

## Capítulo 11

# Métodos numéricos de resolución de ecuaciones diferenciales ordinarias

### Resumen

En este capítulo se tratan discretizaciones de dos tipos de problemas de ecuaciones diferenciales ordinarias: el problema de valor inicial de Cauchy, y, los problemas de valores en la frontera 1d. Se tratan los métodos clásicos de discretizaciones de problemas de valor inicial de Cauchy como son: la familia de los métodos de Runge-Kutta y el método  $\theta$  que conduce a los métodos de Euler explícito e implícito y al método implícito de Crank-Nicolson. Por otro lado, se presenta el método de discretización del tipo Petrov-Galerkin que conduce a métodos implícitos del tipo Crank-Nicolson. La discretización de los problemas de valores en la frontera 1d se limita a las ecuaciones diferenciales de segundo orden y se aplica el método de diferencias finitas. Se estudia la consistencia, la estabilidad y la convergencia del método. Se concluye con la discretización en mallas no uniformes de ecuaciones diferenciales de segundo orden con valores en la frontera. Al final del capítulo se incluye una amplia bibliografía.

### 11.1. Introducción

En ingeniería y las ciencias físicas y químicas, las ciencias biológicas, la economía y sociales surgen modelos gobernados por ecuaciones diferenciales ordinarias, es decir, ecuaciones de la forma:

$$\begin{aligned}u'(t) &= f(t, u(t)) \quad t \in ]0, T[, \\u(0) &= u_0,\end{aligned}$$

donde  $T > 0$ ,  $f$  es una función real definida en  $[0, T] \times \mathbb{R}$ , que en lo sucesivo supondremos al menos continua,  $u_0 \in \mathbb{R}$  se denomina condición inicial y  $u$  es una función real definida en el intervalo  $[0, T]$ ,  $u$  es la función incógnita. El problema de hallar una función  $u$  solución de la ecuación diferencial y que satisfaga la condición inicial, se conoce con el nombre de problema de Cauchy de valor inicial. Con más generalidad, se considera el problema de Cauchy de valor inicial siguiente:

$$\begin{aligned}\vec{u}'(t) &= \vec{f}(t, \vec{u}(t)) \quad t \in ]0, T[, \\ \vec{u}(0) &= \vec{u}_0,\end{aligned}$$

donde  $\vec{f}$  es una función  $[0, T] \times \mathbb{R}^n$  en  $\mathbb{R}^n$ ,  $\vec{u}_0 \in \mathbb{R}^n$ .

Cuando  $\vec{f}(t, \vec{u}(t)) = A\vec{u}(t)$   $t \in [0, T]$ , se tiene un sistema lineal de ecuaciones diferenciales ordinarias de primer orden, donde  $A = (a_{ij})$  es una matriz real de  $n \times n$ . En forma explícita se escribe como sigue:

$$\begin{cases} u'_1(t) = a_{11}u_1(t) + \cdots + a_{1n}u_n(t) \\ \vdots \\ u'_n(t) = a_{n1}u_1(t) + \cdots + a_{nn}u_n(t) \end{cases} \quad t \in ]0, T[,$$

que se le conoce como sistema de ecuaciones diferencial lineal autónomo.

Una ecuación diferencial lineal homogénea de orden  $n$  con coeficientes constantes es una ecuación de la forma

$$u^{(n)}(t) + a_1 u^{(n-1)}(t) + \cdots + a_n u(t) = 0 \quad t \in [0, T],$$

donde  $a_1, \dots, a_n \in \mathbb{R}$ . Este tipo de ecuaciones diferenciales se transforma en un sistema de ecuaciones

diferenciales lineales como el precedente mediante la siguiente transformación:  $\vec{u}(t) = \begin{bmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(n-1)}(t) \end{bmatrix}$ ,

resulta

$$\begin{aligned} \vec{u}'(t) &= \begin{bmatrix} u'(t) \\ u''(t) \\ \vdots \\ u^{(n)}(t) \end{bmatrix} = \begin{bmatrix} u'(t) \\ u''(t) \\ \vdots \\ -a_1 u^{(n-1)}(t) - \cdots - a_n u(t) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & \cdots & 0 \\ & \vdots & & \\ -a_n & -a_{n-1} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(n-1)}(t) \end{bmatrix}. \end{aligned}$$

Ponemos  $\vec{u}(t) = \begin{bmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(n-1)}(t) \end{bmatrix} \in \mathbb{R}^n$ ,  $A = \begin{bmatrix} 0 & 1 & \cdots & 0 \\ & \vdots & & \\ -a_n & -a_{n-1} & \cdots & -a_1 \end{bmatrix} \in M_{n \times n}[\mathbb{R}]$ , resulta  $\vec{u}'(t) = A \vec{u}(t)$   $t \in [0, T]$ . Con mayor generalidad, una ecuación diferencial lineal no homogénea de orden  $n$  con coeficientes constantes es una ecuación de la forma

$$u^{(n)}(t) + a_1 u^{(n-1)}(t) + \cdots + a_n u(t) = f(t) \quad t \in [0, T],$$

donde  $a_1, \dots, a_n \in \mathbb{R}$ ,  $f$  es una función real definida en  $[0, T]$  y que se supondrá allí continua. Este tipo de ecuaciones diferenciales se transforma en un sistema de ecuaciones diferenciales lineales como el

precedente si se define  $\vec{u}(t) = \begin{bmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(n-1)}(t) \end{bmatrix}$ , entonces

$$\begin{aligned} \vec{u}'(t) &= \begin{bmatrix} u'(t) \\ u''(t) \\ \vdots \\ u^{(n)}(t) \end{bmatrix} = \begin{bmatrix} u'(t) \\ u''(t) \\ \vdots \\ f(t) - a_1 u^{(n-1)}(t) - \cdots - a_n u(t) \end{bmatrix} \\ &= \begin{bmatrix} 0 & 1 & \cdots & 0 \\ & \vdots & & \\ -a_n & -a_{n-1} & \cdots & -a_1 \end{bmatrix} \begin{bmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(n-1)}(t) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ f(t) \end{bmatrix}. \end{aligned}$$

Se pone  $\vec{u}(t) = \begin{bmatrix} u(t) \\ u'(t) \\ \vdots \\ u^{(n-1)}(t) \end{bmatrix} \in \mathbb{R}^n$ ,  $\vec{b}(t) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ f(t) \end{bmatrix} \in \mathbb{R}^n$ ,  $t \in [0, T]$  y  $A$  la matriz arriba definida.

Resulta  $\vec{u}'(t) = A \vec{u}(t) + \vec{b}(t)$   $t \in [0, T]$ . En este caso,  $\vec{f}(t, \vec{u}(t)) = A \vec{u}(t) + \vec{b}(t)$ .

Si  $\vec{f}(t, \vec{u}(t)) = A(t)\vec{u}(t)$ , donde para cada  $t \in [0, T]$ ,  $A(t) = (a_{ij}(t))$  es una matriz real de  $n \times n$  dependiente de  $t$ , el sistema de ecuaciones diferenciales  $\vec{u}'(t) = A(t)\vec{u}(t)$  se conoce como sistema lineal no autónomo.

Si  $\vec{u}'(t) = \vec{f}(t, \vec{u}(t))$  con  $\vec{f}$  una función de  $\mathbb{R}^n$  en  $\mathbb{R}^n$  que no se expresa como  $A\vec{u}$ , se llama sistema de ecuaciones diferenciales no lineal autónomo.

En lo que sigue, supondremos que  $\vec{f}$  es una función  $[0, T] \times \mathbb{R}^n$  en  $\mathbb{R}^n$  lipchisiana respecto de la segunda variable, es decir, existe  $k > 0$  tal que  $\|\vec{f}(t, \vec{u}_1) - \vec{f}(t, \vec{u}_2)\| \leq k \|\vec{u}_1 - \vec{u}_2\| \quad \forall \vec{u}_1, \vec{u}_2 \in \mathbb{R}^n, \forall t \in [0, T]$ . Esta hipótesis garantiza la existencia de una sola solución  $\vec{u} \in C^1([0, T])^n$ .

## 11.2. El método $\theta$

Sean  $T > 0$ ,  $u_0 \in \mathbb{R}$  y  $f$  una función real definida en  $[0, T] \times \mathbb{R}$  que suponemos lipchisiana; esto es, existe  $L > 0$  tal que para todo  $y_1, y_2 \in \mathbb{R}$ , se verifica

$$|f(t, y_1) - f(t, y_2)| \leq L |y_1 - y_2| \quad \forall t \in [0, T].$$

Consideramos el problema de valor inicial de Cauchy:

$$\begin{cases} u'(t) = f(t, u(t)) & t \in ]0, T[, \\ u(0) = u_0. \end{cases}$$

De la hipótesis sobre  $f$ , esta ecuación tiene una única solución  $u \in C^1([0, T])$ .

En muy pocos casos se puede resolver esta ecuación directamente y obtener la solución exacta  $u$ . En la generalidad de los casos, la solución  $u$  no puede obtenerse directamente y debe recurrirse a los métodos numéricos, esto significa que podemos calcular soluciones aproximadas de  $u$ .

Sean  $n \in \mathbb{Z}^+$ ,  $\tau(n) = \{t_j \mid j = 0, 1, \dots, n\}$  una partición del intervalo  $[0, T]$ , esto es,  $t_0 = 0$ ,  $t_{j-1} < t_j \quad j = 1, \dots, n$ ,  $t_n = T$ . Ponemos  $h_j = t_j - t_{j-1} \quad j = 1, \dots, n$  y  $\hat{h} = \max_{j=1, \dots, n} h_j$ . En el caso de la partición uniforme, tenemos  $h = \frac{T}{n}$ ,  $t_j = jh \quad j = 0, 1, \dots, n$  con lo que  $\tau(n) = \{jh \mid j = 0, 1, \dots, n\}$  y  $\hat{h} = h$ .

Denotamos con  $u_j$  una aproximación de  $u(t_j)$ . El valor de  $f(t_j, u(t_j))$  se aproxima como  $f(t_j, u_j)$ .

La derivada  $u'(t_j)$  lo aproximamos mediante una diferencia finita progresiva de primer orden, esto es,

$$u'(t_j) \simeq \frac{u(t_{j+1}) - u(t_j)}{h_{j+1}} \quad j = 0, 1, \dots, n-1,$$

entonces  $u'(t_j)$  se aproxima como

$$u'(t_j) \simeq \frac{u_{j+1} - u_j}{h_{j+1}} \quad j = 0, 1, \dots, n-1.$$

Sea  $\theta \in [0, 1]$ . El método  $\theta$  consiste en discretizar la ecuación diferencial mediante el esquema numérico siguiente:

$$\frac{u_{j+1} - u_j}{h_{j+1}} = \theta f(t_j, u_j) + (1 - \theta) f(t_{j+1}, u_{j+1}) \quad j = 0, 1, \dots, n-1,$$

cuyos datos son los pasos temporales  $h_j$ ,  $j = 1, \dots, n$ , los tiempos  $t_j \quad j = 0, 1, \dots, n$ . De este esquema numérico, se tiene interés en tres métodos numéricos conocidos como Euler explícito, Euler implícito y Crank-Nicolson.

### 1. Método de Euler explícito

Para  $\theta = 1$  se obtiene el siguiente esquema numérico

$$\frac{u_{j+1} - u_j}{h_{j+1}} = f(t_j, u_j) \quad j = 0, 1, \dots, n-1$$

y de este resultado

$$u_{j+1} = u_j + h_{j+1} f(t_j, u_j) \quad j = 0, 1, \dots, n-1,$$

que se conoce como esquema numérico de Euler explícito.

La razón de ser método explícito se explica a continuación. Para  $j = 0$ , el esquema numérico precedente se expresa como

$$u_1 = u_0 + h_1 f(t_0, u_0).$$

Note que a partir de los datos conocidos: condición inicial  $u(0) = u_0$ , paso temporal  $h_1$ , tiempo  $t_0$ , podemos calcular directamente  $u_1$  al instante  $t_1$ . Con  $u_1$  calculado y los datos: paso temporal  $h_2$ , tiempo  $t_1$  pasamos a calcular  $u_2$  al instante  $t_2$  mediante el esquema numérico que se obtiene haciendo  $j = 1$ , esto es,

$$u_2 = u_1 + h_2 f(t_1, u_1).$$

Así sucesivamente.

En el caso particular de una partición uniforme, el método de Euler explícito se escribe como

$$u_{j+1} = u_j + h f(t_j, u_j) \quad j = 0, 1, \dots, n-1.$$

Más adelante estudiamos la convergencia del método de Euler explícito.

## 2. Método de Euler implícito

En el método  $\theta$  hacemos  $\theta = 0$ , obtenemos

$$\frac{u_{j+1} - u_j}{h_{j+1}} = f(t_{j+1}, u_{j+1}) \quad j = 0, 1, \dots, n-1,$$

y de esta igualdad resulta

$$u_{j+1} - h_{j+1} f(t_{j+1}, u_{j+1}) = u_j \quad j = 0, 1, \dots, n-1,$$

que se conoce como esquema numérico de Euler implícito.

Para  $j = 0$  se tiene la siguiente ecuación

$$u_1 - h_1 f(t_1, u_1) = u_0$$

cuya incógnita es  $u_1$  al instante  $t_1$ . Solo en muy pocos casos se puede resolver esta ecuación directamente y calcularse  $u_1$ . En la generalidad de los casos,  $u_1$  se calcula en forma aproximada como solución de dicha ecuación. Con  $u_1$  calculado al instante  $t_1$  se pasa inmediatamente a calcular  $u_2$  como solución aproximada de la ecuación que se obtiene con  $j = 1$ , así

$$u_2 - h_2 f(t_2, u_2) = u_1.$$

El proceso continua hasta calcular  $u_3, \dots, u_{n-1}$  en los instantes  $t_3, \dots, t_n = T$ .

Para una partición uniforme, el esquema numérico de Euler implícito se escribe como

$$u_{j+1} - h f(t_{j+1}, u_{j+1}) = u_j \quad j = 0, 1, \dots, n-1.$$

Para calcular  $u_j$   $j = 1, \dots, n$ , definimos la función real  $G$  como

$$G(x) = x - h_{j+1} f(t_{j+1}, x) - u_j \quad x \in \mathbb{R}.$$

Como  $u_{j+1} - h_{j+1} f(t_{j+1}, u_{j+1}) - u_j = 0$ , se sigue que  $G(u_{j+1}) = 0$ , es decir que  $\hat{x} = u_{j+1}$  es raíz de la ecuación  $G(x) = 0$ . Esta raíz  $\hat{x}$  es aproximada aplicando cualquiera de los métodos numéricos de resolución de ecuaciones no lineales que por supuesto dependen de la regularidad de la función  $f$ .

El análisis de la convergencia del método de Euler implícito lo haremos más adelante.

### 3. Método de Crank-Nicolson

Este método fue propuesto por J. Crank y P. Nicolson en 1947. Para  $\theta = \frac{1}{2}$  se tiene el siguiente esquema numérico

$$\frac{u_{j+1} - u_j}{h_{j+1}} = \frac{1}{2}f(t_j, u_j) + \frac{1}{2}f(t_{j+1}, u_{j+1}) \quad j = 0, 1, \dots, n-1$$

y de esta igualdad obtenemos

$$u_{j+1} - \frac{h_{j+1}}{2}f(t_{j+1}, u_{j+1}) = u_j + \frac{h_{j+1}}{2}f(t_j, u_j) \quad j = 0, 1, \dots, n-1$$

que se conoce como esquema numérico de Crank-Nicolson. Este es un esquema numérico implícito.

Al igual que en el método de Euler implícito, para  $j = 0$  y con los datos: paso temporal  $h_1$ , tiempo  $t_0$ , condición inicial  $u(0) = u_0$  se tiene la ecuación

$$u_1 - \frac{h_1}{2}f(t_1, u_1) = u_0 + \frac{h_1}{2}f(t_0, u_0)$$

cuya incógnita es  $u_1$  la misma que se aproxima como solución numérica de dicha ecuación. Calculado  $u_1$  al instante  $t_1$ , con datos el paso temporal  $h_2$ , los tiempos  $t_1$  y  $t_2$ , podemos calcular el valor aproximado de  $u_2$  como solución numérica de la ecuación

$$u_2 - \frac{h_2}{2}f(t_2, u_2) = u_1 + \frac{h_2}{2}f(t_1, u_1).$$

Así sucesivamente.

En el caso de una partición uniforme, el esquema numérico de Crank-Nicolson se escribe como

$$u_{j+1} - \frac{h}{2}f(t_{j+1}, u_{j+1}) = u_j + \frac{h}{2}f(t_j, u_j) \quad j = 0, 1, \dots, n-1.$$

De manera similar que el método de Euler implícito, definimos la función real  $G$  como

$$G(x) = x - \frac{h}{2}f(t_{j+1}, x) - u_j - \frac{h}{2}f(t_j, u_j) \quad x \in \mathbb{R}.$$

Resulta que  $\hat{x} = u_{j+1}$  es raíz de la ecuación  $G(x) = 0$ . El método de resolución numérica de la ecuación no lineal  $G(x) = 0$  está relacionado con la regularidad de la función  $f$ .

La convergencia de este método será tratado más adelante.

Los esquemas numéricos de Euler implícito, explícito, y de Crank - Nicolson obtenidos para una sola ecuación diferencial pueden extenderse inmediatamente a los sistemas de ecuaciones diferenciales:

$$\begin{cases} \vec{u}'(t) = \vec{F}(t, \vec{u}(t)) & t \in ]0, T[, \\ \vec{u}(0) = \vec{u}_0, \end{cases}$$

donde  $T > 0$ ,  $\vec{u}_0^T = (u_1^{(0)}, \dots, u_m^{(0)}) \in \mathbb{R}^m$ ,  $\vec{F}^T = (f_1, \dots, f_m)$  una función vectorial de  $[0, T] \times \mathbb{R}^m$  en  $\mathbb{R}^m$  que suponemos lipschisiana. Este sistema de ecuaciones diferenciales se expresa como

$$\begin{cases} u_1'(t) = f_1(t, \vec{u}(t)) \\ \vdots \\ u_m'(t) = f_m(t, \vec{u}(t)), \end{cases} \quad t \in ]0, T[.$$

Sea  $\tau(n) = \{t_j \mid j = 0, 1, \dots, n\}$  una partición del intervalo  $[0, T]$ . Denotamos con  $\vec{u}^T = (u_1^{(j)}, \dots, u_m^{(j)})$  una aproximación de  $\vec{u}^T(t_j)$   $j = 1, \dots, n$ . Los esquemas numéricos anteriores se expresan como sigue.

1. Euler explícito

$$\vec{u}_{j+1} = \vec{u}_j + h_{j+1} \vec{F}(t_j, \vec{u}_j) \quad j = 0, 1, \dots, n-1.$$

2. Euler implícito

$$\vec{u}_{j+1} - h_{j+1} \vec{F}(t_{j+1}, \vec{u}_{j+1}) = \vec{u}_j \quad j = 0, 1, \dots, n-1.$$

3. Crank - Nicolson

$$\vec{u}_{j+1} - \frac{h_{j+1}}{2} \vec{F}(t_{j+1}, \vec{u}_{j+1}) = \vec{u}_j + \frac{h_{j+1}}{2} \vec{F}(t_j, \vec{u}_j) \quad j = 0, 1, \dots, n-1.$$

Más particularmente en el caso de las funciones vectoriales del tipo

$$\vec{F}(t, \vec{y}) = A(t) \vec{y} + \vec{g}(t) \quad t \in [0, T],$$

donde para cada  $t \in [0, T]$ ,  $A(t) = (a_{ij}(t))$  es una matriz de  $m \times m$  no nula,  $\vec{g}^T(t) = (g_1(t), \dots, g_m(t))$  con  $g_j \in C([0, T])$   $j = 1, \dots, m$ .

El método de Euler explícito se escribe como sigue:

$$\begin{aligned} \vec{u}_{j+1} &= \vec{u}_j + h_{j+1} [A(t_j) \vec{u}_j + \vec{g}(t_j)] \\ &= [I + h_{j+1} A(t_j)] \vec{u}_j + h_{j+1} \vec{g}(t_j) \quad j = 0, 1, \dots, n-1, \end{aligned}$$

donde  $I$  denota la matriz identidad de  $m \times m$ .

El método de Euler implícito se escribe como

$$\vec{u}_{j+1} - h_{j+1} [A(t_{j+1}) \vec{u}_{j+1} + \vec{g}(t_{j+1})] = \vec{u}_j \quad j = 0, 1, \dots, n-1,$$

que a su vez podemos expresarlo como el siguiente

$$(I - h_{j+1} A(t_{j+1})) \vec{u}_{j+1} = \vec{u}_j + h_{j+1} \vec{g}(t_{j+1}) \quad j = 0, 1, \dots, n-1.$$

El método de Crank - Nicolson se expresa de la manera siguiente:

$$\vec{u}_{j+1} - \frac{h_{j+1}}{2} [A(t_{j+1}) \vec{u}_{j+1} + \vec{g}(t_{j+1})] = \vec{u}_j + \frac{h_{j+1}}{2} (A(t_j) \vec{u}_j + \vec{g}(t_j)) \quad j = 0, 1, \dots, n-1,$$

Luego

$$\left(I - \frac{h_{j+1}}{2} A(t_{j+1})\right) \vec{u}_{j+1} = \left(I + \frac{h_j}{2} A(t_j)\right) \vec{u}_j + \frac{1}{2} h_{j+1} (\vec{g}(t_j) + \vec{g}(t_{j+1})) \quad j = 0, 1, \dots, n-1.$$

Se observa que tanto en el método de Euler implícito como en el de Crank-Nicolson, que para calcular  $\vec{u}_{j+1}$  se requieren que las matrices  $I - h_{j+1} A(t_{j+1})$  e  $I - \frac{h_{j+1}}{2} A(t_{j+1})$  sean invertibles.

### 11.3. Método de Petrov-Galerkin.

Sean  $T > 0$ ,  $u_0 \in \mathbb{R}$ ,  $f$  una función real definida en  $[0, T] \times \mathbb{R}$  que suponemos lipschisiana. Consideramos el problema de valor inicial de Cauchy siguiente: hallar una función  $u$  definida en  $[0, T]$  solución de

$$\begin{cases} u'(t) = f(t, u(t)) & t \in ]0, T[, \\ u(0) = u_0. \end{cases}$$

Antes de describir el método de Petrov-Galerkin para resolver el problema de Cauchy precedente, requerimos introducir algunas notaciones y algunos espacios de funciones.

Recordemos que una función  $s$  se dice escalonada en  $[0, T]$  si y solo si existe una partición  $\tau(n) = \{t_j \mid j = 0, 1, \dots, n\}$  del intervalo  $[0, T]$  tal que  $s(t) = s_i \quad t \in ]t_{j-1}, t_j[ \quad j = 1, \dots, n$ , donde  $s_i \in \mathbb{R}$ . En

los puntos  $t_j$   $j = 0, 1, \dots, n$  la función  $s$  debe estar definida de cualquier modo. En la figura siguiente se muestra una función escalonada en  $[0, T]$ .

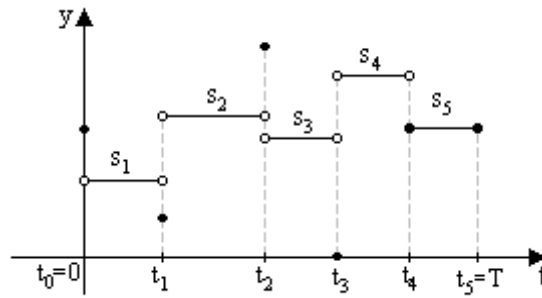


Figura 94

**Definición 1** Diremos que una función real  $f$  definida en  $[0, T]$  es discontinua en  $a \in [0, T]$  con salto de primera especie si y solo si  $f$  es discontinua en  $a$ , y  $|f(a^+) - f(a^-)| < \infty$ , donde  $f(a^+) = \lim_{h \rightarrow 0, h > 0} f(a+h)$ ,  $f(a^-) = \lim_{h \rightarrow 0, h < 0} f(a+h)$ , son los límites por derecha e izquierda respectivamente.

Si  $a = 0$  se tiene únicamente el límite por la derecha finito, esto es  $|f(0^+)| < \infty$  y si  $a = T$ , se tiene el límite por la izquierda  $|f(T^-)| < \infty$ .

**Definición 2** Sea  $f$  una función real definida en  $[0, T]$ . Decimos que  $f$  es acotada y continua a trozos en  $[0, T]$  si y solo si existen  $a_k \in [0, T]$   $k = 1, \dots, m$  tales que  $0 \leq a_1 < a_2 < \dots < a_m \leq T$ ,  $f$  no es continua en  $a_k$ , y  $|f(a_k^+) - f(a_k^-)| < \infty$ .

Se tiene que  $f$  es continua en el conjunto  $[0, T] \setminus \{a_k \mid k = 1, \dots, m\}$ .

En el caso en que  $\tau(m) = \{t_j \mid j = 0, 1, \dots, m\}$  es una partición de  $[0, T]$  y  $f$  es una función acotada y continua a trozos, entonces  $f$  es continua en cada subintervalo  $]t_{j-1}, t_j[$  y  $|f(t_j^+) - f(t_j^-)| < \infty$   $j = 1, \dots, m$ . Una función escalonada en  $[0, T]$  es una función continua a trozos.

Es claro que  $f$  es continua en  $[0, T]$  si y solo si para cada  $t \in [0, T]$ ,  $f(t_j^+) = f(t_j^-) = f(t)$ .

Denotamos con  $C_d([0, T])$  el conjunto de funciones reales acotadas y continuas a trozos. Con las operaciones habituales de funciones: adición " $+$ " y producto de números reales por funciones " $\bullet$ ",  $C_d([0, T])$  es un espacio vectorial real. Se tiene que  $C([0, T]) \subset C_d([0, T])$ .

En el espacio  $C_d([0, T])$  se define el producto escalar siguiente:

$$\langle u, v \rangle = \int_0^T u(t) v(t) dt \quad \forall u, v \in C_d([0, T]).$$

Se denota con  $C_d^1([0, T])$  el espacio de funciones reales  $u$  cuya derivada  $u'$  pertenece a  $C_d([0, T])$ , esto es,

$$C_d^1([0, T]) = \{u \in C_d([0, T]) \mid u' \in C_d([0, T])\}.$$

En la figura siguiente se muestra una función  $u \in C_d^1([0, T])$ .

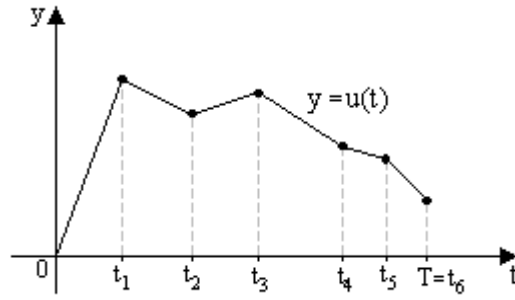


Figura 95

En la siguiente figura se muestra la función  $u'(t)$

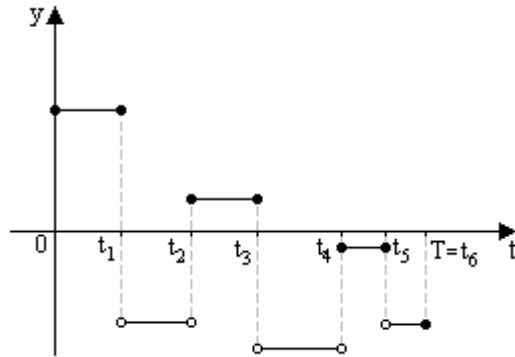


Figura 96

Introducimos el subespacio  $C_*^1([0, T])$  de  $C_d^1([0, T])$  siguiente:

$$C_*^1([0, T]) = \{v \in C_d^1([0, T]) \mid v(T) = 0\}.$$

Pasemos a describir el método de Petrov-Galerkin. Sea  $u \in C^1([0, T])$  la solución del problema de Cauchy arriba propuesto.

Sea  $v \in C_*^1([0, T])$ . Multiplicamos a la ecuación diferencial por  $v$  e integramos sobre el intervalo  $[0, T]$ , esto es,

$$\int_0^T u'(t) v(t) dt = \int_0^T f(t, u(t)) v(t) dt.$$

Apliquemos el método de integración por partes al primer miembro de la igualdad precedente, tenemos

$$\int_0^T u'(t) v(t) dt = u(t) v(t) \Big|_0^T - \int_0^T u(t) v'(t) dt = u(T) v(T) - u(0) v(0) - \int_0^T u(t) v'(t) dt$$

Puesto que  $v \in C_*^1([0, T])$  entonces  $v(T) = 0$  y como  $u$  es solución del problema de Cauchy de valor inicial, se tiene  $u(0) = u_0$ . La igualdad precedente se reduce a la siguiente

$$\int_0^T u'(t) v(t) dt = -u_0 v(0) - \int_0^T u(t) v'(t) dt$$

y en consecuencia

$$-u_0 v(0) - \int_0^T u(t) v'(t) dt = \int_0^T f(t, u(t)) v(t) dt$$



o lo que es lo mismo

$$-\int_0^T u(t) v'(t) dt = u_0 v(0) + \int_0^T f(t, u(t)) v(t) dt.$$

Así, el problema de Cauchy de valor inicial propuesto es equivalente a la ecuación precedente, la misma que es una ecuación integral. Esta ecuación integral es la que da lugar al método de Petrov-Galerkin.

Note que en esta ecuación el orden de derivación ha disminuido en 1 y la función incógnita  $u$  figura bajo el signo de integración. Ahora buscamos una función  $u$  que sea continua en  $[0, T]$  y que verifique la ecuación integral, lo que significa que se ha bajado la regularidad de la función  $u$  (antes se buscaba  $u$  de modo que la sea derivable). Esto permite amplificar el campo de acción de la solución de ecuaciones diferenciales en las que la función  $f$  cumpla condiciones de regularidad más débiles. Es precisamente esta situación la de mayor interés.

**Definición 3** Diremos que  $u \in C_d^1([0, T])$  es solución de la ecuación integral si y solo si satisface la ecuación

$$-\int_0^T u(t) v'(t) dt = u_0 v(0) + \int_0^T f(t, u(t)) v(t) dt \quad \forall v \in C_*^1([0, T]).$$

La formulación dada en la definición precedente es la conocida como método del tipo Petrov-Galerkin. Note que la función incógnita  $u$  pertenece al espacio  $C_d([0, T])$  mientras que las denominadas funciones de prueba o funciones test pertenecen al espacio  $C_*^1([0, T])$ .

Pasemos a la discretización de la formulación del método del tipo Petrov-Galerkin arriba enunciado.

Sea  $n \in \mathbb{Z}^+$ ,  $\tau(n) = \{t_j \mid j = 0, 1, \dots, n\}$  una partición del intervalo  $[0, T]$ . Ponemos  $h_j = t_j - t_{j-1}$ ,  $j = 1, \dots, n$ ,  $\hat{h} = \max_{j=1, \dots, n} h_j$ . En el caso de una partición uniforme del intervalo  $[0, T]$ , ponemos  $h = \frac{T}{n}$ ,  $\tau(n) = \{jh \mid j = 0, 1, \dots, n\}$  y  $\hat{h} = h$ .

Definimos el subespacio  $U_h$  de  $C_d([0, T])$  como sigue:

$$U_h = \left\{ u_h \in C_d([0, T]) \mid u_h|_{]t_{j-1}, t_j[} = cte \quad j = 1, \dots, n \right\}$$

donde  $u_h|_{]t_{j-1}, t_j[} = cte$  denota la restricción de la función  $u_h$  al intervalo abierto  $]t_{j-1}, t_j[$  y que en dicho intervalo la función es constante.

Una base del espacio  $U_h$  es la familia de funciones  $\{\chi_1, \dots, \chi_n\}$  definidas como se indica:

$$\chi_j(t) = \begin{cases} 1, & \text{si } t \in ]t_{j-1}, t_j[, \\ 0, & \text{si } t \in [0, T] \setminus ]t_{j-1}, t_j[, \end{cases} \quad j = 1, \dots, n.$$

Se advierte inmediatamente que cada función  $\chi_j$  es la función indicatriz del intervalo  $]t_{j-1}, t_j[$ . En la figura siguiente se muestra esta función

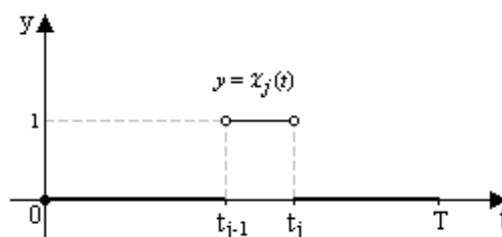


Figura 97

De la definición de espacio  $U_h$  y de la base  $\{\chi_1, \dots, \chi_n\}$  de  $U_h$  se tiene que  $u_h \in U_h$  si y solo si existe  $u_1, \dots, u_n \in \mathbb{R}$  tales que  $u_h = \sum_{j=1}^n u_j \chi_j$ .

Introducimos el subespacio  $V_h$  de  $C_*^1([0, T])$  :

$$V_h = \left\{ v_h \in C_*^1([0, T]) \mid v_h|_{]t_{j-1}, t_j[} \in \mathcal{P}_1, \quad j = 1, \dots, n \right\},$$

donde  $\mathcal{P}_1$  denota el espacio de polinomios reales de grado  $\leq 1$ , y  $v_h|_{]t_{j-1}, t_j[} \in \mathcal{P}_1$  designa la restricción de la función  $v_h$  al intervalo cerrado  $[t_{j-1}, t_j]$  en el que  $v_h$  es un polinomio de grado 1, o dicho de otro modo,  $v_h$  restringido al intervalo  $[t_{j-1}, t_j]$  es una función afín de la forma  $v_h(t) = a_j + b_j t$   $t \in [t_{j-1}, t_j]$  y  $a_j, b_j \in \mathbb{R}$  son constantes escogidas apropiadamente,  $j = 1, \dots, n$ . Una base del espacio  $V_h$  es la familia de funciones  $\{\varphi_0, \dots, \varphi_{n-1}\}$  definidas como sigue:

$$\varphi_0(t) = \begin{cases} -\frac{t-t_i}{h_i}, & \text{si } t \in [0, t_i], \\ 0, & \text{si } t \in [0, T] \setminus [0, t_1], \end{cases}$$

y para  $j = 1, \dots, n-1$  se tiene

$$\varphi_j(t) = \begin{cases} \frac{t-t_{j-1}}{h_j}, & \text{si } t \in [t_{j-1}, t_j], \\ -\frac{t-t_{j+1}}{h_{j+1}}, & \text{si } t \in ]t_j, t_{j+1}], \\ 0, & \text{si } t \in [0, T] \setminus [t_{j-1}, t_{j+1}]. \end{cases}$$

En la figura de la izquierda se muestra la gráfica de la función  $\varphi_0$  y en la derecha se muestra la gráfica de la función  $\varphi_j$ .

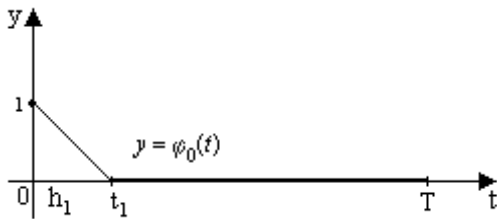


Figura 98

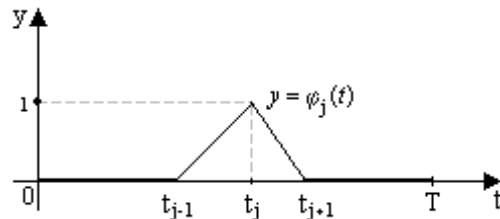


Figura 99

Estas funciones  $\varphi_j$  se las denomina funciones techo.

De la definición del espacio  $V_h$  y de la base  $\{\varphi_0, \dots, \varphi_{n-1}\}$  resulta que  $v_h \in V_h$  si y solo si existen  $v_0, \dots, v_{n-1} \in \mathbb{R}$  tales que  $v_h = \sum_{i=0}^{n-1} v_i \varphi_i$ . Note que  $v_h(T) = 0$  pues  $\varphi_i(T) = 0 \quad \forall i = 0, \dots, n-1$ .

**Definición 4** Diremos que  $u_h \in U_h$  es solución aproximada del problema de Cauchy de valor inicial si y solo si satisface la ecuación integral

$$-\int_0^T u_h(t) v_h'(t) dt = u_0 v_h(0) + \int_0^T f(t, u_h(t)) v_h(t) dt \quad \forall v_h \in V_h.$$

Esta es la formulación discreta de la formulación del tipo Petrov-Galerkin arriba definida, a la que nos referimos como formulación discreta del tipo Petrov-Galerkin.

Observe que la función incógnita  $u_h$  se busca en el espacio  $U_h$  mientras que las funciones  $v_h$ , denominadas funciones test, están en el espacio  $V_h$ .

Aplicando la formulación discreta del tipo Petrov-Galerkin se construye a continuación un esquema numérico.

Sea  $u_h \in V_h$  la solución. Existe  $u_1, \dots, u_n \in \mathbb{R}$  tales que  $u_h = \sum_{j=1}^n u_j \chi_j$ . Remplazando en la formulación discreta del tipo Petrov-Galerkin, tenemos

$$-\int_0^T \left( \sum_{j=1}^n u_j \chi_j(t) \right) v_h'(t) dt = u_0 v_h(0) + \int_0^T f \left( t, \sum_{j=1}^n u_j \chi_j(t) \right) v_h(t) dt \quad \forall v_h \in V_h,$$

y tomando en consideración la linealidad de la integral en el primer miembro, esta ecuación se expresa como

$$-\sum_{j=1}^n u_j \int_0^T \chi_j(t) v_h'(t) dt = u_0 v_h(0) + \int_0^T f \left( t, \sum_{j=1}^n u_j \chi_j(t) \right) v_h(t) dt.$$

Debemos calcular  $u_1, \dots, u_n$  y tenemos una ecuación válida para todo  $v_h \in V_h$ , en particular lo es para los elementos de la base  $\{\varphi_0, \dots, \varphi_{n-1}\}$  de  $V_h$  lo que nos permite obtener las  $n$  ecuaciones requeridas. En efecto, hacemos  $v_h = \varphi_i \in V_h$ . Tenemos

$$-\sum_{j=1}^n u_j \int_0^T \chi_j(t) \varphi_i'(t) dt = u_0 \varphi_i(0) + \int_0^T f \left( t, \sum_{j=1}^n u_j \chi_j(t) \right) \varphi_i(t) dt \quad i = 0, \dots, n-1,$$

lo que da lugar a las  $n$  ecuaciones. Determinemos estas ecuaciones.

Para  $i = 0$  se tiene  $\varphi_0(0) = 1$ ,  $\varphi_0'(t) = \begin{cases} -\frac{1}{h_1}, & \text{si } t \in ]0, t_1[, \\ 0, & \text{en otro caso,} \end{cases}$  además en el intervalo  $]0, t_1[$  intervienen  $u_1$  y  $\chi_1(t) = 1 \quad t \in ]0, t_1[$  con lo que se obtiene

$$-u_1 \int_0^{t_1} -\frac{1}{h_1} dt = u_0 + \int_0^{t_1} f(t, u_1) \varphi_0(t) dt.$$

Como  $h_1 = t_1$  e  $\int_0^{t_1} -\frac{1}{h_1} dt = -1$ , entonces

$$u_1 = u_0 + \int_0^{t_1} f(t, u_1) \varphi_0(t) dt.$$

Para  $i = 2, \dots, n$  se tiene  $\varphi_i(0) = 0$ , más aún  $\varphi_i(t_j) = \begin{cases} 1, & \text{si } i = j \\ 0, & \text{si } i \neq j, \end{cases} \quad \varphi_i'(t) = \begin{cases} \frac{1}{h_i}, & \text{si } t \in ]t_{i-1}, t_i[, \\ -\frac{1}{h_{i+1}}, & \text{si } t \in ]t_i, t_{i+1}[, \\ 0, & \text{en otro caso.} \end{cases}$  Además en los intervalos  $]t_{i-1}, t_i[$  y  $]t_i, t_{i+1}[$  intervienen  $u_i, u_{i+1}$ ,  $\chi_i(t) = 1 \quad t \in ]t_{i-1}, t_i[, \chi_{i+1}(t) = 1 \quad t \in ]t_i, t_{i+1}[$ . Entonces

$$-\left( u_i \int_{t_{i-1}}^{t_i} \frac{1}{h_i} dt + u_{i+1} \int_{t_i}^{t_{i+1}} -\frac{1}{h_{i+1}} dt \right) = \int_{t_{i-1}}^{t_i} f(t, u_i) \varphi_i(t) dt + \int_{t_i}^{t_{i+1}} f(t, u_{i+1}) \varphi_i(t) dt.$$

Puesto que  $\int_{t_{i-1}}^{t_i} \frac{1}{h_i} dt = 1$ ,  $\int_{t_i}^{t_{i+1}} -\frac{1}{h_{i+1}} dt = 1$ , se sigue que

$$-(u_i - u_{i+1}) = \int_{t_{i-1}}^{t_i} f(t, u_i) \varphi_i(t) dt + \int_{t_i}^{t_{i+1}} f(t, u_{i+1}) \varphi_i(t) dt$$

que a su vez se expresa como

$$u_{i+1} - \int_{t_i}^{t_{i+1}} f(t, u_{i+1}) \varphi_i(t) dt = u_i + \int_{t_{i-1}}^{t_i} f(t, u_i) \varphi_i(t) dt \quad i = 1, \dots, n-1.$$

En resumen, el esquema numérico que se obtiene es el siguiente:

$$\begin{cases} u_i - \int_0^{t_i} f(t, u_i) \varphi_0 dt = u_0, \\ \vdots \\ u_{i+1} - \int_{t_i}^{t_{i+1}} f(t, u_{i+1}) \varphi_i(t) dt = u_i + \int_{t_{i-1}}^{t_i} f(t, u_i) \varphi_i(t) dt \quad i = 1, \dots, n-1. \end{cases}$$

Lastimosamente en este esquema numérico se debe aún calcular las integrales. Para ello aplicamos el método de los trapecios; esto es, si  $g \in C([a, b])$  entonces  $\int_a^b g(t) dt \simeq \frac{b-a}{2} (g(a) + g(b))$ . Entonces

$$\int_0^{t_1} f(t, u_1) \varphi_0(t) dt \simeq \frac{t_1}{2} [f(0, u_1) \varphi_0(0) + f(t_1, u_1) \varphi_0(t_1)]$$

y como  $h_1 = t_1$ ,  $\varphi_0 = 1$ ,  $\varphi_0(t_1) = 0$  resulta

$$\int_0^{t_1} f(t, u_1) \varphi_0(t) dt \simeq \frac{h_1}{2} f(0, u_1).$$

De manera similar

$$\begin{aligned} \int_{t_{i-1}}^{t_i} f(t, u_i) \varphi_i(t) dt &\simeq \frac{t_i - t_{i-1}}{2} [f(t_{i-1}, u_i) \varphi_i(t_{i-1}) + f(t_i, u_i) \varphi_i(t_i)], \\ \int_{t_i}^{t_{i+1}} f(t, u_{i+1}) \varphi_i(t) dt &\simeq \frac{t_{i+1} - t_i}{2} [f(t_i, u_{i+1}) \varphi_i(t_i) + f(t_{i+1}, u_i) \varphi_i(t_{i+1})], \end{aligned}$$

y por la definición de  $h_i$ ,  $h_{i+1}$ , y de las funciones  $\varphi_1, \dots, \varphi_{n-1}$ , tenemos  $\varphi_i(t_{i-1}) = \varphi_i(t_{i+1}) = 0$ ,  $\varphi_i(t_i) = 0$ .

Entonces

$$\begin{aligned} \int_{t_{i-1}}^{t_i} f(t, u_i) \varphi_i(t) dt &\simeq \frac{h_i}{2} f(t_i, u_i), \\ \int_{t_i}^{t_{i+1}} f(t, u_{i+1}) \varphi_i(t) dt &\simeq \frac{h_{i+1}}{2} f(t_i, u_{i+1}). \end{aligned}$$

Por abuso de lenguaje designamos nuevamente con  $u_1, \dots, u_n$  a las incógnitas que satisfacen el esquema numérico siguiente:

$$\begin{cases} u_1 - \frac{h_1}{2} f(0, u_1) = u_0 \\ \vdots \\ u_{i+1} - \frac{h_{i+1}}{2} f(t_i, u_{i+1}) = u_i + \frac{h_i}{2} f(t_i, u_i) \quad i = 1, \dots, n-1. \end{cases}$$

Se observa que este esquema numérico es similar al de Crank-Nicolson al que nos referimos como esquema numérico del tipo Crank-Nicolson, el mismo que fue propuesto por HB-PB.

Otra forma de obtener este esquema numérico es la siguiente. Definimos la función  $\varphi_n$  como sigue:

$$\varphi_n(t) = \begin{cases} 0, & \text{si } t \in [0, t_{n-1}[, \\ \frac{t - t_{n-1}}{h_n}, & \text{si } t \in [t_{n-1}, T]. \end{cases}$$

En la figura siguiente se muestra la gráfica de la función  $\varphi_n$ .

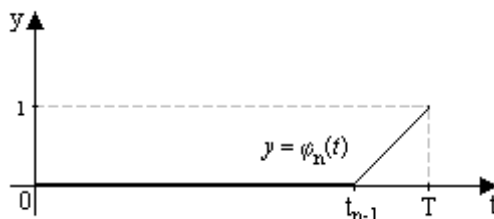


Figura 100

Introducimos los subespacios  $U_h$  y  $V_h$  de  $C_d^1([0, T])$  y  $C_*^1([0, T])$  respectivamente como sigue:

$$\begin{aligned} U_h &= \left\{ \sum_{i=0}^n u_i \varphi_i \mid u_i \in \mathbb{R} \quad i = 0, \dots, n \right\}, \\ V_h &= \left\{ \sum_{i=0}^{n-1} \alpha_i \varphi_i \mid \alpha_i \in \mathbb{R} \quad i = 0, 1, \dots, n-1 \right\}. \end{aligned}$$

Se tiene  $\dim U_h = n+1$ ,  $\dim V_h = n$ . una base de  $U_h$  es la familia de funciones  $\{\varphi_0, \dots, \varphi_n\}$  y una de  $V_h$  es  $\{\varphi_0, \dots, \varphi_{n-1}\}$ . La formulación discreta de la formulación del tipo Petrov-Galerkin se expresa como sigue: hallar una función  $u_h \in U_h$  solución de

$$-\int_0^T u_h(t) v_h'(t) dt = u_0 v_h(0) + \int_0^T f(t, u_h(t)) v_h(t) dt \quad \forall v_h \in V_h.$$

A esta acción nos referimos como formulación discreta del tipo Petrov-Galerkin.

La función  $u_h \in U_h$  se escribe como  $u_h = \sum_{i=0}^n u_i \varphi_i = u_0 \varphi_0 + \sum_{i=1}^n u_i \varphi_i$ , con  $u(0) = u_0$  la condición inicial.

Luego

$$-\int_0^T \left( u_0 \varphi_0(t) + \sum_{i=1}^n u_i \varphi_i(t) \right) v_h'(t) dt = u_0 v_h(0) + \int_0^T f \left( t, u_0 \varphi_0(t) + \sum_{i=1}^n u_i \varphi_i(t) \right) v_h(t) dt \quad \forall v_h \in V_h.$$

Se observa en esta ecuación que las incógnitas son  $u_1, \dots, u_n \in \mathbb{R}$ . Para poder calcular estas incógnitas requerimos generar  $n$  ecuaciones, para ello remplazamos sucesivamente  $v_h$  por los elementos de la base  $\{\varphi_0, \dots, \varphi_{n-1}\}$  de  $V_h$ , es decir  $v_h = \varphi_j \quad j = 0, 1, \dots, n-1$ . Así,

$$\begin{aligned} -u_0 \int_0^T \varphi_0(t) \varphi_j'(t) dt - \sum_{i=1}^{n-1} u_i \int_0^T \varphi_i(t) \varphi_j'(t) dt &= u_0 \varphi_j(0) + \\ + \int_0^T f \left( t, u_0 \varphi_0(t) + \sum_{i=1}^n u_i \varphi_i(t) \right) \varphi_j(t) dt &\quad j = 0, 1, \dots, n-1. \end{aligned}$$

Para  $j = 0$ , de la definición de  $\varphi_0$  se tiene  $\varphi_0(0) = 1$ ,  $\varphi_0(t) = 0$  si  $t \in [t_1, T]$ ,  $\varphi_0'(t) = \begin{cases} -\frac{1}{h_1}, & \text{si } 0 < t < t_1, \\ 0, & \text{si } t_1 < t < T. \end{cases}$  Tomando en consideración esta información, la ecuación precedente se reduce a la siguiente

$$-u_0 \int_0^{t_1} \varphi_0(t) \left( -\frac{1}{h_1} \right) dt - u_1 \int_0^{t_1} \varphi_1(t) \left( -\frac{1}{h_1} \right) dt = u_0 + \int_0^{t_1} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t)) \varphi_0(t) dt$$

e integrando, obtenemos

$$\frac{1}{2} u_0 + \frac{1}{2} u_1 = u_0 + \int_0^{t_1} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t)) \varphi_0(t) dt.$$

Para  $j = 1$ , de la definición de  $\varphi_1$  se tiene  $\varphi_1(t_1) = 1$ ,  $\varphi_1(0) = 0$ ,  $\varphi_1(t) = 0$  si  $t \in [t_2, T]$  y  $t = 0$ , la derivada

de  $\varphi_1$  está definida como sigue:  $\varphi_1'(t) = \begin{cases} \frac{1}{h_1}, & \text{si } 0 < t < t_1, \\ -\frac{1}{h_2}, & \text{si } t_1 < t < t_2, \\ 0, & \text{si } t_2 < t < T. \end{cases}$  Entonces  $u_h = u_0 \varphi_0 + u_1 \varphi_1 + u_2 \varphi_2$  sobre

$[0, t_2]$  en consecuencia, por la aditividad respecto del dominio de integración, tenemos

$$\begin{aligned} -u_0 \int_0^{t_1} \varphi_0(t) \varphi_1'(t) dt - u_1 \int_0^{t_1} \varphi_1(t) \varphi_1'(t) dt - u_1 \int_{t_1}^{t_2} \varphi_1(t) \varphi_1'(t) dt - u_2 \int_{t_1}^{t_2} \varphi_2(t) \varphi_1'(t) dt \\ = u_0 \varphi_1(0) + \int_0^{t_2} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t) + u_2 \varphi_2(t)) \varphi_1(t) dt, \end{aligned}$$

y de esta , resulta

$$\begin{aligned}
& -u_0 \int_0^{t_1} \varphi_0(t) \left( \frac{1}{h_1} \right) dt - u_1 \int_0^{t_1} \varphi_1(t) \frac{1}{h_1} dt - u_1 \int_{t_1}^{t_2} \varphi_1(t) \left( -\frac{1}{h_2} \right) dt - u_2 \int_{t_1}^{t_2} \varphi_2(t) \left( -\frac{1}{h_2} \right) dt \\
& = u_0 \varphi_1(0) + \int_0^{t_1} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t) + u_2 \varphi_2(t)) \varphi_1(t) dt \\
& \quad + \int_{t_1}^{t_2} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t) + u_2 \varphi_2(t)) \varphi_1(t) dt
\end{aligned}$$

Integrando los tres primeros términos obtenemos el siguiente resultado

$$\begin{aligned}
-\frac{u_0}{2} + \frac{u_2}{2} & = \int_0^{t_1} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t) + u_2 \varphi_2(t)) \varphi_1(t) dt \\
& \quad + \int_{t_1}^{t_2} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t) + u_2 \varphi_2(t)) \varphi_1(t) dt.
\end{aligned}$$

Continuando con este proceso, obtenemos para  $1 \leq j \leq n-1$ ,

$$\begin{aligned}
-\frac{1}{2}u_{j-1} + \frac{1}{2}u_{j+1} & = \int_{t_{j-1}}^{t_j} f(t, u_{j-1} \varphi_{j-1}(t) + u_j \varphi_j(t) + u_{j+1} \varphi_{j+1}(t)) \varphi_j(t) dt \\
& \quad + \int_{t_j}^{t_{j+1}} f(t, u_{j-1} \varphi_{j-1}(t) + u_j \varphi_j(t) + u_{j+1} \varphi_{j+1}(t)) \varphi_j(t) dt.
\end{aligned}$$

Para el cálculo de las integrales aplicamos la fórmula de integración del punto medio, esto es, si  $g \in C([a, b])$ ,

$$\int_a^b g(t) dt \simeq (b-a) g\left(\frac{a+b}{2}\right).$$

Entonces, de la definición de las funciones  $\varphi_0, \varphi_1$  se tiene

$$\int_0^{t_1} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t)) u_0(t) dt \simeq \frac{h_1}{2} f\left(\tilde{t}_0, \frac{1}{2}(u_0 + u_1)\right),$$

donde  $\tilde{t}_0 = \frac{1}{2}(t_0 + t_1)$ . Note que  $\varphi_0(\tilde{t}_0) = \frac{1}{2}$ .

De manera similar, de la definición de las funciones  $\varphi_0, \varphi_1, \varphi_2$  se tiene

$$\begin{aligned}
\int_0^{t_1} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t) + u_2 \varphi_2(t)) \varphi_1(t) dt & \simeq \frac{h_1}{2} f\left(\tilde{t}_0, \frac{1}{2}(u_0 + u_1)\right), \\
\int_{t_1}^{t_2} f(t, u_0 \varphi_0(t) + u_1 \varphi_1(t) + u_2 \varphi_2(t)) \varphi_2(t) dt & \simeq \frac{h_2}{2} f\left(\tilde{t}_1, \frac{1}{2}(u_1 + u_2)\right),
\end{aligned}$$

donde  $\tilde{t}_1 = \frac{1}{2}(t_1 + t_2)$ ,  $\varphi_1(\tilde{t}_0) = \frac{1}{2}$ ,  $\varphi_1(\tilde{t}_1) = \frac{1}{2}$ .

De manera general, obtenemos

$$\begin{aligned}
\int_{t_{j-1}}^{t_j} f(t, u_{j-1} \varphi_{j-1}(t) + u_j \varphi_j(t) + u_{j+1} \varphi_{j+1}(t)) \varphi_j(t) dt & = \frac{h_j}{2} f\left(\tilde{t}_{j-1}, \frac{1}{2}(u_{j-1} + u_j)\right) \\
\int_{t_j}^{t_{j+1}} f(t, u_{j-1} \varphi_{j-1}(t) + u_j \varphi_j(t) + u_{j+1} \varphi_{j+1}(t)) \varphi_j(t) dt & \simeq \frac{h_{j+1}}{2} f\left(\tilde{t}_j, \frac{1}{2}(u_j + u_{j+1})\right)
\end{aligned}$$

con  $\tilde{t}_{j-1} = \frac{1}{2}(t_{j-1} + t_j)$ ,  $\tilde{t}_j = \frac{1}{2}(t_j + t_{j+1})$ .

Resulta que para  $j = 0, j = 1, \dots, n-1$  el esquema numérico se expresa como

$$\begin{cases} \frac{1}{2}(u_0 + u_1) \simeq u_0 + \frac{h_1}{2}f\left(\tilde{t}_0, \frac{1}{2}(u_0 + u_1)\right), \\ -\frac{1}{2}u_0 + \frac{1}{2}u_2 \simeq \frac{h_1}{2}f\left(\tilde{t}_0, \frac{1}{2}(u_0 + u_1)\right) + \frac{h_2}{2}f\left(\tilde{t}_1, \frac{1}{2}(u_1 + u_2)\right), \\ \vdots \\ -\frac{1}{2}u_{j-1} + \frac{1}{2}u_{j+1} \simeq \frac{h_j}{2}f\left(\tilde{t}_{j-1}, \frac{1}{2}(u_{j-1} + u_j)\right) + \frac{h_{j+1}}{2}f\left(\tilde{t}_j, \frac{1}{2}(u_j + u_{j+1})\right). \end{cases}$$

Sumando y restando  $\frac{1}{2}u_1$  y de manera general  $\frac{1}{2}u_j$ , se tiene

$$\begin{cases} \frac{1}{2}(u_0 + u_1) - \frac{h_1}{2}f\left(\tilde{t}_0, \frac{1}{2}(u_0 + u_1)\right) \simeq u_0, \\ \frac{1}{2}u_0 + \frac{1}{2}u_2 - \frac{h_2}{2}f\left(\tilde{t}_1, \frac{1}{2}(u_1 + u_2)\right) \simeq \frac{1}{2}(u_0 + u_1) + \frac{h_1}{2}f\left(\tilde{t}_0, \frac{1}{2}(u_0 + u_1)\right), \\ \vdots \\ \frac{1}{2}(u_j + u_{j+1}) - \frac{h_{j+1}}{2}f\left(\tilde{t}_j, \frac{1}{2}(u_j + u_{j+1})\right) \simeq \frac{1}{2}(u_{j-1} + u_j) + \frac{h_j}{2}f\left(\tilde{t}_{j-1}, \frac{1}{2}(u_{j-1} + u_j)\right). \end{cases}$$

En base a este resultado obtenemos el siguiente esquema numérico: buscamos  $\tilde{u}_1, \dots, \tilde{u}_n \in \mathbb{R}$  solución del sistema de ecuaciones siguiente:

$$\begin{cases} \tilde{u}_1 - \frac{h_1}{2}f(\tilde{t}_0, \tilde{u}_1) = u_0, \\ \tilde{u}_2 - \frac{h_2}{2}f(\tilde{t}_1, \tilde{u}_2) = \tilde{u}_1 + \frac{h_1}{2}f(\tilde{t}_0, \tilde{u}_1), \\ \vdots \\ \tilde{u}_{j+1} - \frac{h_{j+1}}{2}f(\tilde{t}_j, \tilde{u}_{j+1}) = \tilde{u}_j + \frac{h_j}{2}f(\tilde{t}_{j-1}, \tilde{u}_j) \quad j = 1, 2, \dots, n-1. \end{cases}$$

que es el esquema numérico del tipo Crank-Nicolson que hemos obtenido anteriormente.

El esquema numérico obtenido para el caso escalar de una ecuación diferencial se extiende inmediatamente a sistemas de ecuaciones diferenciales. Así, consideremos el sistema de ecuaciones diferenciales siguiente:

$$\begin{cases} \vec{u}'(t) = \vec{F}(t, \vec{u}(t)) & t \in ]0, T[, \\ \vec{u}(0) = \vec{u}_0, \end{cases}$$

donde  $T > 0$ ,  $\vec{u}_0^T = (u_1^{(0)}, \dots, u_m^{(0)}) \in \mathbb{R}^m$ ,  $\vec{F}$  es una función vectorial de  $[0, T] \times \mathbb{R}^m$  en  $\mathbb{R}^m$  que suponemos lipchisiana, esto es, existe  $L > 0$  tal que  $\forall \vec{y}_1, \vec{y}_2 \in \mathbb{R}^m$ ,

$$\|\vec{F}(t, \vec{y}_1) - \vec{F}(t, \vec{y}_2)\| \leq L \|\vec{y}_1 - \vec{y}_2\| \quad \forall t \in [0, T].$$

Entonces

$$\begin{cases} \vec{u}_1 - \frac{h_1}{2}\vec{F}(\tilde{t}_0, \vec{u}_1) = \vec{u}_0, \\ \vec{u}_{j+1} - \frac{h_{j+1}}{2}\vec{F}(\tilde{t}_j, \vec{u}_{j+1}) = \vec{u}_j + \frac{h_j}{2}\vec{F}(\tilde{t}_{j-1}, \vec{u}_j) \end{cases}$$

donde  $\tilde{t}_j = \frac{1}{2}(t_{j-1} + t_j)$  los puntos medios de los intervalos  $[t_{j-1}, t_j]$  de la partición  $\tau(n)$  de  $[0, T]$ .

Más particularmente, si la función  $\vec{F}$  tiene la forma

$$\vec{F}(t, \vec{y}) = A(t) \vec{y} + \vec{b}(t) \quad t \in [0, T],$$

con  $A(t) = (a_{ij}(t))$  una matriz de  $m \times m$  no nula y cada  $a_{ij}(t)$  función continua,  $\vec{b}$  una función vectorial continua.

El esquema numérico se expresa como sigue;

$$\left\{ \begin{array}{l} \left[ I - \frac{h_1}{2} A(\tilde{t}_0) \right] \vec{u}_1 = \vec{u}_0 + \frac{h_1}{2} \vec{b}(\tilde{t}_0) \\ \vdots \\ \left[ I - \frac{h_{j+1}}{2} A(\tilde{t}_j) \right] \vec{u}_{j+1} = \left( I - \frac{h_j}{2} A(\tilde{t}_{j-1}) \right) \vec{u}_j + \frac{1}{2} h_j \vec{b}(\tilde{t}_{j-1}) + \frac{1}{2} h_{j+1} \vec{b}(\tilde{t}_j) \quad j = 1, \dots, n-1. \end{array} \right.$$

## 11.4. Método de diferencias finitas para problemas de valores en la frontera 1d.

El método de diferencias finitas (MDF) es uno de los primeros métodos que fueron implementados en la resolución numérica tanto de ecuaciones diferenciales ordinarias como en derivadas parciales para problemas uni, bi y tridimensionales. Su popularidad radica en el hecho de la simplicidad con la que se discretizan tales ecuaciones mediante el uso de aproximaciones de las derivadas por medio de cocientes incrementales.

En este capítulo iniciaremos con los operadores en diferencias finitas que luego serán aplicados a una clase de problemas con valores en la frontera unidimensionales siguientes:

$$-\frac{d}{dx} \left( p \frac{du}{dx} \right) + r \frac{du}{dx} + qu = f \quad \text{sobre } ]0, L],$$

de donde  $L > 0$ ,  $p, q, r, f \in C^0([0, L])$  tales que

- i.  $p(x) \geq \alpha > 0 \quad \forall x \in [0, L]$ ,
- ii.  $q(x) \geq 0 \quad \forall x \in [0, L]$ .

Para el problema propuesto consideramos cuatro condiciones de frontera que precisamos a continuación.

1. Condiciones de frontera de Dirichlet:  $u(0) = a_0, \quad u(L) = a_1$ .
2. Condiciones de frontera de Neumann:  $u'(0) = a, \quad u'(L) = b$ .
3. Condiciones de frontera mixtas:  $u'(0) + \alpha u(0) = a, \quad u'(L) + \beta u(L) = b$ .
4. Condiciones de frontera periódicas:  $u(0) = u(L), \quad u'(0) = u'(L)$ . En este último problema debemos suponer que las funciones  $p, q, r, f$  se extienden por periodicidad a todo  $\mathbb{R}$  y conservan la continuidad. Nótese que

$$\begin{aligned} u(x+L) &= u(x) \quad \forall x \in \mathbb{R}, \\ u'(x+L) &= u'(x) \quad \forall x \in \mathbb{R}. \end{aligned}$$

Suponemos que para el problema propuesto, se conocen resultados de existencia, unicidad, regularidad de la solución.

Con el propósito de introducir las nociones de consistencia, estabilidad y convergencia que serán abordados más adelante, consideramos el problema modelo siguiente:

$$\text{hallar } u \in C^2([0, L]) \text{ solución de } \begin{cases} -u'' + qu = f & \text{sobre } ]0, L[, \\ u(0) = u(L) = 0, \end{cases} \quad (\text{P})$$

donde  $f, q \in C^0([0, L])$  con  $q(x) \geq 0 \quad \forall x \in [0, L]$ .

Este problema es una simplificación del problema planteado en la introducción. Con este problema abordaremos otros desde el punto de vista informático que consiste en la puesta en marcha del método de diferencias finitas.



Sea  $n \in \mathbb{Z}^+$ ,  $h = \frac{L}{n}$  y  $x_j = jh$ ,  $j = 0, 1, \dots, n$ . Ponemos

$$\tau(n) = \{x_j = jh \mid j = 0, 1, \dots, n\}.$$

El conjunto  $\tau(n)$  se llama discretización del intervalo  $[0, L]$  o también malla de  $[0, L]$ .

Puesto que

$$u''(x_j) = \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1}))}{h^2} + r_h(x_j), \quad j = 1, \dots, n-1,$$

(véase el capítulo 2, diferencias finitas centrales) se sigue que (P) se discretiza del modo siguiente:

$$\begin{cases} -\frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1}))}{h^2} + q(x_j)u(x_j) + r_h(x_j) = f(x_j) & j = 1, \dots, n-1, \\ u(0) = 0, \quad u(L) = 0 \end{cases} \quad (P_1)$$

En el esquema numérico precedente  $(P_1)$  se desconocen  $u(x_j)$  y  $r_h(x_j)$ ,  $j = 1, \dots, n-1$ . Deseamos que  $r_h(x_j) \rightarrow_{h \rightarrow 0} 0$   $j = 1, \dots, n-1$ .

Denotamos con  $u_j$  una aproximación de  $u(x_j)$ . Asumimos que  $r_h(x_j) \simeq 0$  y en consecuencia, se tiene el esquema numérico siguiente.

$$\begin{cases} u_0 = 0, \quad u_n = 0, \\ -\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + q(x_j)u_j = f(x_j) & j = 1, \dots, n-1. \end{cases}$$

Para  $j = 1$ , se tiene

$$\left[ \frac{2}{h^2} + q(x_1) \right] u_1 - \frac{1}{h^2} u_2 = f(x_1).$$

Para  $1 < j < n-1$ ,

$$-\frac{1}{h^2} u_{j-1} + \left[ \frac{2}{h^2} + q(x_j) \right] u_j - \frac{1}{h^2} u_{j+1} = f(x_j).$$

Para  $j = n-1$ ,

$$-\frac{1}{h^2} u_{n-2} + \left[ \frac{2}{h^2} + q(x_{n-1}) \right] u_{n-1} = f(x_{n-1}).$$

El conjunto de ecuaciones precedente, en forma matricial se escribe en la siguiente forma:

$$\begin{bmatrix} \frac{2}{h^2} + q(x_1) & -\frac{1}{h^2} & 0 & \cdots & 0 \\ -\frac{1}{h^2} & \frac{2}{h^2} + q(x_2) & -\frac{1}{h^2} & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & -\frac{1}{h^2} \\ 0 & \cdots & \cdots & -\frac{1}{h^2} & \frac{2}{h^2} + q(x_{n-1}) \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ \vdots \\ u_{n-2} \\ u_{n-1} \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ \vdots \\ f(x_{n-2}) \\ f(x_{n-1}) \end{bmatrix} \quad (1)$$

Ponemos  $\vec{u}_h = (u_1, \dots, u_{n-1})^T$ ,  $\vec{b} = (f(x_1), \dots, f(x_{n-1}))^T$ ,  $A_h = (a_{ij}(h))$  con

$$\begin{aligned} a_{ii}(h) &= \frac{2}{h^2} + q(x_i) \quad i = 1, \dots, n-1, \\ a_{i-1i}(h) &= a_{ii-1}(h) = -\frac{1}{h^2}, \quad i = 2, \dots, n-1. \end{aligned}$$

El sistema de ecuaciones lineales (1) se escribe en forma compacta como

$$A_h \vec{u}_h = \vec{b}. \quad (2)$$

La matriz  $A_h$  es de  $(n-1) \times (n-1)$ , tridiagonal, simétrica y se demostrará más adelante que es definida positiva, por lo que el sistema de ecuaciones (2) tiene una única solución  $\vec{u}_n \in \mathbb{R}^{n-1}$ .

### 11.4.1. Aspectos informáticos del método de diferencias finitas

El problema propuesto en la sección precedente sugiere proponer la siguiente metodología para su resolución.

#### 1. Lectura de datos de entrada.

Leer  $L > 0$ ,  $n \in \mathbb{Z}^+$ , funciones  $q$  y  $f$ .

#### 2. Preparación de los datos de entrada.

i. Generación de la malla  $\tau(n) = \{x_j \mid j = 0, 1, \dots, n\}$ .

a. Generación manual.

b. Generación automática.

ii. Construcción de los vectores  $\vec{q}$  y  $\vec{b}$ .

$$\begin{aligned}\vec{q} &= (q(x_1), \dots, q(x_{n-1})), \\ \vec{b} &= (f(x_1), \dots, f(x_{n-1})).\end{aligned}$$

iii. Condiciones de frontera:  $u(0)$ ,  $u(L)$ .

$$\begin{aligned}u(0) &= 0, \\ u(L) &= 0.\end{aligned}$$

iv. Construcción de la matriz  $A_h = (a_{ij}(h))$ .

$$\begin{aligned}a_{ii}(h) &= \frac{2}{h^2} + q(x_i) \quad i = 1, \dots, n-1, \\ a_{i-1,i}(h) &= a_{i,i-1}(h) = -\frac{1}{h^2} \quad i = 2, \dots, n-1.\end{aligned}$$

#### 3. Ejecución del algoritmo.

Resolución del sistema de ecuaciones

$$A_h \vec{u}_h = \vec{b}. \quad (*)$$

a) Método directo: factorización  $LU$ .

b) Método iterativo:  $S.O.R$ .

#### 4. Preparación de los datos de salida.

i. Construir el archivo que contiene  $n$ ,  $\tau(n)$  y  $\vec{u}_h$ ; por ejemplo  $u_h.dat$  contiene

$$\begin{array}{ccccccc} & & & n & & & \\ x_i & u_{h,i} & i & = & 0, \dots, n. & & \end{array}$$

ii. Con el propósito de efectuar pruebas, es recomendable conocer la solución exacta  $u$ . Con esta solución se debe construir el vector  $\vec{u} = (u(0), \dots, u(L))$  y en consecuencia generar el archivo  $u.dat$  que contiene

$$\begin{array}{ccccccc} & & & n & & & \\ x_i & u_i & i & = & 0, \dots, n. & & \end{array}$$

iii. Supongamos que  $u \in V$ , donde  $V$  es un espacio de funciones provisto de la norma  $\|\cdot\|_V$ .

Calcular el error  $r(n) = \|u - u_h\|_V$  para diferentes discretizaciones. Generar un archivo que contiene  $n$  y  $e(n)$ , por ejemplo:  $e_h.dat$ :

$$n \quad e(n) \quad n = n_1, 2n_1, 4n_1, 8n_1, \dots, 2^{n_0}n_1,$$

donde  $n_0 \in \mathbb{Z}^+$  (por ejemplo:  $n_1 = 20$ ,  $n_0 = 5$ ).

### 5. Presentación de resultados.

- i. Gráficas de  $u$  y  $u_h$ .
- ii. Curva de errores.

**Nota:** Se conocen las soluciones  $u$  y  $u_h$  en los puntos  $x_i \in \tau(n)$ ; esto es, se tienen los conjuntos de puntos:

$$\begin{aligned} G_1 &= \{(x_i, u(x_i)) \mid i = 0, 1, \dots, n\}, \\ G_2 &= \{(x_i, u_{h,i}) \mid i = 0, 1, \dots, n\}. \end{aligned}$$

Al representar  $G_1$  y  $G_2$  en el sistema de coordenadas rectangulares  $XY$ , se tienen puntos en el plano. Estos puntos pueden ser unidos con segmentos de recta, polinomios de grado 2, polinomios cúbicos, etc. En definitiva, se utilizarán B-splines de orden 1, 2, 3, etc. Por sencillez se utilizarán los B-splines de orden 1. Para ello definimos

$$\varphi_i(x) = \begin{cases} \frac{x - x_{i-1}}{h_i}, & x \in [x_{i-1}, x_i], \\ -\frac{x - x_{i+1}}{h_{i+1}}, & x \in ]x_i, x_{i+1}], \quad i = 1, \dots, n-1. \\ 0, & x \in [0, L] \setminus [x_{i-1}, x_{i+1}] \end{cases}$$

Note que  $\varphi_i(x_j) = \delta_{ij}$   $i, j = 1, \dots, n-1$ ; además  $h_i = h$ ,  $i = 1, \dots, n$ .

Ponemos

$$\begin{aligned} \tilde{u}(x) &= \sum_{i=1}^{n-1} u(x_i) \varphi_i(x), \\ u_h(x) &= \sum_{i=1}^{n-1} u_{n,i} \varphi_i(x) \\ e(n) &= \sum_{k=1}^{n_0} e(n_j) \psi_k(n) \end{aligned}$$

con  $\psi_k(n)$  definida de manera análoga a  $\varphi_i(x)$ .

#### 11.4.2. Consistencia, estabilidad, convergencia

1. Sea  $L > 0$ . Consideramos el problema siguiente:

$$\text{Hallar } u \in C^2([0, L]) \text{ solución de } \begin{cases} -u'' + qu = f & \text{sobre } ]0, L[, \\ u(0) = u(L) = 0, \end{cases} \quad (1)$$

donde  $q, f \in C^0([0, L])$  con  $q(x) \geq 0 \quad \forall x \in [0, L]$ .

El esquema numérico que aproxima la solución  $u$  es el siguiente:

$$\begin{cases} u_0 = u_n = 0 \\ -\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + q(x_j)u_j = f(x_j), \quad j = 1, \dots, n-1, \end{cases} \quad (2)$$

donde  $\tau(n) = \{x_j = jh \mid j = 0, 1, \dots, n\}$  es el conjunto de nodos de  $[0, L]$ ,  $h = \frac{L}{n}$  y  $n \in \mathbb{Z}^+$ .

El esquema numérico (2), en forma matricial, se escribe:

$$A_h \vec{u}_h = \vec{b}, \quad (3)$$

con  $\vec{u}_h = (u_1, \dots, u_{n-1})^T$ ,  $\vec{b} = (f(x_1), \dots, f(x_{n-1}))$ ,  $A_h = (a_{ij}(h))$  la matriz definida por:

$$\begin{aligned} a_{ii}(h) &= \frac{2}{h^2} + q(x_i) \quad i = 1, \dots, n-1, \\ a_{ii-1}(h) &= -\frac{2}{h^2} = a_{i-1,i}(h) \quad i = 2, \dots, n-1. \end{aligned}$$

**Definición 5** Toda función  $g : \begin{cases} \tau(n) & \rightarrow \mathbb{R} \\ x_j & \rightarrow g(x_j) = y_j \end{cases}$  se llama función reticular que se escribirá  $\vec{g} = (g(x_0), \dots, g(x_n))$  o bien  $\vec{g} = (y_0, \dots, y_n)$ .

Se denota con  $V_h$  al conjunto de todas las funciones reticulares definidas en  $\tau(n)$ . Con las operaciones habituales de funciones: adición y producto por escalares,  $V_h$  es un espacio vectorial de dimensión  $n+1$ .

Se denota con  $V_0 = \{g \in V_h \mid g(0) = g(L) = 0\}$ .

**Teorema 1** Las siguientes son normas en  $V_0$ .

$$\begin{aligned} i) \quad \|g\|_\infty &= \max_{i=1, \dots, n-1} |g(x_i)| \\ ii) \quad \|g\|_2 &= \left[ \sum_{i=1}^{n-1} h g^2(x_i) \right]^{\frac{1}{2}}. \\ iii) \quad \|g\|_{1,2} &= \left[ \sum_{i=0}^{n-1} h \left( \frac{g(x_{i+1}) - g(x_i)}{h} \right)^2 \right]^{\frac{1}{2}}. \end{aligned}$$

**Demostración.** Son inmediatas (véase el apéndice, espacios normados). ■

Nota: para todo  $f, g \in V_0$ ,

$$\begin{aligned} \langle f, g \rangle &= \sum_{i=1}^{n-1} h f(x_i) g(x_i), \\ \langle f, g \rangle_{1,2} &= \sum_{i=0}^{n-1} \frac{[f(x_{i+1}) - f(x_i)][g(x_{i+1}) - g(x_i)]}{h} \end{aligned}$$

son productos escalares en  $V_0$ . Además,

$$\begin{aligned} \|f\|_2 &= \langle f, f \rangle^{\frac{1}{2}}. \\ \|f\|_{1,2} &= \langle f, f \rangle_{1,2}^{\frac{1}{2}}. \\ |\langle f, g \rangle| &\leq \|f\|_2 \|g\|_2 \quad \forall f, g \in V_0, \\ |\langle f, g \rangle| &\leq \|f\|_{1,2} \|g\|_{1,2} \quad \forall f, g \in V_0 \end{aligned}$$

**Teorema 2** Para todo  $f \in V_0$ , se verifica que

$$L^{-\frac{1}{2}} \|f\|_2 \leq \|f\|_\infty \leq L^{\frac{1}{2}} \|f\|_{1,2}.$$

**Demostración.**

i. Probemos primeramente que  $\|f\|_\infty \leq L^{\frac{1}{2}} \|f\|_{1,2} \quad \forall f \in V_0$ . En efecto,  $y_i = \sum_{j=0}^{i-1} (y_{j+1} - y_j)$  ya que  $y_0 = 0$ ,  $i = 1, \dots, n$ . Luego, aplicando la desigualdad de Cauchy-Schwarz, se tiene

$$\begin{aligned} |y_i| &= \sum_{j=0}^{i-1} |y_{j+1} - y_j| \leq \left( \sum_{j=0}^{i-1} 1^2 \right)^{\frac{1}{2}} \left( \sum_{j=0}^{i-1} (y_{j+1} - y_j)^2 \right)^{\frac{1}{2}} \\ &\leq n^{\frac{1}{2}} \left( \sum_{j=0}^{n-1} (y_{j+1} - y_j)^2 \right)^{\frac{1}{2}} = \left[ \sum_{j=0}^{n-1} n (y_{j+1} - y_j)^2 \right]^{\frac{1}{2}}. \end{aligned}$$

Como  $h = \frac{L}{n} \Rightarrow n = \frac{L}{h}$ . Obtenemos

$$|y_i| \leq \left( \sum_{j=0}^{n-1} \frac{L}{h} (y_{j+1} - y_j)^2 \right)^{\frac{1}{2}} = L^{\frac{1}{2}} \left( \sum_{j=0}^{n-1} h \left( \frac{y_{j+1} - y_j}{h} \right)^2 \right)^{\frac{1}{2}} = L^{\frac{1}{2}} \|f\|_{1,2} \quad i = 1, \dots, n-1.$$

En consecuencia  $\|f\|_{\infty} \leq L^{\frac{1}{2}} \|f\|_{1,2}$ .

ii. Probemos que  $\|f\|_2 \leq L^{\frac{1}{2}} \|f\|_{\infty}$ .

Puesto que

$$\|f\|_2^2 = \sum_{i=1}^{n-1} h y_i^2 = h \sum_{i=1}^{n-1} y_i^2 \leq h \sum_{i=1}^{n-1} \|f\|_{\infty}^2 = h n \|f\|_{\infty}^2 = L \|f\|_{\infty}^2.$$

De la no negatividad de la norma, se sigue que  $\|f\|_2 \leq L^{\frac{1}{2}} \|f\|_{\infty}$ .

De i) y ii) se deduce

$$L^{-\frac{1}{2}} \|f\|_2 \leq \|f\|_{\infty} \leq L^{\frac{1}{2}} \|f\|_{1,2} \quad \forall f \in V_0.$$

■

Observación. Puesto que  $|y_i| \leq L^{\frac{1}{2}} \|f\|_{1,2}$   $i = 1, \dots, n$ , se sigue que

$$\|f\|_2^2 = \sum_{i=1}^{n-1} h y_i^2 \leq \sum_{i=1}^{n-1} h L \|f\|_{1,2}^2 = n h L \|f\|_{1,2}^2 = L^2 \|f\|_{1,2}^2,$$

así,

$$\|f\|_2^2 \leq L^2 \|f\|_{1,2}^2$$

de donde

$$\|f\|_2 \leq L \|f\|_{1,2}.$$

que es la análoga a la desigualdad de Poincaré:  $f \in H_0^1(o, L)$ ,  $\exists c > 0$  tal que  $\|f\|_{L^2(o, L)} \leq c \|f'\|_{L^2(0, L)}$ .

### Definición de consistencia

Sea  $\vec{U}_h = (u(x_0), \dots, u(x_n))^T$  el vector constituido por la solución exacta en los puntos  $x_i$ ,  $i = 0, \dots, n$  de la malla  $\tau(n)$ . Definimos

$$\begin{aligned} \vec{e}_h &= \vec{U}_h - \vec{u}_h && \text{el error sobre la solución numérica,} \\ \vec{r}_h &= A_h \vec{U}_h - \vec{b} && \text{error de consistencia.} \end{aligned}$$

**Definición 6** 1. Diremos que (1) y (3) son consistentes, para una norma  $\|\cdot\|$  de  $V_0$ , si  $\lim_{h \rightarrow 0} \|\vec{r}_h\| = 0$ .

**Definición 7** 2. Diremos que (1) y (3) tienen una consistencia de orden  $m > 0$ , si existe una constante  $c_1 > 0$  independiente de  $h$ , tal que

$$\|\vec{r}_h\| \leq c_1 h^m \quad \forall h > 0.$$

La consistencia es necesaria, pero no suficiente para que el sistema discreto (esquema numérico (2)) sea convergente.

Sea  $\mathcal{L} : C^2([0, L]) \rightarrow C^0([0, L])$  el operador diferencial definido por  $\mathcal{L}u = -u'' + qu$ .

Para  $x \in [0, L]$ , escribiremos  $\mathcal{L}u(x) = -u''(x) + q(x)u(x)$ .

Denotamos con  $L_{\pi} : C^0([0, L]) \rightarrow \mathbb{R}$  el operador definido por:

$$L_{\pi}u(x) = -\frac{u(x+h) - 2u(x) + u(x-h))}{h^2} + q(x)u(x),$$

donde  $h > 0$ ,  $x-h, x, x+h \in [0, L]$ .

Ponemos  $L_\pi u_i = f(x_i)$ ,  $i = 1, \dots, n-1$ . Observe que

$$L_\pi u(x_i) = -\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))}{h^2} + q(x_i)u(x_i) \quad i = 1, \dots, n-1.$$

La consistencia establece que el operador  $L_\pi$  aproxima al operador diferencial  $\mathcal{L}$  cuando  $h \rightarrow 0$ . Más precisamente, tenemos el siguiente teorema.

**Teorema 3** *Supongamos que  $u$ , solución de (1), es de clase  $C^4([0, L])$ . Entonces*

$$\begin{aligned} \|\vec{r}_h\|_\infty &\leq \frac{h^2}{12} \|u^{(iv)}\|_{L^\infty(0,L)}, \\ \|L_\pi u_h - \mathcal{L}u\|_\infty &\leq \frac{h^2}{12} \|u^{(iv)}\|_{L^\infty(0,L)}. \end{aligned}$$

**Demostración.** Esto es, el esquema numérico (2) es consistente.

De la definición de  $\vec{r}_h$ , se sigue que

$$\begin{aligned} \vec{r}_h &= A_h \vec{U}_h - \vec{b} \Leftrightarrow \\ r_{ih} &= -\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))}{h^2} + q(x_i)u(x_i) - f(x_i). \end{aligned}$$

Por otro lado,

$$-u''(x) + q(x)u(x) = f(x) \quad x \in ]0, L[,$$

entonces

$$0 = -u''(x_i) + q(x_i)u(x_i) - f(x_i), \quad i = 1, \dots, n-1,$$

de donde

$$r_{ih} = u''(x_i) - \frac{1}{h^2} (u(x_{i+1}) - 2u(x_i) + u(x_{i-1})).$$

Además,

$$\begin{aligned} u(x_{i+1}) &= u(x_i) + hu'(x_i) + \frac{h^2}{2!}u''(x_i) + \frac{h^3}{3!}u'''(x_i) + \frac{h^4}{4!}u^{(iv)}(\mathfrak{S}_1), \\ u(x_{i-1}) &= u(x_i) - hu'(x_i) + \frac{h^2}{2!}u''(x_i) - \frac{h^3}{3!}u'''(x_i) + \frac{h^4}{4!}u^{(iv)}(\mathfrak{S}_2), \end{aligned}$$

con  $\mathfrak{S}_1 \in [x_i, x_{i+1}]$ ,  $\mathfrak{S}_2 \in [x_{i-1}, x_i]$ ,  $i = 1, \dots, n-1$ .

Entonces

$$\frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))}{h^2} = u''(x_i) + \frac{h^2}{4!} (u^{(iv)}(\mathfrak{S}_1) + u^{(iv)}(\mathfrak{S}_2)),$$

con lo cual

$$\begin{aligned} r_{ih} &= -\frac{h^2}{4!} (u^{(iv)}(\mathfrak{S}_1) + u^{(iv)}(\mathfrak{S}_2)) \\ |r_{ih}| &\leq \frac{h^2}{12} \|u^{(iv)}\|_{L^\infty(0,L)}, \quad i = 1, \dots, n-1 \end{aligned}$$

de donde

$$\|\vec{r}_h\|_\infty = \max_{i=1, \dots, n-1} |r_{ih}| \leq \frac{h^2}{12} \|u^{(iv)}\|_{L^\infty(0,L)}.$$

Puesto que  $\mathcal{L}u = f$  y  $L_\pi u(x_i) = f(x_i)$   $i = 1, \dots, n-1$ , se sigue que

$$L_\pi u(x_i) - \mathcal{L}u(x_i) = u''(x_i) - \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1}))}{h^2} = r_{ih} \quad i = 1, \dots, n-1.$$

Luego

$$\|L_\pi \vec{U}_h - \mathcal{L}u\|_\infty \xrightarrow{h \rightarrow 0} 0.$$

■

Nótese que  $\|\vec{r}_h\|_\infty = \|L_\pi \vec{U}_h - \mathcal{L}u\|_\infty \leq \frac{h^2}{12} \max_{x \in [0, L]} |u^{iv}(x)|$ , que muestra que el esquema numérico es de orden 2.

**Definición 8** Sean  $\|\cdot\|_{h,1}$ ,  $\|\cdot\|_{h,2}$  dos normas en  $V_0$ . Diremos que el esquema numérico (2) es estable con respecto de las normas  $\|\cdot\|_{h,1}$ ,  $\|\cdot\|_{h,2}$ , si existe una constante  $C_2 > 0$  independiente de  $h$ , tal que

$$\|A_h^{-1} \vec{u}\|_{h,1} \leq C_2 \|\vec{u}\|_{h,2} \quad \forall \vec{u} \in V_0, \quad \forall h > 0.$$

**Definición 9** Sea  $\|\cdot\|$  una norma en  $V_0$ .

i. Se dice que el esquema numérico (2) es convergente con respecto de  $\|\cdot\|$  si  $\lim_{h \rightarrow 0} \|\vec{e}_h\| = 0$ .

ii. Se dice que el esquema numérico tiene un orden de convergencia  $p > 0$ , si existe una constante  $C_3 > 0$  independiente de  $h$  tal que  $\|\vec{e}_h\| \leq C_3 h^p \quad \forall h > 0$ .

**Teorema 4** El sistema de ecuaciones  $A_h \vec{u}_h = \vec{b}$  tiene una única solución  $u_h$ .

**Demostración.** Probemos que  $A_h$  es invertible. Para el efecto, mostremos que  $A_h$  es definida positiva, esto es,

$$\vec{u}^T A_h \vec{u} > 0 \quad \forall \vec{u} \in \mathbb{R}^{n-1} \quad \text{con} \quad \vec{u} \neq 0.$$

Sea  $\vec{u} = (u_1, \dots, u_n) \in \mathbb{R}^{n-1}$ ,  $u_0 = u_n = 0$ . Entonces

$$\begin{aligned} \vec{u}^T A_h \vec{u} &= (u_1, \dots, u_n) \begin{bmatrix} -\frac{u_0}{h^2} + \left(\frac{2}{h^2} + q(x_1)\right) u_1 - \frac{u_2}{h^2} \\ \vdots \\ -\frac{u_{i-1}}{h^2} \left(\frac{2}{h^2} + q(x_i)\right) u_i - \frac{u_{i+1}}{h^2} \\ \vdots \\ -\frac{u_{n-2}}{h^2} + \left(\frac{2}{h^2} + q(x_{n-1})\right) u_{n-1} - \frac{u_n}{h^2} \end{bmatrix} \\ &= \sum_{i=1}^{n-1} u_i \left( -\frac{u_{i-1}}{h^2} \left( \frac{2}{h^2} + q(x_i) \right) u_i - \frac{u_{i+1}}{h^2} \right) \\ &= \sum_{i=1}^{n-1} \frac{1}{h^2} u_i (-u_{i-1} + 2u_i - u_{i+1}) + \sum_{i=1}^{n-1} q(x_i) u_i^2 \\ &= \frac{1}{h^2} \left( \sum_{i=1}^{n-1} u_i (u_i - u_{i-1}) - \sum_{i=1}^{n-1} u_i (u_{i+1} - u_i) \right) + \sum_{i=1}^{n-1} q(x_i) u_i^2 \\ &= \frac{1}{h^2} \left( \sum_{i=0}^{n-2} u_{i+1} (u_{i+1} - u_i) - \sum_{i=1}^{n-1} u_i (u_{i+1} - u_i) \right) + \sum_{i=1}^{n-1} q(x_i) u_i^2. \end{aligned}$$

Tomando en cuenta que  $u_0 = 0$  u  $u_n = 0$ , se sigue que

$$\begin{aligned} \vec{u}^T A_h \vec{u} &= \frac{1}{h^2} \left( \sum_{i=0}^{n-1} u_{i+1} (u_{i+1} - u_i) - \sum_{i=0}^{n-1} u_i (u_{i+1} - u_i) \right) + \sum_{i=1}^{n-1} q(x_i) u_i^2 \\ &= \frac{1}{h^2} \sum_{i=0}^{n-1} (u_{i+1} - u_i)^2 + \sum_{i=1}^{n-1} q(x_i) u_i^2 \\ &= \frac{1}{h} \sum_{i=0}^{n-1} h \left( \frac{u_{i+1} - u_i}{h} \right)^2 + \sum_{i=1}^{n-1} q(x_i) u_i^2 \\ &\geq \frac{1}{h} \|\vec{u}\|_{1,2}^2. \end{aligned}$$

Luego  $\vec{u}^T A_h \vec{u} \geq \frac{1}{h} \|\vec{u}\|_{1,2}^2 \quad \forall \vec{u} \in \mathbb{R}^n \text{ con } \vec{u} \neq 0$

Si  $\ker(A_h) \neq \{0\}$ , existe  $\tilde{u} \in \mathbb{R}^{n-1}$  con  $u_0 = u_n = 0$  tal que  $\tilde{u} \neq 0$  y  $A_h \tilde{u} = 0$ . Resulta que

$$0 \geq \frac{1}{h} \|\tilde{u}\|_{1,2}^2 \Rightarrow \tilde{u} = 0$$

en otra contradicción con lo supuesto. En consecuencia  $\ker(A_h) = \{0\}$ , que muestra que  $A_h$  es invertible. Por lo tanto,  $\forall \vec{b} \in \mathbb{R}^{n-1}$ , el sistema de ecuaciones  $A_h \vec{u} = \vec{b}$  tiene solución única. ■

Observación: La estabilidad del esquema numérico (2) significa que pequeños errores en los datos de entrada producen pequeños errores en los datos de salida, esto a su vez que existe una constante  $c > 0$  tal que  $\|A_h^{-1}\| \leq c \quad \forall h > 0$ . Note que  $\lim_{h \rightarrow 0} \|A_h^{-1}\| \leq c$ .

La consistencia y la estabilidad son dos nociones independientes.

**Teorema 5** Para todo  $\vec{u} \in V_0$ ,  $\|A_h^{-1} \vec{u}\|_{1,2} \leq L \|\vec{u}\|_2$ ; esto es, se tiene estabilidad con respecto de las normas  $\|\cdot\|_{1,2}$  y  $\|\cdot\|_2$ .

**Demostración.** Probaremos el teorema en dos etapas.

i. Probemos que para todo  $\vec{u} \in V_0$ ,

$$h \vec{u}^T B_h \vec{u} = \|\vec{u}\|_{1,2}^2,$$

donde  $B_h = (b_{ij}(h))$  denota la matriz de  $n-1 \times n-1$  definida por

$$\begin{aligned} b_{ii}(h) &= \frac{2}{h^2} \quad i = 1, \dots, n-1, \\ b_{i,i-1}(h) &= -\frac{1}{h^2} = b_{i-1,i}(h) \quad i = 2, \dots, n-1. \end{aligned}$$

Procediendo de manera similar al teorema precedente, tenemos

$$h \vec{u}^T B_h \vec{u} = \frac{1}{h} \sum_{i=1}^{n-1} u_i (-u_{i-1} + 2u_i - u_{i+1}) = \|\vec{u}\|_{1,2}^2.$$

ii. Probemos que  $\|A_h^{-1} \vec{u}\|_{1,2} \leq c \|\vec{u}\|_{1,2}$ .

Sea  $\vec{v} = A_h^{-1} \vec{u}$  entonces  $\vec{u} = A_h \vec{v}$ . Ponemos  $A_h = B_h + Q_h$  con  $B_h$  definida en i) precedente y  $Q_h = \text{diag}(q(x_1), \dots, q(x_{n-1}))$ .

Multiplicando por  $\vec{v}^T$ , se tiene

$$\vec{v}^T \vec{u} = \vec{v}^T A_h \vec{v} = \vec{v}^T (B_h + Q_h) \vec{v} = \vec{v}^T B_h \vec{v} + \vec{v}^T Q_h \vec{v}.$$

Luego

$$h \vec{v}^T \vec{u} = h \vec{v}^T B_h \vec{v} + h \vec{v}^T Q_h \vec{v} = \|\vec{v}\|_{1,2}^2 + \sum_{i=1}^{n-1} h q(x_i) v_i^2 \geq 0$$

de donde

$$0 \leq \|\vec{v}\|_{1,2}^2 + \sum_{i=1}^{n-1} h q(x_i) v_i^2 = h \vec{v}^T \vec{u} \leq \|\vec{v}\|_2 \|\vec{u}\|_2.$$

Además,

$$\|\vec{v}\|_{1,2}^2 \leq \|\vec{v}\|_{1,2}^2 + \sum_{i=1}^{n-1} h q(x_i) v_i^2 \leq \|\vec{v}\|_2 \|\vec{u}\|_2.$$

Puesto que  $\|\vec{v}\|_2 \leq L \|\vec{v}\|_{1,2}$ , entonces

$$\|\vec{v}\|_{1,2}^2 \leq L \|\vec{v}\|_{1,2} \|\vec{u}\|_2,$$



de donde

$$\|\vec{v}\|_{1,2} \leq L \|\vec{u}\|_2,$$

o bien

$$\|A_h^{-1}\vec{u}\|_{1,2} \leq L \|\vec{u}\|_2.$$

Estabilidad + consistencia  $\Rightarrow$  convergencia. ■

**Teorema 6** *El esquema numérico (2) es convergente para las normas  $\|\cdot\|_2$ ,  $\|\cdot\|_\infty$  y  $\|\cdot\|_{1,2}$ .*

**Demostración.** Puesto que  $U_h = (u(x_1), \dots, u(x_{n-1}))$ ,  $u(0) = u(L) = 0$ , y

$$\vec{r}_h = A_h \vec{U}_h - \vec{b}$$

con lo cual

$$A_h \vec{U}_h = \vec{r}_h + \vec{b}.$$

Además,

$$A_h \vec{u}_h = \vec{b}.$$

Luego, el error  $\vec{e}_h = \vec{U}_h - \vec{u}_h$  satisface la ecuación

$$A_h \vec{e}_h = \vec{r}_h,$$

pues

$$A_h (\vec{U}_h - \vec{u}_h) = \vec{r}_h.$$

Se tiene  $\vec{e}_h \in V_0$  y  $\vec{e}_h = A_h^{-1} \vec{r}_h$ .

Por el teorema precedente (estabilidad para las normas  $\|\cdot\|_{1,2}$  y  $\|\cdot\|_2$ ) y por el teorema relativo a la consistencia, se tiene

$$\|\vec{e}_h\|_{1,2} = \|A_h^{-1} \vec{r}_h\|_{1,2} \leq L \|\vec{r}_h\|_2.$$

Puesto que

$$L^{-\frac{1}{2}} \|\vec{r}_h\|_2 \leq \|\vec{r}_h\|_\infty \leq L^{\frac{1}{2}} \|\vec{r}_h\|_{1,2},$$

resulta que

$$\|\vec{r}_h\|_2 \leq L^{\frac{1}{2}} \|\vec{r}_h\|_\infty \leq \frac{L^{\frac{1}{2}}}{12} h^2 \max_{x \in (0,L)} |u^{iv}(x)|$$

y en consecuencia

$$\|\vec{e}_h\|_{1,2} \leq L \|\vec{r}_h\|_2 \leq \frac{L^{\frac{3}{2}}}{12} h^2 \max_{x \in [0,L]} |u^{iv}(x)|. \quad (*)$$

Por lo tanto,

$$\|\vec{e}_h\|_{1,2} \rightarrow_{h \rightarrow 0} 0.$$

Por otro lado,

$$L^{-\frac{1}{2}} \|\vec{e}_h\|_\infty \leq \|\vec{e}_h\|_{1,2} \leq \frac{L^{\frac{3}{2}}}{12} h^2 \max_{x \in [0,L]} |u^{iv}(x)|, \quad (**)$$

$$\|\vec{e}_h\|_\infty \leq \frac{L^2}{12} h^2 \max_{x \in [0,L]} |u^{iv}(x)| \rightarrow_{h \rightarrow 0} 0. \quad (11.1)$$

Finalmente,

$$L^{-1} \|\vec{e}_h\|_2 \leq \|\vec{e}_h\|_{1,2} \leq \frac{L^{\frac{3}{2}}}{12} h^2 \max_{x \in [0,L]} |u^{iv}(x)|,$$

$$\|\vec{e}_h\|_2 \leq \frac{L^{\frac{5}{2}}}{12} h^2 \max_{x \in [0,L]} |u^{iv}(x)| \rightarrow_{h \rightarrow 0} 0. \quad (***)$$

■

Observe que

$$L^{-\frac{1}{2}} \|\vec{e}_h\|_2 \leq \|\vec{e}_h\|_\infty \leq L^{-\frac{1}{2}} \|\vec{e}_h\|_{1,2} \leq ch^2,$$

con  $c = \frac{c'}{12} \max_{x \in [0, L]} |u^{iv}(x)|$  y  $c' = L^{\frac{3}{2}}, L^2, L^{\frac{5}{2}}$  de acuerdo a (\*), (\*\*) y (\*\*\*) respectivamente.

Adicionalmente, el teorema muestra que el esquema numérico (2) tiene un orden de convergencia 2 para las normas  $\|\cdot\|_\infty$ ,  $\|\cdot\|_2$  y  $\|\cdot\|_{1,2}$ .

### 11.4.3. Orden de convergencia

#### 1. Ecuación diferencial con condiciones de frontera de Dirichlet no homogéneas.

Consideramos el problema siguiente:

$$\text{hallar } u \in C^2([0, L]) \text{ solución de } \begin{cases} -u'' + qu = f \text{ sobre } ]0, L[, \\ u(0) = a, \quad u(L) = b. \end{cases} \quad (\text{P.})$$

Sean  $n \in \mathbb{Z}^+$ ,  $h = \frac{L}{n}$ ,  $\tau(h) = \{x_j = jh \mid j = 0, 1, \dots, n\}$  la malla de  $[0, L]$ .

El esquema numérico del problema (P) es el siguiente:

$$\begin{cases} u_0 = a, \quad u_n = b, \\ -\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + q(x_j)u_j = f(x_j) \quad j = 1, \dots, n-1. \end{cases} \quad (\text{P}_1.)$$

El esquema numérico (P<sub>1</sub>) en forma explícita se escribe: para  $j = 1$ ,  $u_0 = a$ , y

$$-\frac{u_2 - 2u_1 + a}{h^2} + q(x_1)u_1 = f(x_1)$$

con lo cual

$$\left[ \frac{2}{h^2} + q(x_1) \right] u_1 - \frac{1}{h^2} u_2 = f(x_1) + \frac{a}{h^2}.$$

Para  $j = n-1$ ,  $u_n = b$ , luego

$$-\frac{b - 2u_{n-1} + u_{n-2}}{h^2} + q(x_{n-1})u_{n-1} = f(x_{n-1})$$

de donde

$$-\frac{1}{h^2} u_{n-2} + \left[ \frac{2}{h^2} + q(x_{n-1}) \right] u_{n-1} = f(x_{n-1}) + \frac{b}{h^2}.$$

El esquema numérico (P<sub>1</sub>) es:

$$\begin{cases} \left[ \frac{2}{h^2} + q(x_1) \right] u_1 - \frac{1}{h^2} u_2 = f(x_1) + \frac{a}{h^2} \\ -\frac{1}{h^2} u_{j-1} + \left[ \frac{2}{h^2} + q(x_j) \right] u_j - \frac{1}{h^2} u_{j+1} = f(x_j) \quad j = 2, \dots, n-2 \\ -\frac{1}{h^2} u_{n-2} + \left[ \frac{2}{h^2} + q(x_{n-1}) \right] u_{n-1} = f(x_{n-1}) + \frac{b}{h^2}. \end{cases} \quad (\text{P}_2.)$$

Sea  $w : [0, L] \Rightarrow \mathbb{R}$  la función definida por

$$w(x) = \frac{x}{L}b + \left(1 - \frac{x}{L}\right)a.$$

Entonces  $w(0) = a$ ,  $w(L) = b$ . Se define  $\tilde{u} = u - w$ . Se tiene

$$\begin{cases} \tilde{u}(0) = u(0) - w(0) = 0, \\ \tilde{u}(L) = u(L) - w(L) = 0. \end{cases}$$

La función  $\tilde{u}$  satisface las condiciones de frontera de Dirichlet homogénea. La ecuación lineal del problema (P) se escribe

$$-\tilde{u}'' + q(\tilde{u} + w) = f$$

de donde

$$-\tilde{u}'' + q\tilde{u} = f - qw.$$

En consecuencia

$$\begin{cases} -\tilde{u}'' + q\tilde{u} = g & \text{sobre } ]0, L[, \\ \tilde{u}(0) = \tilde{u}(L) = 0 \end{cases} \quad (\tilde{\text{P}}.)$$

con  $g = f - qw$ .

Note que  $g(x_j) = f(x_j) - (qw)(x_j) = f(x_j) - q(x_j) \left[ \frac{x_j}{L}b + \left(1 - \frac{x_j}{L}\right)a \right]$ .

Desde el punto de vista informático, si se utiliza  $(\tilde{\text{P}})$ , el esquema numérico es idéntico al establecido en la sección 3, el sistema de ecuaciones  $A_h \vec{u}_h = \vec{b}$  es similar al establecido en esa sección a condición de remplazar  $\vec{b} = (f(x_1), \dots, f(x_n))^T$  por el vector  $\vec{b} = (g(x_1), \dots, g(x_n))^T$ .

Si se utiliza el esquema numérico  $(\text{P}_2)$ , la matriz  $A_h$  es la misma que la de la sección 3, el vector  $\vec{b}$  se remplaza por el vector  $(f(x_1) + \frac{a}{h^2}, f(x_2), \dots, f(x_{n-2}), f(x_1) + \frac{b}{h^2})$ .

La estabilidad, consistencia y convergencia para el problema  $(\tilde{\text{P}})$  ha sido discutida en la sección 5. Por lo tanto, el esquema numérico  $(\text{P}_2)$  es convergente, con orden de convergencia igual a 2.

## 2. Ecuación diferencial con condiciones de frontera de Neumann.

Consideramos el problema siguiente:

$$\text{hallar } u \in C^2([0, L]) \text{ solución de } \begin{cases} -u'' + qu = f & \text{sobre } ]0, L[, \\ u'(0) = a, \quad u'(L) = b. \end{cases} \quad (\text{P}_n.)$$

Sea  $w$  la función real definida por:

$$w(x) = \frac{x^2}{2L}b - \frac{(L-x)^2}{2L}a, \quad x \in [0, L].$$

Entonces

$$\begin{aligned} w'(x) &= \frac{x}{L}b + \left(1 - \frac{x}{L}\right)a, \\ w''(x) &= \frac{b}{L} - \frac{a}{L} = \frac{b-a}{L}, \\ w'(0) &= a, \quad w'(L) = b. \end{aligned}$$

Se define  $\tilde{u} = u - w$ . Entonces

$$\begin{aligned} \tilde{u}'(x) &= u'(x) - w'(x) \quad x \in [0, L], \\ \tilde{u}'(0) &= 0, \quad \tilde{u}'(L) = 0, \\ \tilde{u}''(x) &= u''(x) - w''(x) = u''(x) - \frac{b-a}{L}. \end{aligned}$$

Consecuentemente, el problema  $(\text{P}_n)$  se escribe

$$-\left(\tilde{u}''(x) + \frac{b-a}{L}\right) + q(x)(\tilde{u}(x) + w(x)) = f(x) \quad x \in ]0, L[,$$

con lo cual

$$\begin{cases} -\tilde{u}'' + q\tilde{u} = f - \frac{b-a}{L} - qw & \text{sobre } ]0, L[, \\ \tilde{u}'(0) = \tilde{u}'(L) = 0. \end{cases} \quad (\tilde{\text{P}}_n)$$

El resultado precedente muestra que debemos estudiar la ecuación diferencial con las condiciones de Neumann homogéneas; eso es, consideramos el problema  $(\text{P}_0)$  siguiente:

$$\begin{cases} -u'' + qu = f & \text{sobre } ]0, L[, \\ u'(0) = u'(L) = 0. \end{cases} \quad (\text{P}_0)$$

Para construir un esquema numérico que aproxime la solución de  $(P_0)$ , comenzamos con la aproximación de  $u'(0)$  y  $u'(L)$ .

### Aproximación de $u'(0)$ y $u''(0)$

Sea  $n \in \mathbb{Z}^+$ ,  $h = \frac{L}{n}$  y  $\tau(n) = \{x_j = jh \mid j = 0, 1, \dots, n\}$ . Entonces

$$u''(0) \simeq \frac{u'(\tilde{x}_1) - u'(0)}{\frac{h}{2}} = \frac{2}{h} (u'(\tilde{x}_1) - a),$$

donde  $\tilde{x}_1$  es el punto medio del intervalo  $[0, x_1]$ . Para la aproximación de  $u'(\tilde{x}_1)$  utilizamos las diferencias finitas centrales. Tenemos

$$u'(\tilde{x}_1) \simeq \frac{u(x_1) - u(0)}{h},$$

luego

$$u''(\tilde{x}_1) \simeq \frac{2}{h} \left[ \frac{u(x_1) - u(0)}{h} - a \right] = \frac{2}{h^2} [u(x_1) - u(0) - ah].$$

Utilizando el polinomio de Taylor, tenemos

$$u(x_1) = u(0) + hu'(0) + \frac{h^2}{2!} u''(0) + \frac{h^3}{3!} u'''(\xi) \quad \text{con } \xi \in [0, x_1],$$

con lo cual

$$\frac{2}{h^2} [u(x_1) - u(0) - ah] = u''(0) + \frac{h}{3} u'''(\xi).$$

### Aproximación de $u'(L)$ y $u''(L)$

Sea  $\tilde{x}_n$  el punto medio del intervalo  $[x_{n-1}, L]$ . Procediendo de modo similar al caso  $u'(L)$ , se deduce que

$$\begin{aligned} u''(L) &\simeq \frac{u'(L) - u'(\tilde{x}_{n-1})}{\frac{h}{2}} = \frac{2}{h} [b - u'(\tilde{x}_{n-1})] \\ &\simeq \frac{2}{h} \left[ b - \frac{u(L) - u(x_{n-1})}{h} \right] = \frac{2}{h^2} (-u(L) + u(x_{n-1}) + bh). \end{aligned}$$

Por el desarrollo de Taylor, se tiene

$$u(x_{n-1}) = u(L) - hu'(L) + \frac{h^2}{2!} u''(L) - \frac{h^3}{3!} u'''(\xi_n) \quad \text{con } \xi_n \in [x_{n-1}, L],$$

entonces

$$\frac{2}{h^2} (-u(L) + u(x_{n-1}) + bh) = u''(L) - \frac{h}{3} u'''(\xi_n).$$

Definimos  $\mathcal{L}u = -u'' + qu$  y suponemos  $u \in C^4([0, L])$ . Entonces

$$\begin{aligned} \mathcal{L}u(x) &= -u''(x) + q(x)u(x) \quad x \in [0, L], \\ \mathcal{L}_\pi u(0) &= -\frac{2}{h^2} (u(x_1) - u(0) - ah) + q(0)u(0), \\ \mathcal{L}_\pi u(L) &= -\frac{2}{h^2} (-u(L) + u(x_{n-1}) + bh) + q(L)u(L). \end{aligned}$$

Se tiene

$$\begin{aligned} |\mathcal{L}u(0) - \mathcal{L}_\pi u(0)| &= \left| -u''(0) + q(0)u(0) + \frac{2}{h^2} (u(x_1) - u(0) - ah) - q(0)u(0) \right| \\ &= \frac{h}{3} |u'''(\xi_1)| \quad \text{con } \xi_1 \in [0, x_1], \\ |\mathcal{L}u(L) - \mathcal{L}_\pi u(L)| &= \left| -\frac{h}{3} u'''(\xi_n) \right| = \frac{h}{3} |u'''(\xi_n)| \quad \text{con } \xi_n \in [x_{n-1}, L]. \end{aligned}$$

Para  $j = 0, \dots, n$ , el esquema numérico tiene la forma siguiente:

$$-\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + q(x_j) u_j = f(x_j)$$

Luego

$$\begin{cases} \left( \frac{2}{h^2} + q(0) \right) u(0) - \frac{2}{h^2} u_1 = f(0) - \frac{2a}{h}, \\ -\frac{1}{h^2} u_{j-1} + \left( \frac{2}{h^2} + q(x_j) \right) u_j - \frac{1}{h} u_{j+1} = f(x_j), \\ -\frac{2}{h^2} u_{n-1} + \left( \frac{2}{h^2} + q(L) \right) u_n = f(L) + \frac{2b}{h}. \end{cases} \quad (\text{P}_1)$$

Poniendo  $\mathcal{L}_\pi u(x_j) = -\frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1}))}{h^2} + q(x_j) u(x_j)$ , se tiene

$$|\mathcal{L}u(x_j) - \mathcal{L}_\pi u(x_j)| \leq \frac{h^2}{12} \max_{x \in [0, L]} |u^{(iv)}(x)| \quad j = 1, \dots, n-1.$$

Como

$$\begin{aligned} |\mathcal{L}u(0) - \mathcal{L}_\pi u(0)| &\leq \frac{h}{3} \max_{x \in [0, L]} |u'''(x)|, \\ |\mathcal{L}u(L) - \mathcal{L}_\pi u(L)| &\leq \frac{h}{3} \max_{x \in [0, L]} |u'''(x)|, \end{aligned}$$

se sigue

$$\|\vec{r}_h\| \leq \frac{h}{3} \max \left\{ \|u'''\|_{L^\infty(0, L)}, \|u^{(iv)}\|_{L^\infty(0, L)} \right\},$$

que muestra que el esquema numérico es de orden 1.

Se puede observar que la matriz  $A_h = (a_{ij}(h))$  del esquema numérico (P<sub>1</sub>) está definida por

$$\begin{aligned} a_{ii}(h) &= -\frac{2}{h^2} + q(x_i), \quad i = 0, 1, \dots, n, \\ a_{12}(h) &= -\frac{2}{h^2}, \quad a_{nn-1}(h) = -\frac{2}{h^2}, \\ a_{i-1i}(h) &= a_{ii-1}(h) = -\frac{1}{h^2} \quad i = 2, \dots, n-1. \end{aligned}$$

La matriz  $A_h$  no es simétrica. Además,

$$\begin{cases} a_{ii}(h) \geq -a_{i-1i}(h) - a_{ii-1}(h), \quad i = 1, \dots, n-1, \\ a_{11}(h) \geq -a_{12}(h), \\ a_{nn}(h) \geq -a_{nn-1}(h). \end{cases}$$

Si  $q \neq 0$ , existe  $j = 0, 1, \dots, n$  tal que  $q(x_j) > 0$  con lo cual

$$a_{jj}(h) > -a_{j-1j}(h) - a_{jj-1}(h)$$

que prueba que la matriz  $A_h$  es diagonalmente dominante, y en consecuencia  $A_h$  es invertible.

Conclusión: El esquema numérico (P<sub>1</sub>) es convergente con orden de convergencia igual a 1.

#### 11.4.4. Método de diferencias finitas en mallas no uniformes

##### Posición del problema

Sea  $L > 0$ . Consideramos el problema (P) siguiente:

$$\text{hallar } u : [0, L] \rightarrow \mathbb{R} \text{ solución de } \begin{cases} -\frac{d}{dx} \left( k(x) \frac{du}{dx}(x) \right) + q(x) u(x) = f(x) & x \in ]0, L[, \\ u(0) = 0, \quad u(L) = 0, \end{cases} \quad (\text{P})$$

donde  $k, q, f \in C^0([0, L])$  tales que

$$\begin{aligned} k(x) &\geq \alpha > 0 \quad \forall x \in [0, L], \\ q(x) &\geq 0 \quad \forall x \in [0, L]. \end{aligned}$$

El término  $-k \frac{du}{dx}$  se interpreta como el flujo.

Sean  $n \in \mathbb{Z}^+$  y  $\tau(n) = \{x_0 = 0 < x_1 < \dots < x_{n-1} < x_n = L\}$  la malla de  $[0, L]$  no necesariamente uniforme. Ponemos  $h_j = x_j - x_{j-1}$   $j = 1, \dots, n$ ; y  $h = \max_{j=1, \dots, n} h_j$ .

Deseamos aproximar la solución de  $u$  de (P) en los nodos  $x_1, \dots, x_{n-1}$ .

### Discretizando de (P)

Sean  $a, b \in [0, L]$  tales que  $a < b$ . Entonces

$$\begin{aligned} - \int_a^b \frac{d}{dx} \left( k(x) \frac{du}{dx} \right) dx + \int_a^b q(x) u(x) dx &= \int_a^b f(x) dx \\ -k(x) \frac{du}{dx} \Big|_a^b + \int_a^b q(x) u(x) dx &= \int_a^b f(x) dx. \end{aligned}$$

Para  $a = x_i - \frac{h_i}{2}$ ,  $b = x_i + \frac{h_{i+1}}{2}$  ( $a$  y  $b$  son los puntos medios de los intervalos  $[x_{i-1}, x_i]$  y  $[x_i, x_{i+1}]$  respectivamente) se tiene

$$\begin{aligned} & - \left[ k \left( x_i + \frac{h_{i+1}}{2} \right) \frac{du}{dx} \left( x_i + \frac{h_{i+1}}{2} \right) - k \left( x_i - \frac{h_i}{2} \right) \frac{du}{dx} \left( x_i - \frac{h_i}{2} \right) \right] + \int_{x_i - \frac{h_i}{2}}^{x_i + \frac{h_{i+1}}{2}} q(x) u(x) dx \\ &= \int_{x_i - \frac{h_i}{2}}^{x_i + \frac{h_{i+1}}{2}} f(x) dx \end{aligned}$$

Las derivadas  $\frac{du}{dx} \left( x_i + \frac{h_{i+1}}{2} \right)$  y  $\frac{du}{dx} \left( x_i - \frac{h_i}{2} \right)$  se aproximan mediante diferencias finitas centrales

$$\begin{aligned} \frac{du}{dx} \left( x_i - \frac{h_i}{2} \right) &\simeq \frac{u(x_i) - u(x_i - h_i)}{h_i} = \frac{u(x_i) - u(x_{i-1}))}{h_i}, \\ \frac{du}{dx} \left( x_i + \frac{h_{i+1}}{2} \right) &\simeq \frac{u(x_i + h_{i+1}) - u(x_i)}{h_{i+1}} = \frac{u(x_{i+1}) - u(x_i)}{h_{i+1}}. \end{aligned}$$

Las integrales del modo siguiente

$$\int_{x_i - \frac{h_i}{2}}^{x_i + \frac{h_{i+1}}{2}} q(x) u(x) dx \simeq \frac{h_i + h_{i+1}}{2} q(x_i) u(x_i), \quad i = 1, \dots, n-1, \quad (1)$$

$$\int_{x_i - \frac{h_i}{2}}^{x_i + \frac{h_{i+1}}{2}} f(x) dx \simeq \frac{h_i + h_{i+1}}{2} f(x_i) \quad i = 1, \dots, n-1. \quad (2)$$

Las fórmulas (1) y (2) son una variante de la fórmula del punto medio siguiente: si  $g \in C^0([\alpha, \beta])$ ,

$$\int_{\alpha}^{\beta} g(x) dx \simeq (\beta - \alpha) g\left(\frac{\alpha + \beta}{2}\right).$$

Denotamos con  $u_i$  una aproximación de  $u(x_i)$ . Se establece el esquema numérico siguiente:

$$\begin{aligned} -k \left( x_i + \frac{h_{i+1}}{2} \right) \frac{u_{i+1} - u_i}{h_{i+1}} + k \left( x_i - \frac{h_i}{2} \right) \frac{u_i - u_{i-1}}{h_i} + \frac{h_i + h_{i+1}}{2} q(x_i) u_i &= \\ \frac{h_i + h_{i+1}}{2} f(x_i) & \quad i = 1, \dots, n-1. \end{aligned}$$

de donde

$$\begin{aligned}
 & -\frac{k\left(x_i - \frac{h_i}{2}\right)}{h_i}u_{i-1} + \left[ \frac{k\left(x_i - \frac{h_i}{2}\right)}{h_i} + \frac{k\left(x_i + \frac{h_{i+1}}{2}\right)}{h_{i+1}} + \frac{h_i + h_{i+1}}{2}q(x_i) \right] u_i - \frac{k\left(x_i + \frac{h_{i+1}}{2}\right)}{h_{i+1}}u_{i+1} \\
 & = \frac{h_i + h_{i+1}}{2}f(x_i) \quad i = 1, \dots, n-1.
 \end{aligned}$$

Tomando en consideración las condiciones de frontera  $u(0) = u_0 = 0$  y  $u(L) = u_n = 0$ , resulta:

$$\left\{ \begin{aligned} & \left[ \frac{k\left(x_1 - \frac{h_1}{2}\right)}{h_1} + \frac{k\left(x_1 + \frac{h_2}{2}\right)}{h_2} + \frac{1}{2}(h_1 + h_2)q(x_1) \right] u_1 - \frac{k\left(x_1 + \frac{h_2}{2}\right)}{h_2}u_2 = \frac{1}{2}(h_1 + h_2)f(x_1), \\ & -\frac{k\left(x_i - \frac{h_i}{2}\right)}{h_i}u_{i-1} + \left[ \frac{k\left(x_i - \frac{h_i}{2}\right)}{h_i} + \frac{k\left(x_i + \frac{h_{i+1}}{2}\right)}{h_{i+1}} + \frac{h_i + h_{i+1}}{2}q(x_i) \right] u_i - \frac{k\left(x_i + \frac{h_{i+1}}{2}\right)}{h_{i+1}}u_{i+1} = \frac{1}{2}(h_i + h_{i+1})f(x_i), \\ & -\frac{k\left(x_{n-1} - \frac{h_{n-1}}{2}\right)}{h_{n-1}}u_{n-1} + \left[ \frac{k\left(x_{n-1} - \frac{h_{n-1}}{2}\right)}{h_{n-1}} + \frac{k\left(x_{n-1} + \frac{h_n}{2}\right)}{h_n} + \frac{h_{n-1} + h_n}{2}q(x_{n-1}) \right] u_{n-1} = \frac{1}{2}(h_{n-1} + h_n)f(x_n) \end{aligned} \right. \quad (3)$$

Ponemos

$$\begin{aligned}
 a_i &= -\frac{k\left(x_i - \frac{h_i}{2}\right)}{h_i} \quad i = 2, \dots, n-1, \quad a_1 = 0, \\
 b_i &= \frac{k\left(x_i - \frac{h_i}{2}\right)}{h_i} + \frac{k\left(x_i + \frac{h_{i+1}}{2}\right)}{h_{i+1}} \frac{1}{2}(h_i + h_{i+1})q(x_i) \quad i = 1, \dots, n-1, \\
 c_i &= -\frac{k\left(x_i + \frac{h_{i+1}}{2}\right)}{h_{i+1}}, \quad i = 1, \dots, n-2, \quad c_{n-1} = 0, \\
 f_i &= \frac{1}{2}(h_i + h_{i+1})f(x_i) \quad i = 1, \dots, n-1
 \end{aligned}$$

El esquema numérico (3) se escribe

$$a_i u_{i-1} + b_i u_i + c_i u_{i+1} = f_i \quad i = 1, \dots, n-1. \quad (4)$$

Por otro lado, se pone  $\vec{u} = (u_1, \dots, u_{n-1})^T$ , y se define  $A_h = (a_{ij}(h))$  con

$$\begin{aligned}
 a_{ii}(h) &= \frac{k\left(x_i - \frac{h_i}{2}\right)}{h_i} + \frac{k\left(x_i + \frac{h_{i+1}}{2}\right)}{h_{i+1}} + \frac{h_i + h_{i+1}}{2}q(x_i) \quad i = 1, \dots, n-1 \\
 a_{i-i}(h) &= a_{ii-1}(h) = -\frac{k\left(x_i - \frac{h_i}{2}\right)}{h_i} \quad i = 2, \dots, n-1, \\
 \vec{v} &= (b_1, \dots, b_{n-1})^T \quad \text{con } b_i = \frac{h_i + h_{i+1}}{2}f(x_i).
 \end{aligned}$$

El esquema numérico (3) se escribe en la forma

$$A_h \vec{u} = \vec{b}. \quad (5)$$

Se verifican las condiciones siguientes:

- i.  $A_h = A_h^T$ , es decir que  $A_h$  es simétrica.
- ii.  $b_i > 0 \quad i = 1, \dots, n-1$ .
- iii.  $c_1 < 0, \quad a_0 < 0, \quad c_i < 0 \quad i = 2, \dots, n-2, \quad a_{n-1} < 0$ .

$$\text{iv. } \begin{cases} b_i > -a_i - c_i & i = 2, \dots, n-2 \\ b_1 > -c_1, & b_{n-1} > -a_{nn-1}. \end{cases}$$

Se demuestra que  $A_h$  es positiva, esto es,  $A_h^{-1} > 0$ .

### Observación

Si se considera el problema (P) siguiente:

$$\begin{cases} -\frac{d}{dx} \left( k(x) \frac{du(x)}{dx} \right) + q(x) u(x) = f(x) & x \in ]0, L[, \\ u(0) = a, & u(L) = b \end{cases} \quad (\text{P})$$

donde  $a, b \in \mathbb{R}$ ,  $k, q, f$  funciones que satisfacen las condiciones citadas precedentemente.

Se define  $w(x) = \frac{x}{L}b + \frac{(L-x)}{L}a$   $x \in [0, L]$ . Se tiene

$$w(0) = a, w(L) = b, w'(x) = \frac{b-a}{L} \quad \forall x \in [0, L].$$

Se define  $\tilde{u} = u - w$ . Entonces

$$\begin{cases} -\frac{d}{dx} \left( k \frac{d\tilde{u}}{dx} \right) + q\tilde{u} = f - qw + \frac{b-a}{L} \frac{dk}{dx}. \\ \tilde{u}(0) = 0, & \tilde{u}(L) = 0 \end{cases} \quad (\tilde{\text{P}})$$

Note que en este caso  $k \in C^1([0, L])$ .

El problema  $(\tilde{\text{P}})$  tiene condiciones de frontera de Dirichlet homogéneas.

### Condiciones de frontera mixtas

Consideramos ahora el caso en el que las condiciones de frontera son las siguientes:

$$\begin{cases} \alpha_1 u'(0) + \alpha_2 u(0) = g_1 \\ \beta_1 u'(L) + \beta_2 u(L) = g_2 \end{cases}$$

con  $\alpha_1, \alpha_2, \beta_1, \beta_2, g_1, g_2 \in \mathbb{R}$ .

### Condiciones de frontera en $x = 0$ .

Tomando  $a = 0, b = \frac{h_1}{2}$  se tiene

$$\begin{aligned} -\int_0^{\frac{h_1}{2}} \frac{d}{dx} \left( k(x) \frac{du}{dx} \right) dx + \int_0^{\frac{h_1}{2}} q(x) u(x) dx &= \int_0^{\frac{h_1}{2}} f(x) dx \\ -k(x) \frac{du}{dx} \Big|_0^{\frac{h_1}{2}} + \int_0^{\frac{h_1}{2}} q(x) u(x) dx &= \int_0^{\frac{h_1}{2}} f(x) dx. \end{aligned}$$

Consideramos el término  $-k(x) \frac{du}{dx} \Big|_0^{\frac{h_1}{2}}$ . Tenemos

$$-k(x) \frac{du}{dx} \Big|_0^{\frac{h_1}{2}} = -k\left(\frac{h_1}{2}\right) \frac{du}{dx}\left(\frac{h_1}{2}\right) + k(0) \frac{du}{dx}(0).$$

Puesto que  $\alpha_1 u'(0) + \alpha_2 u(0) = g_1$ , suponemos que  $\alpha_1 \neq 0$ , con lo cual  $u'(0) = \frac{1}{\alpha_1} (g_1 - \alpha_2 u(0))$ .

Por otro lado, la derivada  $\frac{du}{dx}\left(\frac{h_1}{2}\right)$  se aproxima mediante diferencias finitas centrales. Se tiene

$$\frac{du}{dx}\left(\frac{h_1}{2}\right) \simeq \frac{u(x_1) - u(0)}{h_1},$$



en consecuencia  $-k(x) \frac{du}{dx} \Big|_0^{\frac{h_1}{2}}$  se aproxima como

$$-k(x) \frac{du}{dx} \Big|_0^{\frac{h_1}{2}} \simeq -k\left(\frac{h_1}{2}\right) \frac{u(x_1) - u(0)}{h_1} + k(0) \frac{1}{\alpha_1} (g_1 - \alpha_2 u(0)).$$

Además,

$$\begin{aligned} \int_0^{\frac{h_1}{2}} q(x) u(x) dx &\simeq \frac{h_1}{2} q(0) u(0), \\ \int_0^{\frac{h_1}{2}} f(x) dx &\simeq \frac{h_1}{2} f(0). \end{aligned}$$

Resulta

$$-k\left(\frac{h_1}{2}\right) \frac{u_1 - u_0}{h_1} + k(0) \frac{1}{\alpha_1} (g_1 - \alpha_2 u_0) + \frac{h_1}{2} q(0) u_0 = \frac{h_1}{2} f(0),$$

o bien

$$\left[ \frac{k\left(\frac{h_1}{2}\right)}{h_1} - \frac{\alpha_2}{\alpha_1} k(0) + \frac{h_1}{2} q(0) \right] u_0 - \frac{k\left(\frac{h_1}{2}\right)}{h_1} u_1 = \frac{h_1}{2} f(0).$$

En la práctica  $\alpha_1 = -1$ ,  $\alpha_2 \geq 0$  con lo cual

$$\left[ \frac{k\left(\frac{h_1}{2}\right)}{h_1} + \alpha_2 k(0) + \frac{h_1}{2} q(0) \right] u_0 - \frac{k\left(\frac{h_1}{2}\right)}{h_1} u_1 = \frac{h_1}{2} f(0) + k(0) g_1.$$

**Condición de frontera en  $x = L$ .**

Tomamos  $a = L - \frac{h_n}{2}$ ,  $b = L$ . Entonces

$$-k(x) \frac{du}{dx} \Big|_{L-\frac{h_n}{2}}^L + \int_{L-\frac{h_n}{2}}^L q(x) u(x) dx = \int_{L-\frac{h_n}{2}}^L f(x) dx,$$

con lo cual

$$-k(x) \frac{du}{dx} \Big|_{L-\frac{h_n}{2}}^L = -k(L) \frac{du}{dx}(L) + k\left(L - \frac{h_n}{2}\right) \frac{du}{dx}\left(L - \frac{h_n}{2}\right).$$

La derivada  $\frac{du}{dx}\left(L - \frac{h_n}{2}\right)$  se aproxima mediante diferencias finitas centrales, esto es,

$$\frac{du}{dx}\left(L - \frac{h_n}{2}\right) \simeq \frac{u(L) - u(x_{n-1})}{h_n}.$$

Por otro lado,  $\beta_1 u'(L) + \beta_2 u(L) = g_2 \Rightarrow u'(L) = \frac{1}{\beta_1} (g_2 - \beta_2 u(L))$  con  $\beta_1 \neq 0$ . Luego,

$$-k(x) \frac{du}{dx} \Big|_{L-\frac{h_n}{2}}^L = -\frac{k(L)}{\beta_1} (g_2 - \beta_2 u(L)) + k\left(L - \frac{h_n}{2}\right) \frac{u(L) - u(x_{n-1})}{h_n}.$$

Por otro lado,

$$\begin{aligned} \int_{L-\frac{h_n}{2}}^L q(x) u(x) dx &\simeq \frac{h_n}{2} q(L) u(L), \\ \int_{L-\frac{h_n}{2}}^L f(x) dx &\simeq \frac{h_n}{2} f(L). \end{aligned}$$

Entonces

$$-\frac{k(L)}{\beta_1} (g_2 - \beta_2 u(L)) + k\left(L - \frac{h_n}{2}\right) \frac{u(L) - u(x_{n-1})}{h_n} + \frac{h_n}{2} q(L) u(L) \simeq \frac{h_n}{2} f(L).$$

de donde

$$\frac{-k(L - \frac{h_n}{2})}{h_n} u_{n-1} + \left[ \frac{k(L - \frac{h_n}{2})}{h_n} + \frac{k(L)\beta_2}{\beta_1} + \frac{h_n}{2} q(L) \right] u_n = \frac{h_n}{2} f(L) + \frac{k(L)g_2}{\beta_1}.$$

En la práctica  $\beta_1 = 1$ . Así

$$\frac{-k(L - \frac{h_n}{2})}{h_n} u_{n-1} + \left[ \frac{k(L - \frac{h_n}{2})}{h_n} + k(L)\beta_2 + \frac{h_n}{2} q(L) \right] u_n = \frac{h_n}{2} f(L) + k(L)g_2.$$

La matriz  $A_h = (a_{ij}(h)) \in M_{(n+1) \times (n+1)}(\mathbb{R})$  satisface las siguientes propiedades:

- i.  $A_h = A_h^T$ .
- ii.  $a_{ii}(h) > 0 \quad i = 0, 1, \dots, n$ ,
- iii.  $a_{i-1i}(h) = a_{ii-1}(h) < 0 \quad i = 2, \dots, n$ ,
- iv.  $\begin{cases} a_{ii}(h) \geq -a_{i-1i}(h) - a_{ii-1}(h) & i = 2, \dots, n-1, \\ a_{11}(h) > -a_{21}(h), \quad a_{nn}(h) > -a_{nn-1}(h). \end{cases}$

La matriz  $A_h$  es positiva.

## 11.5. Ejercicios resueltos

1. Considerar el problema de valores de frontera:

$$\begin{cases} -u''(x) + x^2 u(x) = 1 + x & x \in ]0, 1[, \\ u(0) = 0, \quad u(1) = 1. \end{cases}$$

Aplicar el método de diferencias finitas para aproximar la solución con  $n = 5$ . La matriz debe factorarse con el método de Choleski.

### Solución

Utilizando diferencias finitas centrales, se tiene

$$u''(x_j) \simeq \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1}))}{h_x^2},$$

donde  $h_x = \frac{1}{5} = 0,2$  y  $x_j = jh_x \quad j = 0, 1, \dots, 5$ .

Sea  $u_j$  una aproximación de  $u(x_j)$ . Entonces el problema de valores de frontera:

$$\begin{cases} -u''(x) + x^2 u(x) = 1 + x & x \in ]0, 1[, \\ u(0) = 0, \quad u(1) = 1. \end{cases}$$

se discretiza del modo siguiente:

$$\begin{cases} u_0 = 0, \quad u(1) = u_5 = 1, \\ -\frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1}))}{h_x^2} + x_j^2 u_j = 1 + x_j \end{cases}$$

o de manera explícita:

$$\begin{aligned} j &= 1, & -\frac{u_2 - 2u_1}{h_x^2} + x_1^2 u_1 &= 1 + x_1 \\ j &= 2, & -\frac{u_3 - 2u_2 + u_1}{h_x^2} + x_2^2 u_2 &= 1 + x_2 \\ j &= 3, & -\frac{u_4 - 2u_3 + u_2}{h_x^2} + x_3^2 u_3 &= 1 + x_3 \\ j &= 4, & -\frac{1 - 2u_4 + u_3}{h_x^2} + x_4^2 u_4 &= 1 + x_4 \end{aligned}$$

que expresado en forma matricial, se tiene

$$\begin{bmatrix} \frac{2}{h_x^2} + x_1^2 & -\frac{1}{h_x^2} & 0 & 0 \\ -\frac{1}{h_x^2} & \frac{2}{h_x^2} + x_2^2 & -\frac{1}{h_x^2} & 0 \\ 0 & -\frac{1}{h_x^2} & \frac{2}{h_x^2} + x_3^2 & -\frac{1}{h_x^2} \\ 0 & 0 & -\frac{1}{h_x^2} & \frac{2}{h_x^2} + x_4^2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} 1 + x_1 \\ 1 + x_2 \\ 1 + x_3 \\ 1 + x_4 + \frac{1}{h_x^2} \end{bmatrix}.$$

Remplazando  $h_x = 0,2$  y  $x_j = jh_x$   $j = 1, 2, 3, 4$  se obtiene

$$A = \begin{bmatrix} 50,04 & -25 & 0 & 0 \\ -25 & 50,16 & -25 & 0 \\ 0 & -25 & 50,36 & -25 \\ 0 & 0 & -25 & 50,64 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 1,2 \\ 1,4 \\ 1,6 \\ 26,8 \end{bmatrix}$$

Ponemos  $\vec{u} = (u_1, u_2, u_3, u_4)^T$ . Para resolver el sistema de ecuaciones  $A\vec{u} = \vec{b}$  aplicamos el método de factorización  $LU$ .

$$\text{Sean } L = \begin{bmatrix} L_{11} & 0 & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 \\ 0 & L_{32} & L_{33} & 0 \\ 0 & 0 & L_{43} & L_{44} \end{bmatrix}, \quad U = \begin{bmatrix} 1 & u_{12} & 0 & 0 \\ 0 & 1 & u_{23} & 0 \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}. \text{ Entonces}$$

$$A = LU = \begin{bmatrix} L_{11} & L_{22}u_{12} & 0 & 0 \\ L_{21} & L_{21}u_{12} + L_{22} & L_{22}u_{23} & 0 \\ 0 & L_{32} & L_{32}u_{23} + L_{33} & L_{33}u_{34} \\ 0 & 0 & L_{43} & L_{43}u_{34} + L_{44} \end{bmatrix}$$

$$L_{11} = 50,04, \quad u_{12} = -\frac{25}{50,04} = -0,4996003197$$

$$L_{21} = -25, \quad L_{22} = 50,16 - (-25) \left( -\frac{25}{50,04} \right) = 37,66999201,$$

$$u_{23} = -\frac{25}{37,66999201} = -0,663658224$$

$$L_{32} = -25, \quad L_{33} = 50,36 - (-25) (-0,663658224) = 33,7685444,$$

$$u_{34} = -\frac{25}{33,7685444} = -0,740333954$$

$$L_{43} = -25, \quad L_{44} = 50,64 - (-25) (-0,740333954) = 32,13165115$$

$$L = \begin{bmatrix} 50,04 & 0 & 0 & 0 \\ -25 & 37,66999201 & 0 & 0 \\ 0 & -25 & 33,7685444 & 0 \\ 0 & 0 & -25 & 32,13165115 \end{bmatrix}$$

$$U = \begin{bmatrix} 1 & -0,4996003197 & 0 & 0 \\ 0 & 1 & -0,663658224 & 0 \\ 0 & 0 & 1 & -0,740333954 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

El sistema de ecuaciones  $A\vec{x} = \vec{b}$  se transforma en los siguientes:  $LU\vec{u} = \vec{b} \Leftrightarrow \begin{cases} L\vec{y} = \vec{b} \\ U\vec{u} = \vec{y} \end{cases}.$

Solución del sistema lineal de ecuaciones  $L\vec{y} = \vec{b}$ . En forma explícita, este sistema de ecuaciones lineales se expresa como sigue:

$$\begin{bmatrix} 50,04 & 0 & 0 & 0 \\ -25 & 37,66999201 & 0 & 0 \\ 0 & -25 & 33,7685444 & 0 \\ 0 & 0 & -25 & 32,13165115 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 1,2 \\ 1,4 \\ 1,6 \\ 26,8 \end{bmatrix},$$

luego

$$\begin{aligned} y_1 &= \frac{1,2}{50,04} = 0,023980815, \\ y_2 &= \frac{1,4 + 25 \times 0,023980815}{37,66999201} = 0,05307992589, \\ y_3 &= \frac{1,6 + 25 \times 0,05307992589}{33,7685444} = 0,0866782444, \\ y_4 &= \frac{26,8 + 25 \times 0,0866782444}{32,13165115} = 0,9015084838, \end{aligned}$$

Solución del sistema lineal de ecuaciones  $U\vec{u} = \vec{y}$ . Este sistema en forma explícita se escribe como sigue:

$$\begin{bmatrix} 1 & -0,4996003197 & 0 & 0 \\ 0 & 1 & -0,663658224 & 0 \\ 0 & 0 & 1 & -0,740333954 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} 0,023980815 \\ 0,05307992589 \\ 0,0866782444 \\ 0,9015084838 \end{bmatrix},$$

$$\begin{aligned} u_4 &= 0,9015084838, \\ u_3 &= 0,0866782444 + 0,740333954 \times 0,9015084838 = 0,750955848, \\ u_2 &= 0,05307992589 + 0,663658224 \times 0,750955848 = 0,5535416624, \\ u_1 &= 0,023980815 + 0,4996003197 \times 0,5535416624 = 0,3005304065, \end{aligned}$$

con 3 cifras  $\vec{u}^T = (0,3, 0,554, 0,751, 0,902)$ ,  $u_0 = 0$  y  $u_5 = 1$ .

2. Considerar el problema de valores de frontera:

$$\begin{cases} -u''(x) + (1+x)u(x) = x^2 & x \in ]0, 1[, \\ u(0) = 1, \quad u(1) = 0. \end{cases}$$

Aplicar el método de diferencias finitas para aproximar la solución con  $n = 5$ . La matriz debe factorarse con el método  $LU$ .

### Solución

Sean  $n = 5$ ,  $h_x = \frac{1}{5} = 0,2$ ,  $x_j = 0,2j$   $j = 0, 1, \dots, 5$ . Utilizando diferencias finitas centrales, se tiene

$$u''(x_j) \simeq \frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1}))}{h_x^2}.$$

Sea  $u_j$  una aproximación de  $u(x_j)$ . Entonces el problema de valores de frontera:

$$\begin{cases} -u''(x) + (1+x)u(x) = x^2 & x \in ]0, 1[, \\ u(0) = 1, \quad u(1) = 0. \end{cases}$$

se aproxima mediante el esquema siguiente:

$$\begin{cases} -\frac{u(x_{j+1}) - 2u(x_j) + u(x_{j-1}))}{h_x^2} + (1+x_j)u_j = x_j^2, & j = 1, 2, 3, 4, \\ u(0) = u_0 = 1, \quad u(1) = u_5 = 0. \end{cases}$$

En forma explícita, este conjunto de ecuaciones se escribe:

$$\begin{aligned} j &= 1, & -\frac{u_2 - 2u_1 + 1}{h_x^2} + (1 + x_1)u_1 &= x_1^2, & u_0 &= 1, \\ j &= 2, & -\frac{u_3 - 2u_2 + u_1}{h_x^2} + (1 + x_2)u_2 &= x_2^2, \\ j &= 3, & -\frac{u_4 - 2u_3 + u_2}{h_x^2} + (1 + x_3)u_3 &= x_3^2, \\ j &= 4, & -\frac{2u_4 + u_3}{h_x^2} + (1 + x_4)u_4 &= x_4^2, & u_5 &= 0, \end{aligned}$$

que en forma matricial, se escribe como sigue:

$$\begin{bmatrix} \frac{2}{h_x^2} + 1 + x_1 & -\frac{1}{h_x^2} & 0 & 0 \\ -\frac{1}{h_x^2} & \frac{2}{h_x^2} + 1 + x_2 & -\frac{1}{h_x^2} & 0 \\ 0 & -\frac{1}{h_x^2} & \frac{2}{h_x^2} + 1 + x_3 & -\frac{1}{h_x^2} \\ 0 & 0 & -\frac{1}{h_x^2} & \frac{2}{h_x^2} + 1 + x_4 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} x_1^2 + \frac{1}{h_x^2} \\ x_2^2 \\ x_3^2 \\ x_4^2 \end{bmatrix}.$$

Poniendo  $h_x = 0,2$  y  $x_1 = 0,2$ ,  $x_2 = 0,4$ ,  $x_3 = 0,6$ ,  $x_4 = 0,8$ , se tiene

$$\begin{bmatrix} 51,20 & -25 & 0 & 0 \\ -25 & 51,4 & -25 & 0 \\ 0 & -25 & 51,6 & -25 \\ 0 & 0 & -25 & 51,8 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} 25,04 \\ 0,16 \\ 0,36 \\ 0,64 \end{bmatrix}.$$

Para hallar la solución de este sistema de ecuaciones lineales, apliquemos el método de factorización  $LU$ . Para el efecto, factoramos  $A = LU$ , donde

$$L = \begin{bmatrix} L_{11} & 0 & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 \\ 0 & L_{32} & L_{33} & 0 \\ 0 & 0 & L_{43} & L_{44} \end{bmatrix}, \quad U = \begin{bmatrix} 1 & u_{12} & 0 & 0 \\ 0 & 1 & u_{23} & 0 \\ 0 & 0 & 1 & u_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Puesto que

$$LU = \begin{bmatrix} L_{11} & L_{22}u_{12} & 0 & 0 \\ L_{21} & L_{21}u_{12} + L_{22} & L_{22}u_{23} & 0 \\ 0 & L_{32} & L_{32}u_{23} + L_{33} & L_{33}u_{34} \\ 0 & 0 & L_{43} & L_{43}u_{34} + L_{44} \end{bmatrix}$$

y de la igualdad  $A = LU$  se obtienen las matrices  $LU$  como siguen:

$$\begin{aligned} L_{11} &= 51,20, & u_{12} &= \frac{a_{12}}{L_{11}} = -\frac{25}{51,20} = -0,48828125, \\ L_{21} &= -25, & L_{22} &= a_{22} - L_{21}u_{12} = 51,4 - (-25)(-0,48828125) = 39,19296875, \\ & & u_{23} &= \frac{a_{23}}{L_{22}} - \frac{25}{39,19296875} = -0,6378695158, \\ L_{32} &= -25, & L_{33} &= a_{33} - L_{32}u_{23} = 51,66 - (-25)(-0,6378695158) = 35,65326211, \\ & & u_{34} &= \frac{a_{34}}{L_{33}} = -\frac{25}{35,65326211} = -0,7011981098, \\ L_{43} &= -25, & L_{44} &= a_{44} - L_{43}u_{34} = 51,8 - (-25)(-0,7011981098) = 34,27004726, \end{aligned}$$

Ponemos  $\vec{u}^T = (u_1, u_2, u_3, u_4)$ . El sistema de ecuaciones  $A\vec{u} = \vec{b}$  es equivalente al siguiente:

$$LU\vec{u} = \vec{b} \Leftrightarrow \begin{cases} L\vec{y} = \vec{b}, \\ U\vec{u} = \vec{y}. \end{cases}$$

Hallemos la solución del sistema de ecuaciones  $L\vec{y} = \vec{b}$ . Tenemos

$$\begin{bmatrix} 51,20 & 0 & 0 & 0 \\ -25 & 39,19296875 & 0 & 0 \\ 0 & -25 & 35,65326211 & 0 \\ 0 & 0 & -25 & 34,27004726 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} 25,04 \\ 0,16 \\ 0,36 \\ 0,64 \end{bmatrix},$$

luego

$$\begin{aligned} y_1 &= \frac{25,04}{51,20} = 0,4890625, \\ y_2 &= \frac{0,16 + 25 \times 0,4890625}{39,19296875} = 0,316040425, \\ y_3 &= \frac{1,6 + 25 \times 0,316040425}{35,65326211} = 0,2317042008, \\ y_4 &= \frac{26,8 + 25 \times 0,2317042008}{34,27004726} = 0,1877034184, \end{aligned}$$

Hallemos la solución del sistema de ecuaciones  $U\vec{u} = \vec{y}$  que en forma explícita se expresa como sigue:

$$\begin{bmatrix} 1 & -0,48828125 & 0 & 0 \\ 0 & 1 & -0,6378695158 & 0 \\ 0 & 0 & 1 & -0,7011981098 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} = \begin{bmatrix} 0,4890625 \\ 0,316040425 \\ 0,2317042008 \\ 0,1877034184 \end{bmatrix},$$

de donde

$$\begin{aligned} u_4 &= 0,1877034184, \\ u_3 &= 0,2317042008 + 0,7011981098 \times 0,1877034184 = 0,363321483, \\ u_2 &= 0,316040425 + 0,6378695158 \times 0,363321483 = 0,5477921234, \\ u_1 &= 0,4890625 + 0,48828125 \times 0,5477921234 = 0,7565391228, \end{aligned}$$

La aproximación de la ecuación diferencial en cuatro nodos internos con una precisión de 3 cifras es

$$\vec{u}^T = (0,757, 0,548, 0,363, 0,188).$$

3. Considerar el problema de valores de frontera no lineal siguiente:

$$\begin{cases} -u'' + u^3 = f & \text{sobre } ]0, 12[, \\ u(0) = u(12) = 0, \end{cases}$$

donde  $f(x) = \begin{cases} 0, & \text{si } x \in [0, 5], \\ 1, & \text{si } x \in ]5, 12[, \end{cases}$  y el esquema numérico siguiente:

$$\begin{cases} -\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + u_j^3 = f(x_j), & j = 1, \dots, n-1 \\ u_0 = 0, & u_n = 0, \end{cases}$$

con  $h = \frac{12}{n}$ ,  $n \in \mathbb{Z}^+$ ,  $x_j = jh$ ,  $j = 0, 1, \dots, n$  y  $u_j$  una aproximación de  $u(x_j)$ .

i) Para  $n = 6$ , construir el sistema no lineal correspondiente al esquema numérico propuesto.

ii) Considere una aproximación inicial  $\vec{u}^{(0)} = (0, 1, 1, 1, 1, 0)$ . Aplique el método de Newton y dos iteraciones.

### Solución

Puesto que se requiere generar un algoritmo general para aproximar la solución del problema de valores de frontera, hemos de proceder en ese sentido para luego particularizar al caso  $n = 6$ .

Sean  $L > 0$ . Suponemos que  $f$  es una función definida en  $[0, L]$ . Sea  $n \in \mathbb{Z}^+$ . Ponemos  $h = \frac{L}{n}$ ,  $x_j = jh$ ,  $j = 0, 1, \dots, n$ ,  $f_j = f(x_j)$ ,  $j = 0, 1, \dots, n$ . Del esquema numérico

$$\begin{cases} -\frac{u_{j+1} - 2u_j + u_{j-1}}{h^2} + u_j^3 = f(x_j), & j = 1, \dots, n-1, \\ u_0 = 0, & u_n = 0, \end{cases}$$

se obtiene las siguientes ecuaciones.

Para  $j = 1$ ,

$$-\frac{u_2 - 2u_1 + u_0}{h^2} + u_1^3 = f_1,$$

o bien

$$\frac{2}{h^2}u_1 - \frac{1}{h^2}u_2 + u_1^3 = f_1.$$

Para  $1 < j < n-1$ ,

$$-\frac{1}{h^2}u_{j-1} + \frac{2}{h^2}u_j - \frac{1}{h^2}u_{j+1} + u_j^3 = f_j.$$

Para  $j = n-1$

$$-\frac{u_n - 2u_{n-1} + u_{n-2}}{h^2} + u_{n-1}^3 = f_{n-1},$$

de donde

$$-\frac{1}{h^2}u_{n-2} + \frac{2}{h^2}u_{n-1} + u_{n-1}^3 = f_{n-1}.$$

Ponemos  $\vec{u}^T = (u_1, \dots, u_{n-1})$ ,  $u_0 = 0$ ,  $u_n = 0$ . Definimos la matriz  $A = (a_{ij}(h))$  y el vector  $B(\vec{u}) \in \mathbb{R}^{n-1}$  como sigue

$$a_{ij}(h) = \begin{cases} 0, & \text{si } |i-j| > 1, \\ \frac{2}{h^2}, & \text{si } i=j, \\ -\frac{1}{h^2}, & \text{si } |i-j|=1 \end{cases} \quad i, j = 1, \dots, n-1, \quad b_j(\vec{u}) = u_j^3, \quad j = 1, \dots, n-1.$$

Se define  $\vec{f}^T = (f(x_1), \dots, f(x_{n-1})) = (f_1, \dots, f_{n-1})$ .

Entonces, el esquema numérico propuesto, discretización del problema de valores de frontera

$$\begin{cases} -u'' + u^3 = f & \text{sobre } ]0, L[ \\ u(0) = 0 = u(L), \end{cases}$$

se transforma en el siguiente sistema de ecuaciones no lineales:

$$A\vec{u} + B(\vec{u}) = \vec{f}.$$

Se define  $F: \mathbb{R}^{n-1} \rightarrow \mathbb{R}^{n-1}$  como  $F(\vec{u}) = A\vec{u} + B(\vec{u}) - \vec{f}$   $\vec{u} \in \mathbb{R}^{n-1}$ . Para aplicar el método de Newton, requerimos de la matriz jacobiana  $DF(\vec{u})$ . Para el efecto, hallemos la derivada de Gâteaux (derivada direccional de  $F$ )  $D_{\vec{y}}F(\vec{u})$  según la dirección  $\vec{y} \in \mathbb{R}^{n-1}$  en  $\vec{u} \in \mathbb{R}^{n-1}$ . Por definición

$$D_{\vec{y}}F(\vec{u}) = \lim_{t \rightarrow 0} \frac{F(\vec{u} + t\vec{y}) - F(\vec{u})}{t},$$

siempre que el límite exista.

Sea  $t \neq 0$ . Entonces

$$\begin{aligned} F(\vec{u} + t\vec{y}) - F(\vec{u}) &= A(\vec{u} + t\vec{y}) + B(\vec{u} + t\vec{y}) - \vec{f} - (A\vec{u} + B(\vec{u}) - \vec{f}) \\ &= tA\vec{y} + B(\vec{u} + t\vec{y}) - B(\vec{u}). \end{aligned}$$

El  $j$ -ésimo componente de  $B(\vec{u} + t\vec{y}) - B(\vec{u})$  es

$$\begin{aligned} (u_j + ty_j)^3 - u_j^3 &= u_j^3 + 3tu_j^2y_j + 3t^2u_jy_j^2 + t^3y_j^3 - u_j^3 \\ &= ty_j(3u_j^2 + 3tu_jy_j + t^2y_j^2), \end{aligned}$$

de donde

$$\lim_{t \rightarrow 0} \frac{(u_j + ty_j)^3 - u_j^3}{t} = 3y_j u_j^2, \quad j = 1, \dots, n-1,$$

luego,

$$\begin{aligned} D\vec{y} F(\vec{u}) &= \lim_{t \rightarrow 0} \frac{F(\vec{u} + t\vec{y}) - F(\vec{u})}{t} = \lim_{t \rightarrow 0} \left( A\vec{y} + \frac{B(\vec{u} + t\vec{y}) - B(\vec{u})}{t} \right) \\ &= A\vec{y} + D\vec{y} B(\vec{u}), \end{aligned}$$

donde

$$D\vec{y} B(\vec{u}) = 3 \begin{bmatrix} u_1^2 & 0 & \dots & 0 \\ 0 & u_2^2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & \dots & u_{n-1}^2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \end{bmatrix} = 3c(\vec{u}) \vec{y},$$

con  $c(\vec{u}) = (c_{ij}(\vec{u}))$  la matriz definida como

$$c_{ij}(\vec{u}) = \begin{cases} u_i^3, & \text{si } i = j \\ 0, & \text{si } i \neq j \end{cases}, \quad i, j = 1, \dots, n-1.$$

Así,

$$D\vec{y} F(\vec{u}) = [A + 3c(\vec{u})] \vec{y} \quad \forall \vec{y} \in \mathbb{R}^{n-1},$$

con lo cual, la matriz jacobiana está definida como

$$DF(\vec{u}) = A + 3c(\vec{u}), \quad \forall \vec{u} \in \mathbb{R}^{n-1}.$$

El método de Newton está definido como sigue:

$$\begin{cases} \tilde{u}^{(0)} & \text{aproximación inicial,} \\ DF(\tilde{u}^{(k)}) \tilde{w} = -F(\tilde{u}^{(k)}), & k = 0, 1, \dots, N_{\text{máx}}, \\ \tilde{u}^{(k+1)} = \tilde{u}^{(k)} + \tilde{w}. \end{cases}$$

i) Para  $n = 6$  se tiene  $h = \frac{L}{n} = \frac{12}{6} = 2$ ,  $x_j = jh$ ,  $j = 0, 1, \dots, n$ , la partición del intervalo  $[0, 12]$  está constituida por los siguientes nodos:  $x_0 = 0$ ,  $x_1 = 2$ ,  $x_2 = 4$ ,  $x_3 = 6$ ,  $x_4 = 8$ ,  $x_5 = 10$ ,  $x_6 = 12$ , y la función  $f$  en dichos nodos tiene los valores siguientes:  $f(0) = 0$ ,  $f(2) = 0$ ,  $f(4) = 0$ ,  $f(6) = 1$ ,  $f(8) = 1$ ,  $f(10) = 1$ ,  $f(12) = 1$ .

Luego,  $\tilde{f} = (0, 0, 1, 1, 1)$ ,

$$A = \begin{bmatrix} \frac{2}{h^2} & -\frac{1}{h^2} & 0 & 0 & 0 \\ -\frac{1}{h^2} & \frac{2}{h^2} & -\frac{1}{h^2} & 0 & 0 \\ 0 & -\frac{1}{h^2} & \frac{2}{h^2} & -\frac{1}{h^2} & 0 \\ 0 & 0 & -\frac{1}{h^2} & \frac{2}{h^2} & -\frac{1}{h^2} \\ 0 & 0 & 0 & -\frac{1}{h^2} & \frac{2}{h^2} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 0 & -\frac{1}{4} & \frac{1}{2} \end{bmatrix},$$

$$B(\vec{u}) = \begin{bmatrix} u_1^3 \\ u_2^3 \\ u_3^3 \\ u_4^3 \\ u_5^3 \end{bmatrix}, \quad c(\vec{u}) = \begin{bmatrix} u_1^2 & 0 & 0 & 0 & 0 \\ 0 & u_2^2 & 0 & 0 & 0 \\ 0 & 0 & u_3^2 & 0 & 0 \\ 0 & 0 & 0 & u_4^2 & 0 \\ 0 & 0 & 0 & 0 & u_5^2 \end{bmatrix}.$$



El sistema no lineal de ecuaciones es

$$\begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 0 & -\frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} + \begin{bmatrix} u_1^3 \\ u_2^3 \\ u_3^3 \\ u_4^3 \\ u_5^3 \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \\ f_3 \\ f_4 \\ f_5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix},$$

por lo tanto

$$F(\vec{u}) = A\vec{u} + B(\vec{u}) - \vec{f} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 0 & -\frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ u_5 \end{bmatrix} + \begin{bmatrix} u_1^3 \\ u_2^3 \\ u_3^3 \\ u_4^3 \\ u_5^3 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}.$$

ii) Sea  $(\vec{u}^{(0)})^T = (0, 1, 1, 1, 1, 0)$  una aproximación inicial. Ponemos  $(\vec{u}^{(0)})^T = (1, 1, 1, 1, 1)$ . Entonces

$$B(\vec{u}^{(0)}) = \begin{bmatrix} 1^3 \\ 1^3 \\ 1^3 \\ 1^3 \\ 1^3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad A\vec{u}^{(0)} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 0 & -\frac{1}{4} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{1}{4} \\ 0 \\ 0 \\ 0 \\ \frac{1}{4} \end{bmatrix},$$

$$F(\vec{u}^{(0)}) = A\vec{u}^{(0)} + B(\vec{u}^{(0)}) - \vec{f} = \begin{bmatrix} \frac{1}{4} \\ 0 \\ 0 \\ 0 \\ \frac{1}{4} \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{5}{4} \\ 1 \\ 0 \\ 0 \\ \frac{1}{4} \end{bmatrix},$$

$$\begin{aligned}
DF\left(\tilde{u}^{(0)}\right) &= A + 3c\left(\tilde{u}^{(0)}\right) = \begin{bmatrix} \frac{1}{2} & -\frac{1}{4} & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{1}{2} & -\frac{1}{4} \\ 0 & 0 & 0 & -\frac{1}{4} & \frac{1}{2} \end{bmatrix} + 3 \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} \frac{7}{2} & -\frac{1}{4} & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{7}{2} & -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{7}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{7}{2} & -\frac{1}{4} \\ 0 & 0 & 0 & -\frac{1}{4} & \frac{7}{2} \end{bmatrix}
\end{aligned}$$

El sistema de ecuaciones lineales es:

$$\begin{bmatrix} \frac{7}{2} & -\frac{1}{4} & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{7}{2} & -\frac{1}{4} & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{7}{2} & -\frac{1}{4} & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{7}{2} & -\frac{1}{4} \\ 0 & 0 & 0 & -\frac{1}{4} & \frac{7}{2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} = - \begin{bmatrix} \frac{5}{4} \\ 1 \\ 0 \\ 0 \\ \frac{1}{4} \end{bmatrix}.$$

Método de Crout. Sea  $D$  la matriz del sistema de ecuaciones precedente. Entonces,  $D = LU$  y

$$D\vec{w} = \vec{b} \Leftrightarrow LU\tilde{w} = \tilde{b} \Leftrightarrow \begin{cases} L\tilde{y} = \tilde{b} \\ U\tilde{w} = \tilde{y}. \end{cases}$$

Comencemos con la factorización

$$\begin{aligned}
D &= LU = \begin{bmatrix} L_{11} & 0 & 0 & 0 & 0 \\ L_{21} & L_{22} & 0 & 0 & 0 \\ 0 & L_{32} & L_{33} & 0 & 0 \\ 0 & 0 & L_{43} & L_{44} & 0 \\ 0 & 0 & 0 & L_{54} & L_{55} \end{bmatrix} \begin{bmatrix} 1 & u_{12} & 0 & 0 & 0 \\ 0 & 1 & u_{23} & 0 & 0 \\ 0 & 0 & 1 & u_{34} & 0 \\ 0 & 0 & 0 & 1 & u_{45} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} L_{11} & L_{11}u_{12} & 0 & 0 & 0 \\ L_{21} & L_{21}u_{12} + L_{22} & L_{22}u_{23} & 0 & 0 \\ 0 & L_{32} & L_{32}u_{23} + L_{33} & L_{33}u_{34} & 0 \\ 0 & 0 & L_{43} & L_{43}u_{34} + L_{44} & L_{44}u_{45} \\ 0 & 0 & 0 & L_{54} & L_{44}u_{45} + L_{55} \end{bmatrix}
\end{aligned}$$

Obtenemos

$$\begin{aligned}
 L_{11} &= \frac{7}{2}, \\
 u_{12} &= \frac{a_{12}}{L_{11}} = \frac{-\frac{1}{4}}{\frac{7}{2}} = -\frac{1}{14}, \\
 L_{21} &= -\frac{1}{4}, \\
 L_{22} &= a_{22} - L_{21}u_{12} = \frac{7}{2} - \left(-\frac{1}{4}\right)\left(-\frac{1}{14}\right) = \frac{195}{56}, \\
 u_{23} &= \frac{a_{23}}{L_{22}} = \frac{-\frac{1}{4}}{\frac{195}{56}} = -\frac{14}{195}, \\
 L_{32} &= -\frac{1}{4}, \\
 L_{33} &= a_{33} - L_{32}u_{23} = \frac{7}{2} - \left(-\frac{1}{4}\right)\left(-\frac{14}{195}\right) = \frac{1358}{390} = \frac{679}{195}, \\
 u_{34} &= \frac{a_{34}}{L_{33}} = \frac{-\frac{1}{4}}{\frac{1358}{390}} = -\frac{195}{2716}, \\
 L_{43} &= -\frac{1}{4}, \\
 L_{44} &= a_{44} - L_{43}u_{34} = \frac{7}{2} - \left(-\frac{1}{4}\right)\left(-\frac{195}{2716}\right) = \frac{37829}{10864}, \\
 u_{45} &= \frac{a_{45}}{L_{44}} = \frac{-\frac{1}{4}}{\frac{37829}{10864}} = -\frac{2716}{37829}, \\
 L_{54} &= -\frac{1}{4}, \\
 L_{55} &= a_{55} - L_{54}u_{45} = \frac{7}{2} - \left(-\frac{1}{4}\right)\left(-\frac{2716}{37829}\right) = \frac{263445}{75658}.
 \end{aligned}$$

Resolución del sistema triangular inferior  $L\vec{y} = \vec{b}$ , esto es,

$$\begin{bmatrix} \frac{7}{2} & 0 & 0 & 0 & 0 \\ -\frac{1}{4} & \frac{195}{56} & 0 & 0 & 0 \\ 0 & -\frac{1}{4} & \frac{679}{195} & 0 & 0 \\ 0 & 0 & -\frac{1}{4} & \frac{37829}{10864} & 0 \\ 0 & 0 & 0 & -\frac{1}{4} & \frac{263445}{75658} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} = \begin{bmatrix} -\frac{5}{4} \\ -1 \\ 0 \\ 0 \\ -\frac{1}{4} \end{bmatrix}.$$

Se tiene

$$\begin{aligned}
 y_1 &= \frac{-\frac{5}{4}}{\frac{7}{2}} = -\frac{5}{14} = -0,3571428571, \\
 -\frac{1}{4}y_1 + \frac{195}{56}y_2 &= -1 \Rightarrow y_2 = \frac{56}{195} \left( -1 + \frac{1}{4}y_1 \right) = -0,3128205128, \\
 -\frac{1}{4}y_2 + \frac{679}{195}y_3 &= 0 \Rightarrow y_3 = \frac{195}{679} \times \frac{1}{4}y_2 = -0,02245941926, \\
 -\frac{1}{4}y_3 + \frac{37829}{10864}y_4 &= 0 \Rightarrow y_4 = \frac{10864}{37829} \times \frac{1}{4}y_3 = -0,0004031298739, \\
 -\frac{1}{4}y_4 + \frac{263445}{75658}y_5 &= -\frac{1}{4} \Rightarrow y_5 = \frac{75658}{263445} \left( -\frac{1}{4} + \frac{1}{4}y_4 \right) = -0,07182571318,
 \end{aligned}$$

Resolución del sistema triangular superior  $U\vec{w} = \vec{y}$  :

$$\begin{bmatrix} 1 & -\frac{1}{14} & 0 & 0 & 0 \\ 0 & 1 & -\frac{14}{195} & 0 & 0 \\ 0 & 0 & 1 & -\frac{195}{2716} & 0 \\ 0 & 0 & 0 & 1 & -\frac{2716}{37829} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} = \begin{bmatrix} -0,3571428571 \\ -0,3128205128 \\ -0,02245941926 \\ -0,0004031298739 \\ -0,07182571318 \end{bmatrix}$$

Obtenemos

$$\begin{aligned}
 w_5 &= -0,07182571318, \\
 w_4 - \frac{2716}{37829}w_5 &= -0,0004031298739, \\
 w_4 &= \frac{2716}{37829}w_5 - 0,0004031298739 = -0,00555998406, \\
 w_3 - \frac{195}{2716}w_4 &= -0,02245949926, \\
 w_3 &= \frac{195}{2716}w_4 - 0,02245949926 = -0,228586881, \\
 w_2 - \frac{14}{195}w_3 &= -0,3128205128, \\
 w_2 &= \frac{14}{195}w_3 - 0,3128205128 = -0,3144616494, \\
 w_1 - \frac{1}{14}w_2 &= -0,3571428571, \\
 w_1 &= \frac{1}{14}w_2 - 0,3571428571 = -0,3796044035,
 \end{aligned}$$

Luego

$$\tilde{u}^{(1)} = \tilde{u}^{(0)} + \tilde{w} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -0,3796044035 \\ -0,3144616494 \\ -0,228586881 \\ -0,00555998406 \\ -0,07182571318 \end{bmatrix} = \begin{bmatrix} 0,6203955965 \\ 0,6855383506 \\ 0,9771413119 \\ 0,9944400159 \\ 0,9281742868 \end{bmatrix}.$$

Segunda iteración

La matriz  $A$  y el vector  $\tilde{f}$  no cambian. Calculemos  $B(\tilde{u}^{(1)})$ . Tenemos

$$B(\tilde{u}^{(1)}) = \begin{bmatrix} \left(u_1^{(1)}\right)^3 \\ \left(u_2^{(1)}\right)^3 \\ \left(u_3^{(1)}\right)^3 \\ \left(u_4^{(1)}\right)^3 \\ \left(u_5^{(1)}\right)^3 \end{bmatrix} = \begin{bmatrix} 0,238784493 \\ 0,3221775434 \\ 0,9341234603 \\ 0,9987910978 \\ 0,7996291156 \end{bmatrix}.$$

Calculemos  $F(\tilde{u}^{(1)}) = A\tilde{u}^{(1)} + B(\tilde{u}^{(1)}) - \tilde{f}$ . Tenemos

$$\begin{bmatrix} 0,5 & -0,25 & 0 & 0 & 0 \\ -0,25 & 0,5 & -0,25 & 0 & 0 \\ 0 & -0,25 & 0,5 & -0,25 & 0 \\ 0 & 0 & -0,25 & 0,5 & -0,25 \\ 0 & 0 & 0 & -0,25 & 0,5 \end{bmatrix} \begin{bmatrix} 0,6203955965 \\ 0,6855383506 \\ 0,9771413119 \\ 0,9944400159 \\ 0,9281742868 \end{bmatrix} + \begin{bmatrix} 0,238784493 \\ 0,3221775434 \\ 0,9341234603 \\ 0,9987910978 \\ 0,7996291156 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix} \\ = \begin{bmatrix} 0,3775977034 \\ 0,2654626945 \\ 0,00160990548 \\ 0,022160836 \\ 0,0138170415 \end{bmatrix}$$

Calculemos  $DF(\tilde{u}^{(1)}) = A + 3c(\tilde{u}^{(1)})$ .

Se tiene

$$A + 3c(\tilde{u}^{(1)}) = \begin{bmatrix} 0,5 & -0,25 & 0 & 0 & 0 \\ -0,25 & 0,5 & -0,25 & 0 & 0 \\ 0 & -0,25 & 0,5 & -0,25 & 0 \\ 0 & 0 & -0,25 & 0,5 & -0,25 \\ 0 & 0 & 0 & -0,25 & 0,5 \end{bmatrix} \\ + 3 \begin{bmatrix} \left(u_1^{(1)}\right)^2 & 0 & 0 & 0 & 0 \\ 0 & \left(u_2^{(1)}\right)^2 & 0 & 0 & 0 \\ 0 & 0 & \left(u_3^{(1)}\right)^2 & 0 & 0 \\ 0 & 0 & 0 & \left(u_4^{(1)}\right)^2 & 0 \\ 0 & 0 & 0 & 0 & \left(u_5^{(1)}\right)^2 \end{bmatrix} \\ = \begin{bmatrix} 1,654672089 & -0,25 & 0 & 0 & 0 \\ -0,25 & 1,90988849 & -0,25 & 0 & 0 \\ 0 & -0,25 & 3,366756292 & -0,25 & 0 \\ 0 & 0 & -0,25 & 3,497581708 & -0,25 \\ 0 & 0 & 0 & -0,25 & 3,08452252 \end{bmatrix}$$

El sistema de ecuaciones lineales correspondiente  $DF(\tilde{u}^{(1)})\tilde{w} = -F(\tilde{u}^{(1)})$  es el siguiente:

$$\begin{bmatrix} 1,654672089 & -0,25 & 0 & 0 & 0 \\ -0,25 & 1,90988849 & -0,25 & 0 & 0 \\ 0 & -0,25 & 3,366756292 & -0,25 & 0 \\ 0 & 0 & -0,25 & 3,497581708 & -0,25 \\ 0 & 0 & 0 & -0,25 & 3,08452252 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} \\ = \begin{bmatrix} -0,3775977034 \\ -0,2654626945 \\ -0,00160990548 \\ -0,02217505055 \\ -0,0138170415 \end{bmatrix}$$

Resolvemos este sistema utilizando el método de Crout. Para ello la matriz del sistema lo factoramos en la forma  $LU$  como se procedió en la primera iteración. Note que la matriz del sistema es simétrica, estrictamente diagonalmente dominante. Se obtienen los siguientes resultados:

$$\begin{aligned}
 L_{11} &= 1,654672089, \\
 u_{12} &= \frac{-0,25}{1,654672089} = -0,1510873373, \\
 L_{21} &= -0,25, \\
 L_{22} &= a_{22} - L_{21}u_{12} = 1,872116656, \\
 u_{23} &= \frac{a_{23}}{L_{22}} = \frac{-0,25}{0,872116656} = -0,1335386869, \\
 L_{32} &= -0,25, \\
 L_{33} &= a_{33} - L_{32}u_{23} = 3,33337162, \\
 u_{34} &= \frac{a_{34}}{L_{33}} = -0,07499913855, \\
 L_{43} &= -0,25, \\
 L_{44} &= a_{44} - L_{43}u_{34} = 3,478831923, \\
 u_{45} &= \frac{a_{45}}{L_{44}} = 0,7186320165, \\
 L_{54} &= -0,25, \\
 L_{55} &= a_{55} - L_{54}u_{45} = 3,06655672.
 \end{aligned}$$

El sistema triangular inferior  $L\vec{y} = -F(\tilde{u}^{(1)})$  es el siguiente:

$$\begin{aligned}
 &\begin{bmatrix} 1,654672089 & 0 & 0 & 0 & 0 \\ -0,25 & 1,872116656 & 0 & 0 & 0 \\ 0 & -0,25 & 3,33337162 & 0 & 0 \\ 0 & 0 & -0,25 & 3,478831923 & 0 \\ 0 & 0 & 0 & -0,25 & 3,06655682 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix} \\
 &= \begin{bmatrix} -0,3775977034 \\ -0,2654626945 \\ -0,00160990548 \\ -0,02217505055 \\ -0,0138170415 \end{bmatrix}
 \end{aligned}$$

cuya solución es

$$\begin{aligned}
 y_1 &= \frac{-0,3775977034}{1,654672089} = -0,2282009263, \\
 y_2 &= \frac{0,25y_1 - 0,2654626945}{1,872116656} = -0,1722718107, \\
 y_3 &= \frac{0,25y_2 - 0,00160990548}{3,33337162} = -0,0134032035, \\
 y_4 &= \frac{0,25y_3 - 0,02217505055}{3,478831923} = -0,007337477635, \\
 y_5 &= \frac{0,25y_4 - 0,0138170415}{3,06655682} = -0,005103903935.
 \end{aligned}$$

Resolución del sistema triangular superior  $U\tilde{w} = \vec{y}$  :

$$\begin{bmatrix} 1 & -0,1510873383 & 0 & 0 & 0 \\ 0 & 1 & -0,1335386869 & 0 & 0 \\ 0 & 0 & 1 & -0,07499913855 & 0 \\ 0 & 0 & 0 & 1 & -0,7186320165 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ w_5 \end{bmatrix} = \begin{bmatrix} -0,2282009263 \\ -0,1722718107 \\ -0,0134032035 \\ -0,007337477635 \\ -0,005103903935 \end{bmatrix}$$

Obtenemos

$$\begin{aligned} w_5 &= -0,005103903935, \\ w_4 &= -0,01100530641, \\ w_3 &= -0,014228592, \\ w_2 &= -0,1741718782, \\ w_1 &= -0,2545160916, \end{aligned}$$

En consecuencia:

$$\tilde{u}^{(2)} = \tilde{u}^{(1)} + \vec{w} = \begin{bmatrix} 0,6203955965 \\ 0,6855383506 \\ 0,9775405007 \\ 0,9995968701 \\ 0,9281742868 \end{bmatrix} + \begin{bmatrix} -0,2545160916 \\ -0,1741718782 \\ -0,014228592 \\ -0,01100530641 \\ -0,005103903935 \end{bmatrix} = \begin{bmatrix} 0,3658795049 \\ 0,5113664724 \\ 0,9633119087 \\ 0,9885915637 \\ 0,9230703829 \end{bmatrix},$$

y con  $u_0^{(2)} = 0$ ,  $u_6^{(2)} = 0$  se obtiene  $\vec{u}^{(2)}$ .

## 11.6. Lecturas complementarias y bibliografía

1. Tom M. Apostol, Calculus, Volumen 1, Segunda Edición, Editorial Reverté, Barcelona, 1977.
2. Uri M. Ascher, Robert M. M. Mattheij, Robert D. Russell, Numerical Solution of Boundary Value Problems for Ordinary Differential Equations, Editorial Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1995.
3. N. Bakhvalov, Metodos Numéricos, Editorial Paraninfo, Madrid, 1980.
4. E. K. Blum, Numerical Analysis and Computation. Theory and Practice, Editorial Addison-Wesley Publishing Company, Reading, Massachusetts, 1972.
5. Richard L. Burden, J. Douglas Faires, Análisis Numérico, Séptima Edición, International Thomson Editores, S. A., México, 2002.
6. Steven C. Chapra, Raymond P. Canale, Numerical Methods for Engineers, Third Edition, Editorial McGraw-Hill, Boston, 1998.
7. S. D. Conte, Carl de Boor, Análisis Numérico, Segunda Edición, Editorial Mc Graw-Hill, México, 1981.
8. M. Crouzeix, A. L. Mignot, Analyse Numérique des Equations Différentielles, Seconde Edition, Editorial Masson, París, 1989.
9. Jean-Pierre Demailly, Analyse Numérique et Equations differentielles, Presses Universitaires de Grenoble, Grenoble, 1991.

10. Peter Deuffhard, Folkmar Bornemann, Scientific Computing with Ordinary Differential Equations, Editorial Springer-Verlag, New York, 2002.
11. B. P. Demidovich, I. A. Maron, E. Cálculo Numérico Fundamental, Editorial Paraninfo, Madrid, 1977.
12. B. P. Demidovich, I. A. Maron, E. S. Schuwalowa, Métodos Numéricos de Análisis, Editorial Paraninfo, Madrid, 1980.
13. James W. Demmel, Applied Numerical Linear Algebra, Editorial Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1997.
14. C. H. Edwards, Jr., David E. Penney, Ecuaciones Diferenciales Elementales y Problemas con Condiciones en la Frontera, Tercera Edición, Editorial Prentice-Hall Hispanoamericana, S. A., México, 1993.
15. Ferruccio Fontanella, Aldo Pasquali, Calcolo Numerico. Metodi e Algoritmi, Volume II Pitagora Editrice Bologna, 1983.
16. M. K. Gavurin, Conferencias sobre los Métodos de Cálculo, Editorial Mir, Moscú, 1973.
17. Curtis F. Gerald, Patrick O. Wheatley, Análisis Numérico con Aplicaciones, Sexta Edición, Editorial Pearson Educación de México, México, 2000.
18. E. Hairer, S. P. Norsett, G. Wanner, Solving Ordinary Differential Equations I, Second Revised Edition, Editorial Springer-Verlag, Berlín, 2000.
19. R. Kent Nagle, Edward B. Saff, Arthur David Snider, Ecuaciones Diferenciales y Problemas con Valores en la Frontera, Tercera Edición, editorial Pearson Educación, México, 2001.
20. David Kincaid, Ward Cheney, Análisis Numérico, Editorial Addison-Wesley Iberoamericana, Wilmington, 1994.
21. Melvin J. Maron, Robert J. López, Análisis Numérico, Tercera Edición, Compañía Editorial Continental, México, 1995.
22. R. M. M. Mattheij, J. Molenaar, Ordinary Differential Equations in Theory and Practice, Editorial John Wiley & Sons, New York, 1996.
23. Shoichiro Nakamura, Métodos Numérico Aplicados con Software, Editorial Prentice-Hall Hispanoamericana, S. A., México, 1992.
24. Antonio Nieves, Federico C. Dominguez, Métodos Numéricos Aplicados a la Ingeniería, Tercera Reimpresión, Compañía Editorial Continental, S. A. De C. V., México, 1998.
25. S. Nikolski, Fórmulas de Cuadratura, Editorial Mir, Moscú, 1990.
26. J. M. Ortega, W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Editorial Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 2000.
27. Anthony Ralston, Introducción al Análisis Numérico, Editorial Limusa, México, 1978.
28. Werner C. Rheinboldt, Methods for Solving Systems of Nonlinear Equations, Second Edition, Editorial Society for Industrial and Applied Mathematics (SIAM), Philadelphia, 1998.
29. A. A. Samarski, Introducción a los Métodos Numéricos, Editorial Mir, Moscú, 1986.
30. M. Sibony, J. Cl. Mardon, Analyse Numérique II, Approximations et Equations Différentielles, Editorial Hermann, París, 1988.
31. J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, Editorial Springer-Verlag, 1980.