$  y definamos

$$\begin{aligned} A : V \times V &\rightarrow \mathbf{R} \\ A(w, v) &:= \int_{\Omega} w' v' dx \quad \forall w, v \in V, \end{aligned} \quad (3.1)$$

$$F : V \rightarrow \mathbf{R}$$

$$F(v) := \int_{\Omega} f v \, dx \quad \forall v \in V. \quad (3.2)$$

Con esto, la formulación débil (2.5) puede reescribirse como sigue: *Hallar  $u \in V$  tal que*

$$A(u, v) = F(v) \quad \forall v \in V. \quad (3.3)$$

Es fácil ver que  $A$  es una *forma bilineal*, esto es lineal en cada componente, y que  $F$  es un funcional lineal. Además, utilizando la desigualdad de Cauchy-Schwarz en  $L^2(\Omega)$ , se obtiene

$$|A(w, v)| \leq \|w'\|_{L^2(\Omega)} \|v'\|_{L^2(\Omega)} \leq \|w\|_V \|v\|_V \quad \forall w, v \in V,$$

lo cual prueba que  $A$  es acotada. En general, se dice que una forma bilineal  $B$  es *acotada* sobre un Hilbert  $W$  si existe una constante  $M > 0$  tal que

$$|B(w, v)| \leq M \|w\|_W \|v\|_W \quad \forall w, v \in W. \quad (3.4)$$

También, se dice que  $B$  es *coerciva* sobre  $W$ , o  *$W$ -elíptica* (ver [ 3]), si existe una constante  $\alpha > 0$  tal que

$$B(v, v) \geq \alpha \|v\|_W^2 \quad \forall v \in W. \quad (3.5)$$

Ahora, usando de nuevo la desigualdad de Cauchy-Schwarz en  $L^2(\Omega)$ , resulta

$$|F(v)| \leq \|f\|_{L^2(\Omega)} \|v\|_{L^2(\Omega)} \leq \|f\|_{L^2(\Omega)} \|v\|_V,$$

lo cual muestra que  $F$  es acotado. En general, se dice que un funcional lineal  $G$  es *acotado* sobre un Hilbert  $W$ , y por lo tanto un elemento del dual  $W'$ , si existe una constante positiva, que denotamos  $\|G\|_{W'}$ , tal que

$$|G(v)| \leq \|G\|_{W'} \|v\|_V \quad \forall v \in W. \quad (3.6)$$

Con el propósito de mostrar que nuestra forma bilineal  $A$  es  $V$ -elíptica, se necesita el siguiente resultado (ver [ 7, Teorema 1.2-5]).

**Lema 3.1** *Definamos la aplicación  $|\cdot|_{H^1(\Omega)} : H^1(\Omega) \rightarrow \mathbf{R}$ ,*

$$|v|_{H^1(\Omega)} := \left\{ \int_{\Omega} (v')^2 \, dx \right\}^{1/2} \quad \forall v \in H^1(\Omega). \quad (3.7)$$

Entonces  $|\cdot|_{H^1(\Omega)}$  es una seminorma sobre  $H^1(\Omega)$ , la cual es equivalente a la norma usual  $\|\cdot\|_{H^1(\Omega)}$  en el subespacio  $H_0^1(\Omega)$ . Es decir, existe  $C > 0$  tal que

$$|v|_{H^1(\Omega)} \leq \|v\|_{H^1(\Omega)} \leq C |v|_{H^1(\Omega)} \quad \forall v \in H_0^1(\Omega). \quad (3.8)$$

*Demostración.* Basta mostrar la equivalencia de  $|\cdot|_{H^1(\Omega)}$  con  $\|\cdot\|_{H^1(\Omega)}$  en el subespacio denso de  $H_0^1(\Omega)$  dado por  $C_0^\infty(\Omega)$ . En efecto, dado  $\varphi \in C_0^\infty(\Omega)$  y  $x \in \Omega := (0, 1)$ , tenemos

$$|\varphi(x)|^2 = \left| \int_0^x \varphi'(t) dt \right|^2 \leq x \int_0^x (\varphi'(t))^2 dt \leq x |\varphi|_{H^1(\Omega)}^2,$$

donde la desigualdad de Cauchy-Schwarz se utilizó en la segunda estimación. Integrando lo anterior con respecto a  $x$  en  $\Omega$ , se deduce que

$$\|\varphi\|_{L^2(\Omega)}^2 \leq \frac{1}{2} |\varphi|_{H^1(\Omega)}^2$$

y por lo tanto

$$\|\varphi\|_{H^1(\Omega)}^2 = \|\varphi\|_{L^2(\Omega)}^2 + |\varphi|_{H^1(\Omega)}^2 \leq \frac{3}{2} |\varphi|_{H^1(\Omega)}^2.$$

Esto completa la demostración. ■

La  $V$ -elipticidad de  $A$  se sigue entonces de (3.1), (3.5), (3.7) y (3.8).

De acuerdo al análisis anterior, la existencia de una única solución para nuestro problema (3.3) se concluye de manera inmediata a partir del siguiente resultado general (ver [3], [6], [7, Teorema 2.2-1]), cuya demostración será proporcionada en la Sección 6.2.

**Teorema 3.2** (Lema de Lax-Milgram). *Sea  $(V, \langle \cdot, \cdot \rangle_V)$  un Hilbert y sea  $A : V \times V \rightarrow \mathbf{R}$  una forma bilineal acotada y  $V$ -elíptica. Entonces, para cada  $F \in V'$ , existe un único  $u \in V$  tal que*

$$A(u, v) = F(v) \quad \forall v \in V.$$

Además,

$$\|u\|_V \leq \frac{1}{\alpha} \|F\|_{V'},$$

donde  $\alpha > 0$  es la constante de  $V$ -elipticidad de  $A$ .

## 4 El método de Galerkin

En lo que sigue podemos abstraernos del problema de valores de contorno (1.2) y de su formulación débil asociada (3.3), e introducir así un espacio de Hilbert arbitrario  $V$ , una forma bilineal  $A : V \times V \rightarrow \mathbf{R}$  acotada y  $V$ -elíptica, y un funcional acotado  $F \in V'$ . Aplicando entonces el Lema de Lax-Milgram se deduce que: *existe un único  $u \in V$  tal que*

$$A(u, v) = F(v) \quad \forall v \in V. \quad (P)$$

Para deducir un procedimiento que nos permita aproximar la solución de (P), se considera primero una sucesión de subespacios de dimensión finita  $V_n \subseteq V$ , y luego se plantean los siguientes problemas: *Hallar  $u_n \in V_n$  tal que*

$$A(u_n, v_n) = F(v_n) \quad \forall v_n \in V_n. \quad (P_n)$$

Sin pérdida de generalidad, podemos suponer que la dimensión de  $V_n$  es  $n$ . Esta simple idea de reducir el problema original formulado en el espacio de dimensión infinita  $V$ , a una sucesión de problemas análogos en los subespacios de *dimensión finita*  $V_n$ , es lo que se conoce con el nombre de *Método de Galerkin*. El objetivo será, entonces, elegir los  $V_n$  de manera tal que al menos se cumpla la siguiente propiedad

$$\lim_{n \rightarrow \infty} \|u - u_n\|_V = 0. \quad (C)$$

A continuación mostramos que (P<sub>n</sub>) se reduce a un sistema lineal de ecuaciones. En efecto, sea  $\{e_1, e_2, \dots, e_n\}$  una base de  $V_n$ . Entonces, existen escalares  $\alpha_1, \alpha_2, \dots, \alpha_n$  tales que

$$u_n = \sum_{j=1}^n \alpha_j e_j.$$

Luego, (P<sub>n</sub>) se convierte en: *Hallar  $\alpha_1, \dots, \alpha_n \in \mathbf{R}$  tales que*

$$\sum_{j=1}^n \alpha_j A(e_j, v_n) = F(v_n) \quad \forall v_n \in V_n,$$

lo cual es equivalente a: *Hallar  $\alpha_1, \dots, \alpha_n \in \mathbf{R}$  tales que*

$$\sum_{j=1}^n \alpha_j A(e_j, e_i) = F(e_i) \quad \forall i \in \{1, \dots, n\}. \quad (4.1)$$

Así, definiendo la matriz  $\mathbf{A} := (a_{ij})_{n \times n}$  y los vectores  $\alpha := (\alpha_j)_{n \times 1}$  y  $\mathbf{F} := (f_j)_{n \times 1}$ , con

$$a_{ij} := A(e_j, e_i) \quad \text{y} \quad f_j := F(e_j),$$

la formulación (4.1) se reescribe como sigue: *Hallar  $\alpha \in \mathbf{R}^n$  tal que*



$$\mathbf{A} \alpha = \mathbf{F} . \quad (4.2)$$

Notemos aquí que la  $V$ -elipticidad de  $A$  garantiza también que (4.2) siempre tiene una única solución. En la literatura,  $\mathbf{A}$  se llama usualmente *matriz de rigidez* y  $\mathbf{F}$ , el *vector de carga*. Estos nombres se deben a la terminología usada inicialmente por los ingenieros, quienes crearon un caso particular del método de Galerkin para resolver el problema de elasticidad lineal, conocido como el *método de elementos finitos*.

Ahora, es preciso señalar que el diseño de algoritmos eficientes para resolver el sistema lineal (4.2) adquiere una tremenda importancia cuando la dimensión del subespacio  $V_n$  tiende a  $\infty$ . Al respecto, debemos observar que para matrices de gran tamaño, una estructura *rala* o *banda* es muy deseable. También, si la matriz  $\mathbf{A}$  es simétrica y definida positiva entonces la resolución de (4.2) se facilita aún más. Este es precisamente el caso cuando la forma bilineal  $A$ , además de ser acotada y  $V$ -elíptica, resulta *simétrica*. En efecto, dado  $\beta := (\beta_1, \beta_2, \dots, \beta_n)^T \in \mathbf{R}^n$ , definamos  $v := \sum_{j=1}^n \beta_j e_j \in V_n \subseteq V$ . En virtud de la  $V$ -elipticidad de  $A$  se sigue que

$$\alpha \|v\|_V^2 \leq A(v, v) = \sum_{i=1}^n \sum_{j=1}^n a_{ij} \beta_i \beta_j = \beta^T \mathbf{A} \beta ,$$

de donde  $\beta^T \mathbf{A} \beta \geq 0 \forall \beta \in \mathbf{R}^n$ , y  $\beta^T \mathbf{A} \beta = 0$  si y sólo si  $\beta$  es el vector nulo.

**EJEMPLO 4.1.** Para mostrar un ejemplo específico de un esquema de Galerkin asociado a (1.2), consideramos una partición  $0 = x_0 < x_1 < x_2 < \dots < x_n < x_{n+1} = 1$  del dominio  $\Omega := (0, 1)$ , e introducimos el espacio

$$V_n := \{ v \in C(\bar{\Omega}) : v(0) = v(1) = 0 \text{ y } v|_{[x_{j-1}, x_j]} \in P_1([x_{j-1}, x_j]) \forall j = \overline{1, n+1} \} ,$$

donde  $P_1(S)$  denota el espacio de polinomios de grado  $\leq 1$  definidos sobre cualquier subconjunto  $S$  de  $\Omega$ . Es fácil probar que las funciones de  $V_n$  quedan determinadas únicamente por los valores que toman en los nodos  $x_j$  de la partición de  $\Omega$ . En particular,  $\{e_1, e_2, \dots, e_n\}$  constituye una base de  $V_n$ , donde  $e_i \in V_n$  es tal que

$$e_i(x_j) = \begin{cases} 1 & \text{si } j = i , \\ 0 & \text{si } j \neq i . \end{cases}$$

Tales funciones base de  $V_n$  se llaman *funciones techo*. En cuanto al cálculo de la matriz de rigidez  $\mathbf{A}$ , se obtiene

$$a_{ij} := A(e_i, e_j) = \int_0^1 e'_i e'_j dx = \begin{cases} 0 & \text{si } |i - j| \geq 2 , \\ -\frac{1}{h_j} & \text{si } j = i - 1 , \\ \frac{1}{h_j} + \frac{1}{h_{j+1}} & \text{si } j = i , \end{cases}$$

donde  $h_j := x_j - x_{j-1} \forall j = \overline{1, n+1}$ , y por lo tanto  $\mathbf{A}$  resulta tridiagonal. En el caso de una partición uniforme con  $h_j = h := \frac{1}{n+1}$ , el sistema lineal (4.2) queda dado por

$$\frac{1}{h} \begin{bmatrix} 2 & -1 & 0 & \cdot & \cdot & \cdot & 0 \\ -1 & 2 & -1 & \cdot & \cdot & \cdot & 0 \\ 0 & & & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ \cdot & & & & & & \\ & & & & & & 0 \\ 0 & \cdot & \cdot & \cdot & -1 & 2 & -1 \\ 0 & \cdot & \cdot & \cdot & 0 & -1 & 2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \alpha_{n-1} \\ \alpha_n \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ f_{n-1} \\ f_n \end{bmatrix}, \quad (4.3)$$

donde  $f_j := F(e_j) = \int_{x_{j-1}}^{x_j} f e_j dx$ . Notemos finalmente que, debido a la definición del subespacio  $V_n$ , la solución  $\alpha_1, \alpha_2, \dots, \alpha_n$  que entrega el sistema lineal (4.3) coincide con los valores que toma la solución de Galerkin  $u_n$  en cada uno de los nodos de la partición. En otras palabras, se tiene que  $\alpha_i = u_n(x_i) \forall i \in \{1, \dots, n\}$ .

Lo anterior constituye una de las situaciones más simples posibles que muestran el tipo de sistemas lineales que se obtienen al aplicar el método de Galerkin al problema  $(P)$ .

El procedimiento ilustrado a través del EJEMPLO 4.1 representa la misma idea que motivó a los ingenieros en la década del 60 a crear el *Método de Elementos Finitos* (MEF) para resolver la formulación débil del problema de elasticidad. Más precisamente, el MEF es un caso particular del método de Galerkin, en el cual se utilizan subespacios de dimensión finita  $V_n$  cuyos elementos son funciones seccionalmente polinomiales. Algunas de las ventajas, propiedades y aplicaciones de esta técnica constituyen el tema de las restantes secciones de este trabajo.

## 5 Estimaciones del error

### 5.1 El Lema de Cea

Nuestro propósito siguiente es obtener una cota superior del *error* que se origina mediante el método de Galerkin. Este se define como la diferencia en norma entre la solución exacta  $u \in V$  de  $(P)$  y la solución aproximada  $u_n \in V_n$  de  $(P_n)$ . Para ello, vemos primero que de  $(P)$  y  $(P_n)$  se deduce que

$$A(u - u_n, v_n) = 0 \quad \forall v_n \in V_n, \quad (5.1)$$

la cual se conoce como *relación de ortogonalidad*. La razón de este nombre se debe a lo que acontece cuando  $A$  es simétrica, en cuyo caso  $A$  se convierte en un producto interior donde su norma asociada, definida por  $\|v\|_E := A(v, v)^{1/2} \forall v \in V$ , y llamada *norma de la energía*, es equivalente a la norma original del espacio  $V$ . Así, la ecuación (5.1) es la típica condición de ortogonalidad que caracteriza la mejor aproximación de  $u \in V$  por elementos del subespacio  $V_n$ , con respecto a  $\|\cdot\|_E$ . En otras palabras, (5.1) muestra que, en el caso simétrico, la solución de Galerkin  $u_n$  es la mejor aproximación de la solución exacta  $u$  con respecto a la norma de la energía, o equivalentemente,

$$\|u - u_n\|_E = \inf_{v_n \in V_n} \|u - v_n\|_E. \quad (5.2)$$

No obstante lo anterior, es importante reiterar que la *relación de ortogonalidad* (5.1) también se satisface para  $A$  no-simétrica.

**Lema 5.1** (Lema de Cea [3]) *Suponga que  $A$  es una forma bilineal acotada (con constante  $M$ ) y  $V$ -elíptica (con constante  $\alpha$ ). Sean  $u \in V$  y  $u_n \in V_n$  las soluciones de  $(P)$  y  $(P_n)$ , respectivamente. Entonces*

$$\|u - u_n\|_V \leq \frac{M}{\alpha} \inf_{v_n \in V_n} \|u - v_n\|_V. \quad (5.3)$$

*Demostración.* Usando la  $V$ -elipticidad de  $A$ , sumando y restando un elemento arbitrario  $v_n \in V_n$  en el segundo término de  $A$ , y luego aplicando la relación de ortogonalidad (5.1), obtenemos

$$\alpha \|u - u_n\|_V^2 \leq A(u - u_n, u - u_n) = A(u - u_n, u - v_n + v_n - u_n) = A(u - u_n, u - v_n).$$

Puesto que  $A$  es acotada, la desigualdad anterior implica que

$$\alpha \|u - u_n\|_V^2 \leq M \|u - u_n\|_V \|u - v_n\|_V,$$

de donde, suponiendo  $\|u - u_n\|_V \neq 0$ , resulta

$$\|u - u_n\|_V \leq \frac{M}{\alpha} \|u - v_n\|_V \quad \forall v_n \in V_n. \quad (5.4)$$

Naturalmente, (5.4) es válida también para  $\|u - u_n\|_V = 0$ . Finalmente, tomando ínfimo en (5.4) con respecto a  $v_n \in V_n$  se concluye la demostración. ■

La cota de error dada por (5.3) se conoce con el nombre de *estimación de Cea* y ella provee el punto de partida para finalmente derivar las propiedades de convergencia de los métodos de elementos finitos. Como se mencionó antes, lo mínimo que se requerirá en cada caso es la condición  $(C)$ , la cual garantiza que el error se va a cero a medida que la dimensión del subespacio  $V_n$  tiende a  $\infty$ .

## 5.2 Cotas de error en normas más débiles

El error también puede medirse en la norma de un espacio  $W$  que contenga a  $V$ . Esto se hace utilizando lo que se conoce como *argumentos de dualidad*. Para ejemplificar esta idea, sea  $W := L^2(\Omega)$ , y consideremos nuevamente el problema unidimensional (1.2). Sea  $w \in V := H_0^1(\Omega)$  la solución débil del problema de valores de contorno:

$$\|u - u_n\|_V \leq \frac{M}{\alpha} \|u - v_n\|_V \quad \forall v_n \in V_n. \quad (5.4)$$

$$\begin{aligned} -w'' &= u - u_n \quad \text{en } \Omega := (0,1), \\ w(0) &= w(1) = 0. \end{aligned} \quad (5.5)$$

Se sigue que

$$A(v, w) = \int_{\Omega} (u - u_n) v \, dx \quad \forall v \in V,$$

y, en particular, para  $v := u - u_n$ , se obtiene

$$\|u - u_n\|_W^2 = A(u - u_n, w).$$

Ahora, usando la relación de ortogonalidad y el acotamiento de  $A$ , resulta

$$\|u - u_n\|_W^2 = A(u - u_n, w - v_n) \leq M \|u - u_n\|_V \|w - v_n\|_V \quad \forall v_n \in V_n,$$

lo cual implica que

$$\|u - u_n\|_W^2 \leq M \|u - u_n\|_V \inf_{v_n \in V_n} \|w - v_n\|_V. \quad (5.6)$$

Si definimos el subespacio  $Z$  de  $V$  dado por aquellas funciones  $z \in V$  tales que  $z'' \in W$ , es razonable suponer que  $V_n$  satisface la siguiente *propiedad de aproximación*:

$$\inf_{v_n \in V_n} \|z - v_n\|_V \leq \epsilon(n) \|z''\|_W \quad \forall z \in Z, \quad (5.7)$$

donde  $\epsilon(n) \rightarrow 0$  cuando  $n \rightarrow \infty$ . Puesto que  $w \in Z$  se deduce de (5.7) que

$$\inf_{v_n \in V_n} \|w - v_n\|_V \leq \epsilon(n) \|w''\|_W = \epsilon(n) \|u - u_n\|_W,$$

lo cual, junto a (5.6), nos da

$$\|u - u_n\|_W \leq M \epsilon(n) \|u - u_n\|_V.$$

Aplicando la estimación de Cea (5.3) y asumiendo que  $u \in Z$ , se concluye de (5.7) que

$$\|u - u_n\|_V \leq \frac{M}{\alpha} \epsilon(n) \|u''\|_W,$$

o bien,

$$\|u - u_n\|_V \leq \frac{M}{\alpha} \epsilon(n) \|f\|_W,$$

y en consecuencia,

$$\|u - u_n\|_W \leq \frac{M^2}{\alpha} \epsilon(n)^2 \|f\|_W.$$

El principal comentario aquí es notar que  $\|u - u_n\|_V$  es de orden  $\epsilon(n)$  mientras que  $\|u - u_n\|_W$  es de orden  $\epsilon(n)^2$ , lo cual indica que el error medido en la norma de  $W$  converge más rápido a cero cuando  $n \rightarrow \infty$ .

### 5.3 Una propiedad de aproximación

La teoría de interpolación en espacios de Sobolev (ver [7, Capítulo 4]) es la base para el estudio de las propiedades de aproximación de los métodos de elementos finitos. A modo de ejemplo, a continuación mostramos que el espacio de elementos finitos  $V_n$  introducido en el EJEMPLO 4.1 satisface la *propiedad de aproximación* (5.7). Para este efecto, se define el operador de interpolación  $I_n : C(\bar{\Omega}) \rightarrow V_n$  por

$$I_n v := \sum_{j=1}^n v(x_j) e_j \quad \forall v \in C(\bar{\Omega}).$$

Se tiene el siguiente resultado previo (ver [1, Capítulo 0]).

**Lema 5.2.** *Sea  $Z$  el subespacio de  $H^1(\Omega)$  dado por las funciones  $z$  tales que  $z'' \in L^2(\Omega)$ . Sea  $\epsilon(n) := \max_{1 \leq j \leq n+1} (x_j - x_{j-1})$ . Entonces existe una constante  $C > 0$ , independiente de  $n$ , tal que*

$$\|v - I_n v\|_E \leq C \epsilon(n) \|v''\|_{L^2(\Omega)} \quad \forall v \in Z. \quad (5.8)$$

*Demostración.* Dada  $v \in Z$  y  $j \in \{1, \dots, n+1\}$  definamos la función  $w_j : [0, 1] \rightarrow \mathbf{R}$  tal que

$$w_j(t) := (v - I_n v)(x_{j-1} + t(x_j - x_{j-1})) \quad \forall t \in [0, 1].$$

Puesto que  $(I_n v)(x_i) = v(x_i) \forall i \in \{1, \dots, n+1\}$ , se tiene que  $w_j(0) = w_j(1) = 0$ , y luego, por Teorema de Rolle, existe  $\xi \in (0, 1)$  tal que  $w'_j(\xi) = 0$ . Se sigue que

$$w'_j(t) = \int_{\xi}^t w''_j(s) ds,$$

de donde, aplicando la desigualdad de Cauchy-Schwarz, resulta

$$\begin{aligned} |w'_j(t)|^2 &\leq \left| \int_{\xi}^t 1 \cdot ds \right| \left| \int_{\xi}^t [w''_j(s)]^2 ds \right| = |t - \xi| \left| \int_{\xi}^t [w''_j(s)]^2 ds \right| \\ &\leq |t - \xi| \int_0^1 [w''_j(s)]^2 ds, \end{aligned}$$

o bien

$$|w'_j(t)|^2 \leq |t - \xi| \int_0^1 [w''_j(s)]^2 ds \quad \forall t \in (0, 1).$$

Integrando con respecto a  $t$  en  $(0, 1)$  se obtiene

$$\int_0^1 [w'_j(t)]^2 dt \leq C \int_0^1 [w''_j(s)]^2 ds, \quad (5.9)$$

donde  $C := \max_{0 < \xi < 1} \int_0^1 |t - \xi| dt$ . Notemos que  $(v - I_n v)'' = v''$  porque  $(I_n v)|_{[x_{j-1}, x_j]} \in P_1([x_{j-1}, x_j])$ . Entonces, haciendo los cambios de variable  $x := x_{j-1} + t(x_j - x_{j-1})$  y  $x := x_{j-1} + s(x_j - x_{j-1})$  en las integrales de (5.9), esta desigualdad se convierte en

$$\int_{x_{j-1}}^{x_j} [(v - I_n v)'(x)]^2 dx \leq C (x_j - x_{j-1})^2 \int_{x_{j-1}}^{x_j} [v''(x)]^2 dx. \quad (5.10)$$

Así, sumando sobre todos los subintervalos  $[x_{j-1}, x_j]$  en (5.10), y acotando la longitud  $(x_j - x_{j-1})$  por  $\epsilon(n)$ , se concluye que

$$\begin{aligned} \|v - I_n v\|_E^2 &:= \int_0^1 [(v - I_n v)'(x)]^2 dx \\ &\leq C \epsilon^2(n) \int_0^1 [v''(x)]^2 dx = C \epsilon^2(n) \|v''\|_{L^2(\Omega)}^2, \end{aligned}$$

lo que prueba (5.8) y completa la demostración. ■

Ahora, debido a la  $V$ -elipticidad de  $A$  y a (5.8) se deduce que

$$\|v - I_n v\|_V \leq C \|v - I_n v\|_E \leq C \epsilon(n) \|v''\|_W,$$

lo cual, junto al hecho que

$$\inf_{v_n \in V_n} \|v - v_n\|_V \leq \|v - I_n v\|_V,$$

implican la propiedad de aproximación (5.7).

## 6 Algunos elementos de análisis funcional

En lo que sigue se formaliza la definición de algunos conceptos de análisis funcional utilizados en las secciones anteriores, se demuestran algunos teoremas clásicos como el Teorema de Descomposición Ortogonal y el Teorema de Representación de Riesz, y por último se prueba el Lema de Lax-Milgram, ya enunciado en la Sección 3.

### 6.1 Preliminares

Sea  $(V, \langle \cdot, \cdot \rangle_V)$  un espacio de Hilbert real. Toda aplicación (o transformación)  $F : V \rightarrow \mathbf{R}$  se llama **funcional**. Note, por ejemplo, que la norma de cualquier espacio de Hilbert es un funcional.

**DEFINICION 6.1** *Un funcional  $F : V \rightarrow \mathbf{R}$  se dice lineal si  $F(\alpha v + \beta w) = \alpha F(v) + \beta F(w)$ , para todo  $\alpha, \beta \in \mathbf{R}$ ,  $v, w \in V$ .*

**DEFINICION 6.2** *Un funcional lineal  $F : V \rightarrow \mathbf{R}$  se dice acotado si existe una constante  $M > 0$  tal que  $|F(v)| \leq M \|v\|_V \forall v \in V$ .*

El espacio de todos los funcionales lineales y acotados definidos sobre  $V$  se llama **dual** de  $V$  y se denota  $V'$ . Al ínfimo de todas las constantes  $M > 0$  que satisfacen la definición anterior se llama **norma del funcional  $F$** , y se denota  $\|F\|_{V'}$ . Notar que

$$\|F\|_{V'} = \sup_{0 \neq v \in V} \frac{|F(v)|}{\|v\|_V}.$$

Usando la desigualdad de Cauchy-Schwarz es fácil probar que, dado  $x \in V$ , la aplicación que a cada  $v \in V$  asigna el número real  $\langle x, v \rangle_V$ , es un elemento de  $V'$ . El Teorema de Representación de Riesz, que se demuestra más adelante, constituye precisamente el recíproco de este hecho. En otras palabras, dicho teorema establece que todo funcional lineal acotado  $F$  definido sobre  $V$  admite un único representante  $x \in V$  a través de la relación:  $F(v) = \langle x, v \rangle_V \forall v \in V$ .

**DEFINICION 6.3** *Un subconjunto  $U$  de un espacio vectorial normado  $X$  se dice cerrado si para cada sucesión  $\{x_n\}_{n \in \mathbf{N}} \subseteq U$  tal que  $x_n \rightarrow x \in X$ , se tiene que  $x \in U$ .*

**Teorema 6.4.** *Sea  $U \neq \emptyset$  un subespacio cerrado de un Hilbert  $V$ , y sea  $x \in V$  tal que  $x \notin U$ . Entonces existe un único  $z \in U$  tal que*

$$\|x - z\|_V = \inf_{v \in U} \|x - v\|_V.$$

( $z$  se llama la mejor aproximación de  $x$  por elementos de  $U$ ).

*Demostración.* Sea  $d := \inf_{v \in U} \|x - v\|_V$ . Entonces existe una sucesión  $\{z_n\}_{n \in \mathbb{N}} \subseteq U$  tal que  $\|x - z_n\|_V \rightarrow d$  cuando  $n \rightarrow +\infty$ . Aplicando la ley del paralelogramo a  $x - z_n$  y  $x - z_m$ , con  $m, n \in \mathbb{N}$ , se tiene

$$\begin{aligned} & \| (x - z_n) + (x - z_m) \|_V^2 + \| (x - z_n) - (x - z_m) \|_V^2 \\ &= 2 \left\{ \|x - z_n\|_V^2 + \|x - z_m\|_V^2 \right\}, \end{aligned}$$

o bien

$$4 \left\| x - \frac{(z_n + z_m)}{2} \right\|_V^2 + \|z_n - z_m\|_V^2 = 2 \|x - z_n\|_V^2 + 2 \|x - z_m\|_V^2,$$

de donde

$$\|z_n - z_m\|_V^2 = 2 \|x - z_n\|_V^2 + 2 \|x - z_m\|_V^2 - 4 \left\| x - \frac{(z_n + z_m)}{2} \right\|_V^2.$$

Puesto que  $\frac{z_n + z_m}{2} \in U$  resulta  $\|x - \frac{(z_n + z_m)}{2}\| \geq d$ , y luego

$$\|z_n - z_m\|_V^2 \leq 2 \|x - z_n\|_V^2 + 2 \|x - z_m\|_V^2 - 4d^2.$$

Así, tomando límite cuando  $m, n \rightarrow +\infty$ , obtenemos  $\|z_n - z_m\|_V \rightarrow 0$ , lo que indica que  $\{z_n\}_{n \in \mathbb{N}}$  es de Cauchy en  $V$ . Como  $U$  es cerrado, se deduce que existe  $z \in U$  tal que  $z_n \rightarrow z$  y por lo tanto

$$d = \lim_{n \rightarrow +\infty} \|x - z_n\|_V = \|x - z\|_V.$$

Para la unicidad, sean  $z_1, z_2 \in U$  tales que  $\|x - z_1\|_V = \|x - z_2\|_V = d$ . Aplicando nuevamente los mismos cálculos anteriores, se obtiene

$$\begin{aligned} 0 &\leq \|z_1 - z_2\|_V^2 = 2 \|x - z_1\|_V^2 + 2 \|x - z_2\|_V^2 - 4 \left\| x - \frac{(z_1 + z_2)}{2} \right\|_V^2 \\ &= 4d^2 - 4 \left\| x - \frac{(z_1 + z_2)}{2} \right\|_V^2 \leq 0, \end{aligned}$$

con lo cual  $\|z_1 - z_2\|_V = 0$ . ■

**DEFINICION 6.5** Dado un subconjunto  $S$  de un espacio de Hilbert  $(V, \langle \cdot, \cdot \rangle_V)$ , se define el ortogonal de  $S$ , y se denota  $S^\perp$ , como

$$S^\perp := \{x \in V : \langle x, v \rangle_V = 0 \quad \forall v \in S\}.$$



El siguiente resultado establece que un espacio de Hilbert  $V$  siempre puede escribirse como suma directa de todo subespacio cerrado con su ortogonal.

**Teorema 6.6** (Teorema de Descomposición Ortogonal). *Sea  $U \neq \phi$  un subespacio cerrado de un Hilbert  $V$ . Entonces para cada  $x \in V$  existe  $v \in U$  y  $w \in U^\perp$  tal que  $x = v + w$ . Esta descomposición es única.*

*Demostración.* Si  $x \in U$ , basta elegir  $v = x$  y  $w = 0$ . Supongamos que  $x \notin U$  y, en virtud del Teorema 6.4, sea  $z \in U$  tal que

$$d = \|x - z\|_V = \min_{v \in U} \|x - v\|_V.$$

Es claro que podemos escribir  $x = v + w$  con  $v := z$  y  $w := x - z$ . Basta probar ahora que  $(x - z) \in U^\perp$ . En efecto, para cada  $u \in U$ ,  $u \neq 0$ , y  $\alpha \in \mathbf{R}$ , obtenemos

$$\begin{aligned} d^2 &\leq \|x - (z - \alpha u)\|_V^2 = \|w + \alpha u\|_V^2 \\ &= \|w\|_V^2 + 2\alpha \langle w, u \rangle_V + \alpha^2 \|u\|_V^2 \\ &= \|u\|_V^2 \left\{ \alpha^2 + 2\alpha \frac{\langle w, u \rangle_V}{\|u\|_V^2} + \frac{\langle w, u \rangle_V^2}{\|u\|_V^4} \right\} + d^2 - \frac{\langle w, u \rangle_V^2}{\|u\|_V^2} \\ &= \|u\|_V^2 \left\{ \alpha + \frac{\langle w, u \rangle_V}{\|u\|_V^2} \right\}^2 + d^2 - \frac{\langle w, u \rangle_V^2}{\|u\|_V^2}. \end{aligned}$$

Luego, tomando  $\alpha := -\frac{\langle w, u \rangle_V}{\|u\|_V^2}$  se sigue que

$$d^2 \leq d^2 - \frac{\langle w, u \rangle_V^2}{\|u\|_V^2},$$

lo que implica que  $\langle w, u \rangle_V \leq 0 \forall u \neq 0$ , y por lo tanto  $\langle w, u \rangle_V = 0 \forall u \in U$ .

Ahora, si  $x = v_1 + w_1$  con  $v_1 \in U$  y  $w_1 \in U^\perp$ , entonces  $v + w = v_1 + w_1$ , de donde  $v - v_1 = w_1 - w \in U \cap U^\perp = \{0\}$ , y en consecuencia  $v - v_1 = w - w_1 = 0$ . Esto completa la demostración. ■

Un simple corolario del teorema anterior queda dado por lo siguiente.

**Lema 6.7.** *Sea  $U \neq \phi$  un subespacio cerrado propio de un Hilbert  $V$ . Entonces, existe  $\bar{y} \in V$ ,  $\bar{y} \neq 0$ , tal que  $\bar{y} \in U^\perp$ .*

Con los resultados anteriores estamos en condiciones de demostrar a continuación el Teorema de Representación de Riesz.

**Teorema 6.8** (Teorema de Representación de Riesz) *Sea  $(V, \langle \cdot, \cdot \rangle_V)$  un espacio de Hilbert. Entonces, para cada  $F \in V'$  existe un único  $y \in V$  tal que*

$$F(v) = \langle y, v \rangle_V \quad \forall v \in V,$$

y además

$$\|F\|_{V'} = \|y\|_V.$$

*Demostración.* Si  $F(v) = 0$  para todo  $v \in V$ , entonces  $F$  es el funcional nulo y en tal caso basta tomar  $y = 0$ . Sea  $U := \{v \in V : F(v) = 0\}$  y supongamos que  $U \neq V$ . Por la linealidad y acotamiento de  $F$  es fácil ver que  $U$  es un subespacio cerrado de  $V$ .

De acuerdo al Lema 6.7, existe  $\bar{y} \in V$ ,  $\bar{y} \neq 0$  tal que  $\bar{y} \in U^\perp$ . Es claro entonces que  $\bar{y} \notin U$  y luego  $F(\bar{y}) \neq 0$ .

Por otra parte, para cada  $v \in V$  tenemos

$$F(F(\bar{y})v - F(v)\bar{y}) = F(\bar{y})F(v) - F(v)F(\bar{y}) = 0,$$

lo cual muestra que  $F(\bar{y})v - F(v)\bar{y} \in U$ . Puesto que  $\bar{y} \in U^\perp$  se deduce que

$$0 = \langle \bar{y}, F(\bar{y})v - F(v)\bar{y} \rangle \quad \forall v \in V,$$

o bien

$$\langle \bar{y}, F(\bar{y})v \rangle_V = \langle \bar{y}, F(v)\bar{y} \rangle_V = F(v) \|\bar{y}\|_V^2,$$

de donde

$$F(v) = \langle y, v \rangle_V \quad \forall v \in V,$$

con  $y := \frac{F(\bar{y})\bar{y}}{\|\bar{y}\|_V^2}$ . Para la unicidad de  $y$ , sean  $y_1, y_2 \in V$  tales que

$$F(v) = \langle y_1, v \rangle_V = \langle y_2, v \rangle_V \quad \forall v \in V.$$

Se sigue que  $y_1 - y_2 \in V^\perp = \{0\}$ , y por lo tanto  $y_1 = y_2$ .

Por otro lado, usando la desigualdad de Cauchy-Schwarz se obtiene

$$|F(v)| = |\langle y, v \rangle_V| \leq \|y\|_V \|v\|_V,$$

y luego

$$\|F\|_{V'} := \sup_{0 \neq v \in V} \frac{|F(v)|}{\|v\|_V} \leq \|y\|_V.$$

En particular, para  $v = y$  se tiene

$$\frac{|F(v)|}{\|v\|_V} = \frac{\|y\|_V^2}{\|y\|_V} = \|y\|_V,$$

y en consecuencia  $\|F\|_{V'} = \|y\|_V$ . ■

Para continuar nuestro análisis, se introduce ahora el concepto de operadores. Dados  $(V, \langle \cdot, \cdot \rangle_V)$  y  $(W, \langle \cdot, \cdot \rangle_W)$  espacios de Hilbert reales, toda aplicación (o transformación)  $T : V \rightarrow W$  se llama **operador**. En particular, todo funcional es un operador.

**DEFINICION 6.9.** Un operador  $T : V \rightarrow W$  se dice *lineal* si  $T(\alpha v + \beta w) = \alpha T(v) + \beta T(w)$ , para todo  $\alpha, \beta \in \mathbf{R}$ ,  $v, w \in V$ .

**DEFINICION 6.10.** Un operador lineal  $T : V \rightarrow W$  se dice *acotado* si existe una constante  $M > 0$  tal que  $\|T(v)\|_W \leq M \|v\|_V \quad \forall v \in V$ .

El espacio de todos los operadores lineales y acotados definidos sobre  $V$  con valores en  $W$  se denota  $\mathcal{L}(V, W)$ . Al ínfimo de todas las constantes  $M > 0$  que satisfacen la definición anterior se llama *norma del operador*  $T$ , y se denota  $\|T\|_{\mathcal{L}(V, W)}$ . Notar que

$$\|T\|_{\mathcal{L}(V, W)} = \sup_{0 \neq v \in V} \frac{\|T(v)\|_W}{\|v\|_V}.$$

El Teorema 6.8 induce la definición del siguiente operador, llamado **aplicación de Riesz**,

$$\begin{aligned} R &: V' \rightarrow V \\ V' &\ni F \rightarrow RF \in V, \end{aligned} \tag{6.1}$$

donde  $RF$  es el único elemento de  $V$  tal que

$$F(v) = \langle RF, v \rangle_V \quad \forall v \in V.$$

Es claro que  $R$  es un isomorfismo isométrico ya que es lineal, biyectivo y además  $\|RF\|_V = \|F\|_{V'}$  para todo  $F \in V'$ . En realidad, estas propiedades de  $R$  constituyen una forma equivalente de enunciar el Teorema de Representación de Riesz.

## 6.2 Lema de Lax-Milgram

El único propósito de esta sección es probar el Teorema 3.2, el cual se enuncia aquí nuevamente previo a su demostración.

**Lema de Lax-Milgram.** Sea  $(V, \langle \cdot, \cdot \rangle_V)$  un Hilbert y sea  $A : V \times V \rightarrow \mathbf{R}$  una forma bilineal acotada y  $V$ -elíptica. Entonces, para cada  $F \in V'$ , existe un único  $u \in V$  tal que  $A(u, v) = F(v) \quad \forall v \in V$ . Además,

$$\|u\|_V \leq \frac{1}{\alpha} \|F\|_{V'},$$

donde  $\alpha > 0$  es la constante de  $V$ -elipticidad de  $A$ .

*Demostración.* Puesto que  $A$  es acotada podemos introducir el operador lineal  $\mathbf{A} : V \rightarrow V'$ , tal que para cada  $v \in V$ ,  $\mathbf{A}v \in V'$  se define como

$$\mathbf{A}v(w) := A(v, w) \quad \forall w \in V.$$

Claramente  $\mathbf{A}$  es acotado ya que

$$\|\mathbf{A}v\|_{V'} := \sup_{0 \neq w \in V} \frac{|\mathbf{A}v(w)|}{\|w\|_V} = \sup_{0 \neq w \in V} \frac{|A(v, w)|}{\|w\|_V} \leq M \|v\|_V,$$

lo cual indica que  $\|\mathbf{A}\|_{\mathcal{L}(V, V')} \leq M$ . Luego, si  $R : V' \rightarrow V$  es la aplicación de Riesz definida en (6.1), el problema que nos interesa equivale a: *Hallar  $u \in V$  tal que*

$$\mathbf{A}u(v) = \langle RF, v \rangle_{V'} \quad \forall v \in V'. \quad (6.2)$$

Es claro que (6.2) puede reformularse como: *Hallar  $u \in V$  tal que*

$$\mathcal{A}u = RF,$$

donde  $\mathcal{A} := RA : V \rightarrow V$  denota el operador lineal que se ilustra en el siguiente diagrama

$$\begin{array}{ccc} & \mathbf{A} & \\ & \longrightarrow & V' \\ V & & \\ & \mathcal{A} \searrow & \downarrow R \\ & & V \end{array}$$

De este modo, la presente demostración se reduce a probar que  $\mathcal{A}$  es un operador biyectivo con inversa acotada. Para este efecto, notamos primero que  $\mathcal{A}$  es acotado y que  $\|\mathcal{A}\|_{\mathcal{L}(V, V)} \leq M$ . También, por la  $V$ -elipticidad de  $A$ , se tiene

$$\alpha \|v\|_V^2 \leq A(v, v) = \mathbf{A}v(v) \leq \|\mathbf{A}v\|_{V'} \|v\|_V,$$

esto es

$$\alpha \|v\|_V \leq \|\mathbf{A}v\|_{V'} = \|\mathbf{R}\mathbf{A}v\|_V = \|\mathcal{A}v\|_V,$$

o bien

$$\alpha \|v\|_V \leq \|\mathcal{A}v\|_V \quad \forall v \in V. \quad (6.3)$$

La inyectividad de  $\mathcal{A}$  es inmediata a partir de (6.3). Para mostrar que  $\mathcal{A}$  es sobreyectivo procedemos de la siguiente manera. Primero probamos que  $\mathcal{A}(V)$  es

un subespacio cerrado de  $V$ , luego que  $\mathcal{A}(V)^\perp = \{0\}$  y por último aplicamos el Teorema de Descomposición Ortogonal (cf. Teorema 6.6).

Sea  $\{v_n\}_{n \in \mathbb{N}} \subseteq V$  tal que  $\mathcal{A}(v_n) \rightarrow w \in V$  cuando  $n \rightarrow +\infty$ . Usando (6.3) se obtiene que

$$\alpha \|v_n - v_m\|_V \leq \|\mathcal{A}(v_n) - \mathcal{A}(v_m)\|_V,$$

lo cual indica que  $\{v_n\}_{n \in \mathbb{N}}$  es de Cauchy en  $V$  y por lo tanto convergente a un elemento  $v \in V$ . En virtud de la continuidad de  $\mathcal{A}$  se sigue que  $\mathcal{A}(v_n) \rightarrow \mathcal{A}(v)$  cuando  $n \rightarrow +\infty$ , y por la unicidad del límite se concluye que  $w = \mathcal{A}(v) \in \mathcal{A}(V)$ . Esto prueba que  $\mathcal{A}(V)$  es cerrado.

Sea  $v \in \mathcal{A}(V)^\perp$ . Entonces  $\langle w, v \rangle_V = 0$  para todo  $w \in \mathcal{A}(V)$ . En particular, para  $w = \mathcal{A}(v) = RA(v)$  se obtiene

$$0 = \langle \mathcal{A}(v), v \rangle_V = \langle RA(v), v \rangle_V = \mathbf{A}v(v) = A(v, v) \geq \alpha \|v\|_V^2,$$

de donde  $v = 0$ . Esto prueba que  $\mathcal{A}(V)^\perp = \{0\}$ . En consecuencia, la sobreyectividad de  $\mathcal{A}$  resulta de una aplicación directa del Teorema 6.6.

Por otra parte, de (6.3) se deduce que

$$\alpha \|u\|_V \leq \|\mathcal{A}u\|_V = \|RF\|_V = \|F\|_{V'},$$

y por lo tanto  $\|u\|_V \leq \frac{1}{\alpha} \|F\|_{V'}$ . Esto completa la demostración. ■

## 7 La ecuación de Poisson

En esta sección se considera el problema de Poisson con condiciones de contorno de Dirichlet. Se muestran algunas características de la formulación variacional, y se ilustra la aplicación del método de elementos finitos a través de un ejemplo bidimensional.

En lo que sigue,  $\Omega$  denota un abierto acotado de  $\mathbf{R}^N$  con frontera  $\Gamma := \partial\Omega$  de clase  $C^1$ . La definición precisa del grado de suavidad de la frontera, incluyendo ciertamente  $C^1$  y otras clases de funciones, puede verse en [ 1], [ 3] o [ 7].

### 7.1 Identidades de Green

Con el objeto de derivar posteriormente las formulaciones débiles respectivas, se deducen a continuación las *identidades de Green*. Para ello, recordamos primero el Teorema de la Divergencia de Gauss, el cual dice que para un campo vectorial  $G \in [C^1(\Omega) \cap C(\bar{\Omega})]^N$  se tiene

$$\int_{\Omega} \operatorname{div} G \, dx = \int_{\Gamma} G \cdot \nu \, ds,$$

donde  $\operatorname{div} G := \sum_{i=1}^N \frac{\partial G_i}{\partial x_i}$  y  $\nu := (\nu_1, \dots, \nu_N)^T$  es el vector normal unitario exterior a  $\Omega$ . En particular, si  $G := (0, \dots, 0, \overset{(i)}{u}v, 0, \dots, 0)$  con  $u, v \in C^1(\Omega) \cap C(\bar{\Omega})$ , entonces se obtiene

$$\int_{\Omega} v \frac{\partial}{\partial x_i} (uv) \, dx = \int_{\Gamma} uv \nu_i \, ds,$$

es decir

$$\int_{\Omega} v \frac{\partial u}{\partial x_i} \, dx = - \int_{\Omega} u \frac{\partial v}{\partial x_i} \, dx + \int_{\Gamma} uv \nu_i \, ds \quad \forall i = 1, \dots, N.$$

Ahora, reemplazando  $u = \frac{\partial w}{\partial x_i}$  en la ecuación anterior, resulta

$$\int_{\Omega} v \frac{\partial^2 w}{\partial x_i^2} \, dx = - \int_{\Omega} \frac{\partial w}{\partial x_i} \frac{\partial v}{\partial x_i} \, dx + \int_{\Gamma} \frac{\partial w}{\partial x_i} v \nu_i \, ds \quad \forall i = 1, \dots, N,$$

y luego, sumando sobre todos los índices  $i$ , se concluye la *primera identidad de Green*:

$$\int_{\Omega} v \Delta w \, dx = - \int_{\Omega} \nabla w \cdot \nabla v \, dx + \int_{\Gamma} \frac{\partial w}{\partial \nu} v \, ds, \quad (7.1)$$

la cual es válida para todo  $v \in C^1(\Omega) \cap C(\bar{\Omega})$  y para todo  $w \in C^2(\Omega) \cap C^1(\bar{\Omega})$ . Aquí,  $\frac{\partial w}{\partial \nu} := \sum_{i=1}^N \frac{\partial w}{\partial x_i} \nu_i$  es la derivada normal sobre  $\Gamma$ .

Intercambiando los roles de  $v$  y  $w$  en (7.1) y luego restando la ecuación resultante a (7.1) se obtiene la *segunda identidad de Green*:

$$\int_{\Omega} [v \Delta w - w \Delta v] \, dx = \int_{\Gamma} \left[ v \frac{\partial w}{\partial \nu} - w \frac{\partial v}{\partial \nu} \right] \, ds, \quad (7.2)$$

para todo  $v, w \in C^2(\Omega) \cap C^1(\bar{\Omega})$ .

## 7.2 Un problema de Dirichlet

Dada una función  $f \in C(\Omega)$ , nos interesa el siguiente problema de valores de contorno: *Hallar  $u \in C^2(\Omega) \cap C(\bar{\Omega})$  tal que*

$$\begin{aligned} -\Delta u &= f \quad \text{en } \Omega, \\ u &= 0 \quad \text{en } \Gamma. \end{aligned} \quad (7.3)$$

Ahora, consideramos nuevamente el espacio  $C_0^\infty(\Omega)$ , cuya definición en el caso  $N$ -dimensional es idéntica a la dada en la Sección 2 para un abierto  $\Omega$  de  $\mathbb{R}^N$ .

Multiplicando la ecuación diferencial parcial de (7.3) por una función  $\varphi \in C_0^\infty(\Omega)$ , aplicando la *primera identidad de Green* (7.1), y usando el hecho que  $\varphi|_{\Gamma} = 0$ , resulta

$$\int_{\Omega} \nabla u \cdot \nabla \varphi \, dx = \int_{\Omega} f \varphi \, dx \quad \forall \varphi \in C_0^\infty(\Omega). \quad (7.4)$$

DEFINICION 7.1. Dada  $v \in L^2(\Omega)$ , se dice que  $\frac{\partial v}{\partial x_i} \in L^2(\Omega)$  en el sentido distribucional, si existe  $z_i \in L^2(\Omega)$  tal que

$$-\int_{\Omega} v \frac{\partial \varphi}{\partial x_i} \, dx = \int_{\Omega} z_i \varphi \, dx \quad \forall \varphi \in C_0^\infty(\Omega).$$

En este caso, se escribe  $\frac{\partial v}{\partial x_i} = z_i$ .

De acuerdo a la definición anterior, podemos introducir el *espacio de Sobolev* análogo al dado en (2.2),

$$H^1(\Omega) := \left\{ v \in L^2(\Omega) : \frac{\partial v}{\partial x_i} \in L^2(\Omega) \quad \forall i = 1, \dots, N \right\}, \quad (7.5)$$

el cual también es un espacio de Hilbert provisto del siguiente producto escalar

$$\langle v, w \rangle_{H^1(\Omega)} := \int_{\Omega} [\nabla v \cdot \nabla w + v w] \, dx \quad \forall v, w \in H^1(\Omega).$$

Entonces, la norma de  $H^1(\Omega)$  está dada por

$$\|\cdot\|_{H^1(\Omega)} : H^1(\Omega) \rightarrow \mathbf{R}$$

$$\|v\|_{H^1(\Omega)} := \left\{ |v|_{H^1(\Omega)}^2 + \|v\|_{L^2(\Omega)}^2 \right\}^{1/2} \quad \forall v \in H^1(\Omega),$$

donde  $|\cdot|_{H^1(\Omega)}$  es la semi-norma asociada definida como

$$|v|_{H^1(\Omega)} := \left\{ \int_{\Omega} \|\nabla v\|^2 \, dx \right\}^{1/2} = \left\{ \sum_{i=1}^N \left\| \frac{\partial v}{\partial x_i} \right\|_{L^2(\Omega)}^2 \right\}^{1/2} \quad \forall v \in H^1(\Omega).$$

Al igual que en la Sección 2 se introduce ahora la adherencia de  $C_0^\infty(\Omega)$  en  $H^1(\Omega)$  y se denota por  $H_0^1(\Omega)$ . Puede probarse (ver [7, Capítulo 1]) que  $H_0^1(\Omega)$  coincide con el subespacio de elementos de  $H^1(\Omega)$  que se anulan en la frontera en el sentido de trazas, esto es

$$H_0^1(\Omega) := \left\{ v \in H^1(\Omega) : v = 0 \quad \text{en} \quad \partial\Omega \right\}.$$

En virtud de la identidad (7.4) y de la densidad de  $C_0^\infty(\Omega)$  en  $H_0^1(\Omega)$ , la formulación débil del problema de valores de contorno (7.3) se reduce a: *Hallar  $u \in H_0^1(\Omega)$  tal que*

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in H_0^1(\Omega). \quad (7.6)$$

Sea  $V := H_0^1(\Omega)$  y definamos la forma bilineal

$$A : V \times V \rightarrow \mathbf{R}$$

$$A(w, v) := \int_{\Omega} \nabla w \cdot \nabla v \, dx \quad \forall w, v \in V, \quad (7.7)$$

y el funcional lineal

$$F : V \rightarrow \mathbf{R}$$

$$F(v) := \int_{\Omega} f v \, dx \quad \forall v \in V. \quad (7.8)$$

Con estas notaciones, (7.6) se reescribe como sigue: *Hallar  $u \in V$  tal que*

$$A(u, v) = F(v) \quad \forall v \in V. \quad (7.9)$$

Nuestro siguiente propósito es demostrar que (7.9) verifica las hipótesis del Lema de Lax-Milgram (cf. Sección 6.2). En efecto, utilizando la desigualdad de Cauchy-Schwarz en  $L^2(\Omega)$  y luego en  $\mathbf{R}^N$ , se obtiene

$$\begin{aligned} |A(w, v)| &\leq \sum_{i=1}^N \left\{ \left\| \frac{\partial w}{\partial x_i} \right\|_{L^2(\Omega)} \left\| \frac{\partial v}{\partial x_i} \right\|_{L^2(\Omega)} \right\} \\ &\leq \left\{ \sum_{i=1}^N \left\| \frac{\partial w}{\partial x_i} \right\|_{L^2(\Omega)}^2 \right\}^{1/2} \left\{ \sum_{i=1}^N \left\| \frac{\partial v}{\partial x_i} \right\|_{L^2(\Omega)}^2 \right\}^{1/2} \\ &= \|w\|_{H^1(\Omega)} \|v\|_{H^1(\Omega)}, \end{aligned}$$

lo cual prueba que  $A$  es acotada con constante  $M = 1$ . También, usando la desigualdad de Cauchy-Schwarz en  $L^2(\Omega)$ , se muestra fácilmente que  $F$  es acotado.



Para ver que  $A$  es  $V$ -elíptica, probamos a continuación que la norma y la seminorma de  $H^1(\Omega)$  son equivalentes en  $H_0^1(\Omega)$ , resultado que se conoce con el nombre de Desigualdad de Poincaré (ver [7, Teorema 1.2-5]).

**Lema 7.2.** *Sea  $\Omega$  un abierto de  $\mathbf{R}^N$ , el cual es acotado en alguna dirección coordenada. Entonces, existe una constante  $\alpha > 0$ , que depende sólo de  $\Omega$ , tal que*

$$\alpha \|v\|_{H^1(\Omega)}^2 \leq |v|_{H^1(\Omega)}^2 \quad \forall v \in H_0^1(\Omega). \quad (7.10)$$

*Demostración.* Sin pérdida de generalidad, supongamos que  $\Omega$  es acotado en la dirección  $x_1$ . Análogamente a la demostración del Lema 3.1, basta mostrar aquí que (7.10) se satisface en el subespacio denso de  $H_0^1(\Omega)$  dado por  $C_0^\infty(\Omega)$ . En efecto, dada  $\varphi \in C_0^\infty(\Omega)$  y  $\mathbf{x} \in \Omega$ , notemos primero que

$$\varphi(\mathbf{x}) := \varphi(x_1, x_2, \dots, x_N) = \int_{-\infty}^{x_1} \frac{\partial \varphi}{\partial t}(t, x_2, \dots, x_N) dt.$$

Puesto que  $\text{sop } \varphi$  es un compacto contenido en  $\Omega$ , existen  $a_i, b_i \in \mathbf{R}$ ,  $i \in \{1, \dots, N\}$ , tales que  $\text{sop } \varphi \subset [a_1, b_1] \times [a_2, b_2] \times \dots \times [a_N, b_N]$ . Además, en virtud del supuesto inicial sobre el acotamiento uni-direccional de  $\Omega$ ,  $a_1$  y  $b_1$  pueden elegirse independientes de  $\varphi$ . Luego, usando la desigualdad de Cauchy-Schwarz, se obtiene

$$\left| \varphi(x_1, \hat{\mathbf{x}}) \right|^2 = \left| \int_{a_1}^{x_1} \frac{\partial \varphi}{\partial t}(t, \hat{\mathbf{x}}) dt \right|^2 \leq (x_1 - a_1) \int_{a_1}^{b_1} \left| \frac{\partial \varphi}{\partial t}(t, \hat{\mathbf{x}}) \right|^2 dt,$$

donde  $\hat{\mathbf{x}} := (x_2, \dots, x_N) \in \mathbf{R}^{N-1}$ . Ahora, integrando sobre  $\Omega$  a ambos lados de la estimación anterior y utilizando el Teorema de Fubini para intercambiar el orden de integración, resulta

$$\begin{aligned} \int_{\Omega} |\varphi(\mathbf{x})|^2 d\mathbf{x} &\leq \int_{\Omega} (x_1 - a_1) \left\{ \int_{a_1}^{b_1} \left| \frac{\partial \varphi}{\partial t}(t, \hat{\mathbf{x}}) \right|^2 dt \right\} dx_1 d\hat{\mathbf{x}} \\ &= \int_{a_1}^{b_1} (x_1 - a_1) dx_1 \int_{a_2}^{b_2} \dots \int_{a_{N-1}}^{b_{N-1}} \int_{a_N}^{b_N} \left\{ \int_{a_1}^{b_1} \left| \frac{\partial \varphi}{\partial t}(t, \hat{\mathbf{x}}) \right|^2 dt \right\} d\hat{\mathbf{x}} \\ &= \frac{1}{2}(b_1 - a_1)^2 \int_{\Omega} \left| \frac{\partial \varphi}{\partial t}(t, \hat{\mathbf{x}}) \right|^2 dt d\hat{\mathbf{x}} \leq \frac{1}{2}(b_1 - a_1)^2 \int_{\Omega} \|\nabla \varphi\|^2 d\mathbf{x}, \end{aligned}$$

de donde

$$\|\varphi\|_{L^2(\Omega)} \leq \frac{1}{\sqrt{2}}(b_1 - a_1) |\varphi|_{H^1(\Omega)} \quad \forall \varphi \in C_0^\infty(\Omega).$$

Por lo tanto, se concluye que

$$\|\varphi\|_{H^1(\Omega)}^2 = |\varphi|_{H^1(\Omega)}^2 + \|\varphi\|_{L^2(\Omega)}^2 \leq \left[1 + \frac{1}{2}(b_1 - a_1)^2\right] |\varphi|_{H^1(\Omega)}^2 \quad \forall \varphi \in C_0^\infty(\Omega),$$

lo cual termina la demostración. ■

El análisis anterior permite la aplicación del Lema de Lax-Milgram, y en consecuencia la formulación débil (7.6) (o (7.9)) posee una única solución  $u \in H_0^1(\Omega)$ .

### 7.3 Un método de elementos finitos

En esta sección se considera el caso bi-dimensional y se presenta un método de elementos finitos para la formulación débil (7.9) asociada al problema de valores de contorno (7.3). El procedimiento que se describe a continuación constituye la extensión natural a  $\mathbf{R}^2$  del esquema dado en el EJEMPLO 4.1 (cf. Sección 4) para el problema uni-dimensional (3.3).

En lo que sigue, se asume que  $\Omega$  es un dominio poligonal de  $\mathbf{R}^2$ , lo cual significa que la frontera  $\partial\Omega$  está dada por una unión finita de trazos de rectas. Sea  $\mathcal{T} := \{K_1, K_2, \dots, K_p\}$  una familia de  $p$  triángulos contenidos en  $\bar{\Omega}$ . Se dice que  $\mathcal{T}$  constituye una *triangulación* del dominio  $\Omega$  si

- 1)  $\bar{\Omega} = \cup_{k=1}^p K_k$ .
- 2) Para  $k, r \in \{1, \dots, p\}$ ,  $k \neq r$ , se tiene sólo una de las tres posibilidades siguientes:
  - a)  $K_k \cap K_r$  es vacío.
  - b)  $K_k \cap K_r$  es un vértice de ambos triángulos.
  - c)  $K_k \cap K_r$  es un lado común de  $K_k$  y  $K_r$ .

Un ejemplo de triangulación se observa en la Figura 7.1 más abajo. Denotemos por  $\{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_m\}$  los vértices (o nodos) de la triangulación  $\mathcal{T}$ . Supongamos que el conjunto de nodos  $\{\bar{x}_1, \dots, \bar{x}_n\}$  está contenido en el interior de  $\Omega$  y que  $\{\bar{x}_{n+1}, \bar{x}_{n+2}, \dots, \bar{x}_m\} \subseteq \partial\Omega$ . Entonces definimos el subespacio  $V_n$  de  $H_0^1(\Omega)$  dado por

$$V_n := \left\{ v \in C(\bar{\Omega}) : v|_{K_k} \in P_1(K_k) \quad \forall k = \overline{1, p}, v(\bar{x}_i) = 0 \quad \forall i = \overline{n+1, m} \right\}.$$

Además, para cada  $j \in \{1, \dots, n\}$  se introduce la función  $\varphi_j \in V_n$  tal que  $\varphi_j(\bar{x}_j) = 1$  y  $\varphi_j(\bar{x}_i) = 0$  para todo  $i \neq j$ . Se sigue que  $\text{sop } \varphi_j$  es la unión de todos los triángulos de  $\mathcal{T}$  a los cuales pertenece el nodo  $\bar{x}_j$ .

De acuerdo a la definición de  $V_n$  se deduce fácilmente que  $\{\varphi_1, \dots, \varphi_n\}$  es una base de  $V_n$  y que

$$v = \sum_{j=1}^n v(\bar{x}_j) \varphi_j \quad \forall v \in V_n.$$

De este modo, si la solución de Galerkin  $u_n \in V_n$  se escribe  $u_n = \sum_{j=1}^n \mu_j \varphi_j$ , con  $\mu_j = u_n(\bar{x}_j)$ , entonces el esquema asociado se reduce a: Hallar  $\mu := (\mu_1, \dots, \mu_n) \in \mathbb{R}^n$  tal que

$$\mathbf{A} \mu = \mathbf{F}, \quad (7.11)$$

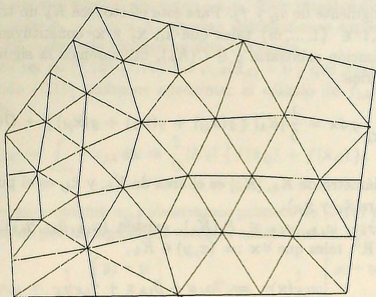


Figura 7.1: Una triangulación típica.

donde la matriz de rigidez  $\mathbf{A} := (a_{ij})_{n \times n}$  y el vector de carga  $\mathbf{F} := (f_j)_{n \times 1}$  están dados por

$$a_{ij} = \int_{\Omega} \nabla \varphi_i \cdot \nabla \varphi_j \, d\mathbf{x} \quad \text{y} \quad f_j = \int_{\Omega} f \varphi_j \, d\mathbf{x}. \quad (7.12)$$

Ahora, puesto que  $\bar{\Omega} = \cup_{k=1}^p K_k$ , podemos escribir

$$a_{ij} = \sum_{k=1}^p a_{ij,k} \quad \text{y} \quad f_j = \sum_{k=1}^p f_{j,k}, \quad (7.13)$$

donde

$$a_{ij,k} := \int_{K_k} \nabla \varphi_{i,k} \cdot \nabla \varphi_{j,k} \, d\mathbf{x} \quad , \quad f_{j,k} := \int_{K_k} f \varphi_{j,k} \, d\mathbf{x}, \quad (7.14)$$

y  $\varphi_{i,k}$  es la restricción de la función base  $\varphi_i$  al triángulo  $K_k$  de  $\mathcal{T}$ . Note que cada uno de los sumandos en la ecuación (7.13) contiene un gran número de términos

ulos. En efecto, puesto que  $\varphi_{i,k} \neq 0$  sólo si  $\bar{x}_i \in K_k$ , se observa que  $a_{ij,k} = 0$  para todos aquellos triángulos  $K_k$  a los cuales no pertenecen ambos nodos  $\bar{x}_i$  y  $\bar{x}_j$ . Esto implica que, para  $i \neq j$ ,  $a_{ij,k}$  es distinto de cero sólo para dos triángulos  $K_k$ . Similarmente,  $a_{jj,k}$  y  $f_{j,k}$  son eventualmente distintos de cero sólo para los triángulos  $K_k$  que conforman  $\text{sop } \varphi_j$ , vale decir, para aquellos que tienen a  $\bar{x}_j$  como uno de sus vértices.

El objetivo siguiente es obtener fórmulas más explícitas para el cálculo de  $a_{ij,k}$  y  $f_{j,k}$ , y por consiguiente de  $a_{ij}$  y  $f_j$ . Para este efecto, sea  $K_k$  un triángulo genérico de  $\mathcal{T}$  y sean  $i, j, l \in \{1, \dots, m\}$  tales que  $\bar{x}_i, \bar{x}_j$  y  $\bar{x}_l$  constituyen los vértices de  $K_k$ . Dada una función arbitraria  $g \in C(K_k)$ , utilizaremos la siguiente fórmula de integración numérica:

$$\int_{K_k} g \, dx = \frac{1}{3} |K_k| \{g(\bar{x}_{ij}) + g(\bar{x}_{il}) + g(\bar{x}_{jl})\} + O(h_k^2), \quad (7.15)$$

donde  $h_k$  es el diámetro de  $K_k$ ,  $|K_k|$  es el área de  $K_k$ , y  $\bar{x}_{ij}$  es el punto medio entre  $\bar{x}_i$  y  $\bar{x}_j$  (idem para  $\bar{x}_{il}$  y  $\bar{x}_{jl}$ ).

Puesto que  $\varphi_{i,k}, \varphi_{j,k}, \varphi_{l,k} \in P_1(K_k)$ , existen  $(\alpha_{i,k}, \beta_{i,k}, \gamma_{i,k}), (\alpha_{j,k}, \beta_{j,k}, \gamma_{j,k}), (\alpha_{l,k}, \beta_{l,k}, \gamma_{l,k}) \in \mathbb{R}^3$ , tales que  $\forall x := (x, y) \in K_k$ ,

$$\begin{aligned} \varphi_{i,k}(x) &= \alpha_{i,k} + \beta_{i,k}x + \gamma_{i,k}y, \\ \varphi_{j,k}(x) &= \alpha_{j,k} + \beta_{j,k}x + \gamma_{j,k}y, \\ \varphi_{l,k}(x) &= \alpha_{l,k} + \beta_{l,k}x + \gamma_{l,k}y. \end{aligned} \quad (7.16)$$

Así, dado que  $\varphi_{r,k}(\bar{x}_s) = \begin{cases} 1 & \text{si } r = s \\ 0 & \text{si } r \neq s \end{cases}$ , para todo  $r, s \in \{i, j, l\}$ , resulta

$$\begin{bmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_l & y_l \end{bmatrix} \begin{bmatrix} \alpha_{i,k} & \alpha_{j,k} & \alpha_{l,k} \\ \beta_{i,k} & \beta_{j,k} & \beta_{l,k} \\ \gamma_{i,k} & \gamma_{j,k} & \gamma_{l,k} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad (7.17)$$

donde  $\bar{x}_i = (x_i, y_i)$ ,  $\bar{x}_j = (x_j, y_j)$  y  $\bar{x}_l = (x_l, y_l)$ . Se sigue que

$$\begin{bmatrix} \alpha_{i,k} & \alpha_{j,k} & \alpha_{l,k} \\ \beta_{i,k} & \beta_{j,k} & \beta_{l,k} \\ \gamma_{i,k} & \gamma_{j,k} & \gamma_{l,k} \end{bmatrix} = \begin{bmatrix} 1 & x_i & y_i \\ 1 & x_j & y_j \\ 1 & x_l & y_l \end{bmatrix}^{-1}$$

$$= \frac{1}{D_{ijl}} \begin{bmatrix} x_j y_l - x_l y_j & x_l y_i - x_i y_l & x_i y_j - x_j y_i \\ y_j - y_l & y_l - y_i & y_i - y_j \\ x_l - x_j & x_i - x_l & x_j - x_i \end{bmatrix}, \quad (7.18)$$

donde  $D_{ijl} := x_j y_l - x_l y_j + x_l y_i - x_i y_l + x_i y_j - x_j y_i$ . Esto muestra que las restricciones de las funciones base  $\varphi_i, \varphi_j, \varphi_l$  a  $K_k$  quedan determinadas explícitamente por las coordenadas de los nodos que conforman este triángulo.

Ahora, es fácil ver de (7.16)-(7.17) que  $\varphi_{j,k}(\bar{x}_i) = 0$  y que  $\varphi_{j,k}(\bar{x}_{ij}) = \varphi_{j,k}(\bar{x}_{jl}) = \frac{1}{2}$ . Luego, usando (7.15) podemos aproximar el cálculo de  $f_{j,k}$  por medio de la siguiente fórmula:

$$f_{j,k} := \int_{K_k} f \varphi_{j,k} dx \approx \frac{1}{6} |K_k| \{ f(\bar{x}_{ij}) + f(\bar{x}_{jl}) \}. \quad (7.19)$$

Por otra parte, para obtener  $a_{ij,k}$  observamos primero de (7.16) y (7.18) que  $\nabla \varphi_{i,k}$  y  $\nabla \varphi_{j,k}$  están dados por los siguientes vectores constantes en  $\mathbf{R}^2$ ,

$$\nabla \varphi_{i,k} = \frac{1}{D_{ijl}} \begin{bmatrix} y_j - y_l \\ x_l - x_j \end{bmatrix} \quad \text{y} \quad \nabla \varphi_{j,k} = \frac{1}{D_{ijl}} \begin{bmatrix} y_l - y_i \\ x_i - x_l \end{bmatrix}.$$

Luego, en este caso no se requiere integración numérica y el cálculo de  $a_{ij,k}$  se reduce a

$$a_{ij,k} := \int_{K_k} \nabla \varphi_{i,k} \cdot \nabla \varphi_{j,k} dx = \frac{|K_k|}{D_{ijl}^2} (\bar{x}_l - \bar{x}_j) \cdot (\bar{x}_i - \bar{x}_l). \quad (7.20)$$

De manera análoga se calculan  $a_{il,k}$  y  $a_{jl,k}$ , lo cual da

$$a_{il,k} := \int_{K_k} \nabla \varphi_{i,k} \cdot \nabla \varphi_{l,k} dx = \frac{|K_k|}{D_{ijl}^2} (\bar{x}_l - \bar{x}_j) \cdot (\bar{x}_j - \bar{x}_i), \quad (7.21)$$

y

$$a_{jl,k} := \int_{K_k} \nabla \varphi_{j,k} \cdot \nabla \varphi_{l,k} dx = \frac{|K_k|}{D_{ijl}^2} (\bar{x}_i - \bar{x}_l) \cdot (\bar{x}_j - \bar{x}_i). \quad (7.22)$$

También, es fácil ver que

$$a_{ii,k} = \frac{|K_k|}{D_{ijl}^2} \|\bar{x}_l - \bar{x}_j\|^2, \quad (7.23)$$

$$a_{jj,k} = \frac{|K_k|}{D_{ij}^2} \|\bar{x}_i - \bar{x}_l\|^2, \quad (7.24)$$

$$a_{ll,k} = \frac{|K_k|}{D_{ij}^2} \|\bar{x}_i - \bar{x}_j\|^2. \quad (7.25)$$

Esto induce la definición de la *matriz de rigidez local*  $\mathbf{A}_k$ , la cual es simétrica, y cuyos coeficientes están determinados por (7.20)-(7.25), esto es

$$\mathbf{A}_k := \begin{bmatrix} a_{ii,k} & a_{ij,k} & a_{il,k} \\ a_{ij,k} & a_{jj,k} & a_{jl,k} \\ a_{il,k} & a_{jl,k} & a_{ll,k} \end{bmatrix}_{3 \times 3}, \quad (7.26)$$

También, se define el *vector de carga local*  $\mathbf{F}_k$  como

$$\mathbf{F}_k := \begin{bmatrix} f_{i,k} \\ f_{j,k} \\ f_{l,k} \end{bmatrix}_{3 \times 1}. \quad (7.27)$$

Naturalmente, si alguno de los nodos  $\bar{x}_i$ ,  $\bar{x}_j$ ,  $\bar{x}_l$  pertenece a  $\partial\Omega$ , entonces se debe eliminar la fila y la columna respectiva de  $\mathbf{A}_k$ , y la fila correspondiente de  $\mathbf{F}_k$ .

El análisis previo sugiere, en virtud de las expresiones dadas en (7.13), que el cálculo de la matriz de rigidez global  $\mathbf{A}$  y del vector de carga global  $\mathbf{F}$  puede hacerse triángulo por triángulo a través de las respectivas fórmulas locales dadas por (7.26) y (7.27). Este procedimiento recibe el nombre de *ensamblamiento*.

En relación a lo anterior, es importante observar, sin embargo, que si el sistema lineal (7.11) se resuelve utilizando el **Método del Gradiente Conjugado**, entonces **no es necesario obtener explícitamente la matriz  $\mathbf{A}$** . En efecto, puesto que este método se basa principalmente en multiplicaciones de  $\mathbf{A}$  por un vector dado  $\mathbf{z} \in \mathbf{R}^n$ , todo lo que se requiere es un algoritmo que permita calcular estos productos de manera eficiente. A continuación mostramos que esto también puede hacerse localmente.

Sea  $\mathbf{w} = \mathbf{A} \mathbf{z}$ , con  $\mathbf{w} := (w_1, w_2, \dots, w_n)^T$  y  $\mathbf{z} := (z_1, z_2, \dots, z_n)^T$ . Entonces para todo  $r \in \{1, \dots, n\}$  se tiene

$$w_r = \sum_{s=1}^n a_{rs} z_s = \sum_{s=1}^n \left\{ \sum_{k=1}^p a_{rs,k} \right\} z_s = \sum_{k=1}^p \left\{ \sum_{s=1}^n a_{rs,k} z_s \right\},$$

y puesto que  $a_{rs,k} = 0$  si  $\bar{x}_s \notin K_k$ , resulta

$$w_r = \sum_{k=1}^p \left\{ \sum_{s \in \{i,j,l\}} a_{rs,k} z_s \right\},$$

o bien

$$w_r = \sum_{k=1}^p w_{r,k} \quad , \quad \text{donde} \quad w_{r,k} = \sum_{s \in \{i,j,l\}} a_{rs,k} z_s. \quad (7.28)$$

Es claro que si  $r \notin \{i, j, l\}$ , equivalentemente  $\bar{x}_r \notin K_k$ , entonces  $w_{r,k} = 0$ . Por lo tanto, para identificar aquellos  $w_{r,k}$  no nulos y calcular así las componentes de  $w$  mediante (7.28), basta recorrer los triángulos  $K_k$  y efectuar la multiplicación de la matriz de rigidez local  $A_k$  por el vector  $(z_i, z_j, z_l)^T \in \mathbf{R}^3$ . Es decir, para cada triángulo  $K_k$ , cuyos vértices son  $\bar{x}_i, \bar{x}_j$  y  $\bar{x}_l$ , se realiza el siguiente producto

$$\begin{bmatrix} w_{i,k} \\ w_{j,k} \\ w_{l,k} \end{bmatrix} := \begin{bmatrix} a_{ii,k} & a_{ij,k} & a_{il,k} \\ a_{ij,k} & a_{jj,k} & a_{jl,k} \\ a_{il,k} & a_{jl,k} & a_{ll,k} \end{bmatrix} \begin{bmatrix} z_i \\ z_j \\ z_l \end{bmatrix}, \quad (7.29)$$

y luego se suman  $w_{i,k}$ ,  $w_{j,k}$  y  $w_{l,k}$  a las expresiones de  $w_i$ ,  $w_j$  y  $w_l$ , respectivamente.

Esto completa el análisis matricial para el cálculo de  $u_n$  a través del sistema lineal (7.11).

Para finalizar esta sección, recordemos que el error de la solución de elementos finitos  $u_n$  queda acotado por la estimación de Cea (cf. (5.3) en Lema 5.1), la cual, en este caso, se reduce a:  $\|u - u_n\|_V \leq \frac{1}{\alpha} \inf_{v_n \in V_n} \|u - v_n\|_V$ . Al igual que para el caso unidimensional (cf. Sección 5.3), aquí también puede demostrarse que  $\inf_{v_n \in V_n} \|u - v_n\|_V \leq Ch$ , donde  $C$  es una constante positiva que depende de  $u$ , y  $h := \max \{h_k \mid k = 1, \dots, p\}$  (ver [3], [7]).

## 7.4 El caso no-homogéneo

Dadas  $f \in C(\Omega)$  y  $g \in C(\partial\Omega)$ , nos interesa ahora el siguiente problema de valores de contorno: *Hallar  $u \in C^2(\Omega) \cap C(\bar{\Omega})$  tal que*

$$\begin{aligned} -\Delta u &= f \quad \text{en } \Omega, \\ u &= g \quad \text{en } \Gamma. \end{aligned} \quad (7.30)$$

Si siguiendo el mismo análisis de la Sección 7.2, puede probarse que la formulación débil de (7.30) está dada por: *Hallar  $u \in H^1(\Omega)$  tal que  $u = g$  en  $\Gamma$  y*

$$\int_{\Omega} \nabla u \cdot \nabla v \, dx = \int_{\Omega} f v \, dx \quad \forall v \in H_0^1(\Omega). \quad (7.31)$$

Ciertamente, no podemos aplicar el Lema de Lax-Milgram, al menos directamente, al problema (7.31). Para remediar esta situación, se considera un elemento  $\bar{g} \in H^1(\Omega)$  tal que  $\bar{g} = g$  en  $\Gamma$ . La existencia de tal  $\bar{g}$  está garantizada por el Teorema de Trazas (ver [7, Capítulo 1]).

Introduciendo la nueva incógnita  $\tilde{u} := u - \bar{g}$ , observamos que  $\tilde{u} \in H_0^1(\Omega)$ , y luego (7.31) es equivalente a: Hallar  $\tilde{u} \in V := H_0^1(\Omega)$  tal que

$$A(\tilde{u}, v) = F(v) \quad \forall v \in V, \quad (7.32)$$

donde  $A$  es la forma bilineal usual (cf. (7.7)) y  $F : V \rightarrow \mathbf{R}$  es el funcional

$$F(v) = \int_{\Omega} f v \, dx - A(\bar{g}, v) \quad \forall v \in V.$$

Utilizando los resultados obtenidos en la Sección 7.2 y aplicando el Lema de Lax-Milgram, se concluye que (7.32) tiene una única solución  $\tilde{u} \in H_0^1(\Omega)$ . De este modo,  $u := \tilde{u} + \bar{g}$  es la única solución de la formulación original (7.31).

Para ilustrar el esquema de Galerkin en este caso no-homogéneo, consideramos nuevamente el ejemplo bi-dimensional de la Sección 7.3. Además de la triangulación  $\mathcal{T}$  y del subespacio  $V_n$ , se introduce ahora la función

$$\bar{g}_n := \sum_{j=n+1}^m g(\bar{x}_j) \varphi_j, \quad (7.33)$$

donde, para  $j \geq n+1$ ,  $\varphi_j$  se define de manera idéntica a como se hizo anteriormente para las funciones base de  $V_n$ . Entonces, la solución de Galerkin asociada a  $V_n$  está dada por

$$u_n := \tilde{u}_n + \bar{g}_n, \quad (7.34)$$

donde  $\tilde{u}_n \in V_n$  es la única solución del problema:

$$A(\tilde{u}_n, v_n) = \int_{\Omega} f v_n \, dx - A(\bar{g}_n, v_n) \quad \forall v_n \in V_n. \quad (7.35)$$

Note que el lado derecho de la formulación anterior se obtiene reemplazando  $\bar{g}$  por  $\bar{g}_n$  en la expresión de  $F$ .



Puesto que  $\tilde{u}_n \in V_n$ , existe  $\mu := (\tilde{\mu}_1, \tilde{\mu}_2, \dots, \tilde{\mu}_n) \in \mathbf{R}^n$  tal que  $\tilde{u}_n = \sum_{j=1}^n \tilde{\mu}_j \varphi_j$ . Luego, (7.35) se reduce a: Hallar  $\mu \in \mathbf{R}^n$  tal que

$$\mathbf{A} \mu = \mathbf{F}, \quad (7.36)$$

donde la matriz de rigidez  $\mathbf{A} := (a_{ij})_{n \times n}$  y el vector de carga  $\mathbf{F} := (f_j)_{n \times 1}$  están dados por

$$a_{ij} = \int_{\Omega} \nabla \varphi_i \cdot \nabla \varphi_j \, dx \quad \text{y} \quad f_j = \int_{\Omega} f \varphi_j \, dx - \sum_{i=n+1}^m g(\bar{x}_i) a_{ij}. \quad (7.37)$$

Aquí, la expresión de  $a_{ij}$  es válida para todo  $i \in \{1, \dots, m\}$  y para todo  $j \in \{1, \dots, n\}$ .

El análisis matricial y los procedimientos indicados en la Sección 7.3 siguen siendo válidos, con algunas modificaciones menores, para el presente sistema (7.36), y por lo tanto se omite cualquier comentario adicional al respecto.

Ahora, con el objeto de deducir una estimación *tipo Cea* para el error  $\|u - u_n\|_{H^1(\Omega)}$ , se necesita el siguiente subconjunto de  $H^1(\Omega)$ ,

$$W_n := \left\{ v \in C(\bar{\Omega}) : v \Big|_{K_k} \in P_1(K_k) \quad \forall k = \overline{1, p}, v(\bar{x}_i) = g(\bar{x}_i) \quad \forall i = \overline{n+1, m} \right\}.$$

**Lema 7.3.** Sea  $u \in H^1(\Omega)$  la solución de (7.31) y sea  $u_n \in W_n$  su aproximación de Galerkin dada por (7.34). Entonces existe  $C > 0$  tal que

$$\|u - u_n\|_{H^1(\Omega)} \leq C \inf_{w_n \in W_n} \|u - w_n\|_{H^1(\Omega)}. \quad (7.38)$$

*Demostración.* Notemos primero que  $u_n$  pertenece efectivamente a  $W_n$  ya que  $\tilde{u}_n \in V_n$  y  $\tilde{g}_n \in W_n$ . También, dado que en general  $g \neq \tilde{g}_n$  en  $\Gamma$ , es preciso enfatizar que  $(u - u_n) \notin H_0^1(\Omega)$ , y por lo tanto la presente demostración difiere de la dada para el Lema de Cea (ver Lema 5.1).

Observemos, sin embargo, de (7.31), (7.34) y (7.35), que la relación de ortogonalidad aún se cumple, esto es

$$A(u - u_n, v_n) = 0 \quad \forall v_n \in V_n. \quad (7.39)$$

Entonces, puesto que para todo  $w_n \in W_n$  se tiene  $(u_n - w_n) \in V_n \subseteq H_0^1(\Omega)$ , podemos aplicar la desigualdad de Poincaré (7.10), y luego usar (7.39) y el acotamiento de  $A$ , para concluir que

$$\alpha \|u_n - w_n\|_{H^1(\Omega)}^2 \leq A(u_n - w_n, u_n - w_n) = A(u - w_n - (u - u_n), u_n - w_n)$$

$$= A(u - w_n, u_n - w_n) \leq \|u - w_n\|_{H^1(\Omega)} \|u_n - w_n\|_{H^1(\Omega)},$$

de donde

$$\|u_n - w_n\|_{H^1(\Omega)} \leq \frac{1}{\alpha} \|u - w_n\|_{H^1(\Omega)}.$$

Finalmente, usando la desigualdad triangular y la estimación anterior, se obtiene

$$\begin{aligned} \|u - u_n\|_{H^1(\Omega)} &\leq \|u - w_n\|_{H^1(\Omega)} + \|u_n - w_n\|_{H^1(\Omega)} \\ &\leq \left[1 + \frac{1}{\alpha}\right] \|u - w_n\|_{H^1(\Omega)} \quad \forall w_n \in W_n, \end{aligned}$$

lo cual completa la demostración. ■

Es posible demostrar también que  $\inf_{w_n \in W_n} \|u - w_n\|_{H^1(\Omega)} \leq Ch$ , donde  $C$  y  $h$  tienen el mismo significado dado al final de la Sección 7.3 (ver [1]).

## Referencias

- [1] Brenner, S.C. and Scott, L.R.: *The Mathematical Theory of Finite Element Methods*. Springer-Verlag, 1994.
- [2] Brezzi, F. and Fortin, M.: *Mixed and Hybrid Finite Element Methods*. Springer-Verlag, 1991.
- [3] Ciarlet, P.: *The Finite Element Method for Elliptic Problems*. North-Holland, 1978.
- [4] Gatica, G.N. and Hsiao, G.C.: *Boundary-Field Equation Methods for a Class of Nonlinear Problems*. Pitman Research Notes in Mathematics Series, vol. 331, Longman, 1995.
- [5] Girault, V. and Raviart, P.-A.: *Finite Element Methods for Navier-Stokes Equations. Theory and Algorithms*. Springer-Verlag, 1986.
- [6] Quarteroni, A. and Valli, A.: *Numerical Approximation of Partial Differential Equations*. Springer-Verlag, 1994.
- [7] Raviart, P.-A. and Thomas, J.-M.: *Introduction à L'Analyse Numérique des Equations aux Dérivées Partielles*. Masson, 1983.