Ciencias Matemáticas, Vol. 32, No. 2, Pag. 115-123, 2018
Recibido 02-2018

Método de las diferencias finitas aplicado a un problema elíptico unidimensional con condiciones de discontinuidad
Finite differences method applied to a one-dimensional elliptic problem with discontinuity conditions

Frank Ernesto Alvarez Borges[1], Julián Bravo Castillero[1]*, Angela León Mecías[2], Raúl Guinovart Díaz[1], Reinaldo Rodriguez Ramos[1]

Resumen En el presente trabajo es planteada la ecuación del calor con condiciones de contacto imperfecto, para luego obtener un esquema en diferencias finitas que aproxime al problema previamente presentado. Es analizada la consistencia, estabilidad y convergencia del método utilizado, para luego discutir algunos ejemplos.

Abstract The heat equation with non-perfect contact conditions is stated, and then, a finite difference scheme is obtained in order to approximate it. Consistency, stability and convergence is studied and some examples are discussed.

Palabras Clave
Problema elíptico — Contacto Imperfecto — Diferencias finitas

[1] *Departamento de Matemática, Facultad de Matemática y Computación, Universidad de la Habana, La Habana, Cuba, jb_castillero@yahoo.com.mx,*
[2] *Departamento de Matemática Aplicada, Facultad de Matemática y Computación, Universidad de la Habana, La Habana, Cuba, angela@matcom.uh.cu*
***Autor para Correspondencia**

Introducción

La ecuación del calor, acoplada con determinadas condiciones de contacto (que puede ser perfecto o imperfecto), aparece en el estudio de la transmisión del calor por regiones compuestas por materiales de distintas propiedades, así como en otros problemas de la mecánica de sólidos.

Esta diferencia de propiedades puede ser modelada a través de condiciones de discontinuidad. Diversos árticulos han estudiado este tipo de problemas mediante un enfoque numérico. En [**?**] y [**?**] se abordan problemas que presentan discontinuidades mediante el método de Elementos Finitos. El mf́odo de las Diferencias Finitas, el cual será abordado en el presente documento, fue utilizado en [**?**] para resolver las ecuaciones de Maxwell en el dominio del tiempo, en [**?**] y [**?**] es considerado un sistema de ecuaciones con coeficientes dependientes de variables temporales y espaciales, con condiciones de discontinuidad, en una y dos dimensiones respectivamente.

A continuación será planteada la ecuación que modela un problema estacionario de conducción del calor, en el caso unidimensional con condiciones de contacto imperfecto. Más adelante, en la Sección 1, el método de las diferencias finitas es aplicado al problema previamente obtenido. En la Sección 2 se analizará la consistencia y estabilidad del método, así como la convergencia de la solución numérica a la solución exacta. Finalmente en la Sección 3 son presentados y discutidos algunos ejemplos.

0.1 Obtención de la ecuación del calor

La ecuación del calor con convección viene dada por la expresión

$$d\frac{\partial u}{\partial t} - \nabla\cdot(a\nabla u) + b\cdot\nabla u + cu = f \text{ en } \Omega, \text{ donde } c = \nabla\cdot b. \tag{1}$$

La deducción de esta ecuación a partir de una ley de conservación y varias relaciones constitutivas puede ser encontrada en la introducción de [3]. Para el caso de un problema estacionario (independiente del tiempo), unidimensional (por ejemplo, el problema de conducción del calor sobre una barra metálica),

donde $b = 0$, la ecuación (1) queda reducida a

$$-\frac{d}{d\xi}\Big(a(\xi)\frac{d}{d\xi}u\Big) = f, \text{ en } \Omega. \tag{2}$$

En dicho caso Ω sería un intervalo. Si además se considera que la barra está conformada por secciones, cada una de diferente material, y por lo tanto, con diferentes propiedades, en cada uno de los puntos de enlace de dichas secciones deben cumplirse condiciones del tipo Robin

$$a\frac{du}{d\xi} + \kappa(u^+ - u^-) = 0,$$

donde κ es un coeficiente de transferencia de calor. Es decir, sobre los enlaces se considerará el flujo proporcional a la diferencia entre las temperaturas de las secciones en cuestión. Además se considerará que el flujo se mantiene constante sobre los puntos de enlace. Denotando como ξ_k los puntos de encuentro de dos secciones consecutivas, estas condiciones quedan como

$$a(\xi)\frac{d}{d\xi}u(\xi)\Big|_{\xi_k^-} = \beta_k[\![u(\xi)]\!]_{\xi_k}, \tag{3}$$

y

$$\Big[\!\Big[a(\xi)\frac{d}{d\xi}u(\xi)\Big]\!\Big]_{\xi_k} = 0. \tag{4}$$

Las constantes β_k son conocidas como números de Biot. Para las condiciones de frontera, que en el caso unidimensional serán dadas sobre los puntos extremos de Ω, serán seleccionadas condiciones de Dirichlet, $u(\xi_0) = u_0$ y $u(\xi_f) = u_f$. De este modo, con la ecuación (2) junto con las condiciones (3) y (4), además de las condiciones de frontera, queda establecido el problema a tratar numéricamente por el método de las diferencias finitas.

Bajo determinadas condiciones de diferenciabilidad de las funciones $a(\xi)$ y $f(\xi)$, y de positividad de $a(\xi)$, el problema (2)-(4) posee solución diferenciable hasta determinado orden, cuya expresión analítica, para $\xi \in (\xi_j, \xi_{j+1})$, es

$$u(\xi) = -\int_{\xi_0}^{\xi}\frac{1}{a(s)}\Big(\int_{\xi_0}^{s} f(t)dt - c\Big)ds - \sum_{i=1}^{j}\frac{1}{\beta_j}\Big(\int_{\xi_0}^{\xi_j} f(\xi)d\xi - c\Big) + u_0, \tag{5}$$

con

$$c = \frac{\Big[u_f - u_0 + \langle\frac{1}{a(\xi)}\Big(\int_{\xi_0}^{\xi} f(s)ds\Big)\rangle + \sum_{i=1}^{l}\frac{1}{\beta_j}\int_{\xi_0}^{\xi_j} f(\xi)d\xi\Big]}{\Big(\langle\frac{1}{a(\xi)}\rangle + \sum_{i=1}^{l}\frac{1}{\beta_j}\Big)},$$

donde

$$\langle F(\xi)\rangle = \int_{\xi_0}^{\xi_f} F(\xi)d\xi.$$

Esta expresión de $u(\xi)$ garantiza su acotación, así como la de sus derivadas.

Nótese que las integrales presentes en (5) podrían no ser expresables en términos de funciones elementales, lo cual representa otro factor para analizar numéricamente el problema (2)-(4).

1. Obtención del sistema de ecuaciones lineales para el problema con condiciones de discontinuidad

1.1 Formulación del problema

Sean $\{\xi_k\}_{k=1}^{l}$ números reales tales que

$$\xi_0 = 0 < \xi_1 < \xi_2 < \ldots < \xi_l < \xi_{l+1} = 1.$$

Sean además $f(\xi)$ y $a(\xi)$ funciones reales infinitamente diferenciables para todo $\xi \in [\xi_i, \xi_{i+1}]$, $i = 0, \ldots, l$, con $a(\xi)$ satisfaciendo la desigualdad

$$0 < \alpha_0 \le a(\xi) \le \alpha_1$$

(donde α_0 y α_1 son constantes reales). Considérese también el operador contraste definido como

$$[\![F(\xi)]\!]_{\xi_i} = \lim_{\xi\to\xi_i^+} F(\xi) - \lim_{\xi\to\xi_i^-} F(\xi).$$

El objetivo perseguido es aplicar el método de las diferencias finitas al problema

$$-\frac{d}{d\xi}\Big(a(\xi)\frac{du}{d\xi}\Big) = f(\xi),\ \xi \in \Omega^* = (0,1)\setminus\{\xi_k\}, \tag{6}$$

$$a(\xi)\frac{du}{d\xi}\Big|_{\xi_k^-} = \beta_k[\![u(\xi)]\!]_{\xi_k}, \tag{7}$$

$$[\![a(\xi)\frac{du}{d\xi}]\!]_{\xi_k} = 0, \tag{8}$$

con las condiciones de frontera $u(0) = t_0$ y $u(1) = t_1$, donde $t_0, t_1, \beta_i \in \mathbb{R}$, $\beta_k > 0$ para todo $k = 1, \ldots, l$.

1.2 Simplificación del problema

Para abordar el problema numéricamente, resultaría provechoso que los puntos de discontinuidad formaran parte de la malla que será seleccionada, para de este modo utilizar la información brindada por las condiciones (7)-(8). Esto impediría en muchos casos que dicha malla forme una red uniforme. La demostración de este hecho será dada a continuación:

Sea $\hat{\xi}$ un punto del intervalo $(0,1)$. Suponiendo que existe una red uniforme de dicho intervalo, que contiene a $\hat{\xi}$, entonces existen naturales p y q, y un número real h, tal que $ph = \hat{\xi}$ y $qh = 1$, consecuentemente, $\hat{\xi}$ puede ser expresado como $\xi_0 = p/q$. Esta expresión conduce a una contradicción en el caso en que $\hat{\xi}$ sea irracional. En conclusión, dado un conjunto de números reales en el intervalo $(0,1)$, no siempre es posible encontrar una red uniforme que contenga a dichos puntos como nodos.

Trabajar con una malla uniforme, simplifica enormemente las expresiones obtenidas para las aproximaciones de las derivadas de las funciones en cuestión. Por esta razón será mostrado a continuación como transformar el problema (6)-(8), donde los puntos de discontinuidad no son equidistantes entre si, en un problema equivalente donde los puntos de discontinuidad formen una red uniforme.

Para ello considérese la función $g(x)$, lineal por tramos y continua, obtenida de unir los puntos $x_i = \frac{i}{l+1}$ con los puntos ξ_i, para $i = 0, \ldots, l+1$. La expresión analítica para esta función es

$$g(x) = (\xi_{i+1} - \xi_i)(l+1)x + \xi_i(i+1) - i\xi_{i+1} \text{ si } \frac{i}{l+1} < x < \frac{i+1}{l+1},$$

para $i = 0, \ldots, l$. Esta función es positiva, diferenciable en todo el intervalo $(0,1)$, excepto por los puntos $\frac{i}{l+1}$, $i = 1, \ldots, l$, y en los puntos donde la derivada existe, es mayor que cero. Definiendo las funciones $a_1(x) = \frac{a(g(x))}{g'(x)}$ y $u_1(x) = u(g(x))$, se tiene entonces que

$$\begin{aligned}
&-\frac{d}{dx}\Big(a_1(x)\frac{d}{dx}u_1(x)\Big) \\
&= -\frac{d}{dx}\Big(\frac{a(g(x))}{g'(x)}\frac{d}{d\xi}u(\xi)\Big|_{\xi=g(x)}g'(x)\Big) \\
&= -\frac{d}{dx}\Big(a(g(x))\frac{d}{d\xi}u(\xi)\Big|_{\xi=g(x)}\Big) \\
&= -\frac{d}{d\xi}\Big(a(\xi)\frac{d}{d\xi}u(\xi)\Big)\Big|_{\xi=g(x)}g'(x).
\end{aligned}$$

Consequentemente, sabiendo que $u(\xi)$ es solución de (6), entonces $u_1(x)$ es solución de la ecuación

$$-\frac{d}{dx}\Big(a_1(x)\frac{d}{dx}u_1(x)\Big) = f(g(x))g'(x) = f_1(x).$$

Además, como $\xi_i = g(x_i)$ también se cumple que

$$\begin{aligned}
\Big(a_1(x)\frac{d}{dx}u_1(x)\Big)\Big|_{x=x_i^-} &= \Big(\frac{a(g(x))}{g'(x)}\frac{d}{d\xi}u(\xi)\Big|_{\xi=g(x)}g'(x)\Big)\Big|_{x=x_i^-} \\
&= \Big(a(g(x))\frac{d}{d\xi}u(\xi)\Big|_{\xi=g(x)}\Big)\Big|_{x=x_i^-} \\
&= \Big(a(g(x_i^-))\frac{d}{d\xi}u(\xi)\Big|_{\xi=g(x_i^-)}\Big) \\
&= \Big(a(\xi_i^-))\frac{d}{d\xi}u(\xi)\Big|_{\xi=\xi_i^-}\Big) \\
&= \beta_i[\![u(\xi)]\!]_{\xi_i} \\
&= \beta_i[\![u_1(x)]\!]_{x_i}.
\end{aligned}$$

De igual modo puede ser comprobado que

$$[\![a_1(x)\frac{du}{dx}]\!]_{x_i^-} = [\![a(\xi)\frac{du}{d\xi}]\!]_{\xi_k^-} = 0.$$

De este modo, ha sido demostrado que el problema (6)-(8) siempre puede ser reducido a un problema de la forma

$$-\frac{d}{dx}\Big(a_1(x)\frac{du_1}{dx}\Big) = f_1(x),\ x \in \Omega^* = (0,1) \setminus \{x_i = \frac{i}{l+1}\}, \tag{9}$$

$$a_1(x)\frac{du_1}{dx}\Big|_{x_i^-} = \beta_i[\![u_1(x)]\!]_{x_i}, \tag{10}$$

$$[\![a_1(x)\frac{du_1}{dx}]\!]_{x_i} = 0, \tag{11}$$

con las condiciones de frontera $u_1(0) = u(g(0)) = u(0) = t_0$ y $u_1(1) = t_1$, donde los coeficientes del nuevo problema satisfacen todas las condiciones de diferenciabilidad por tramos (y positividad en el caso de $a_1(x)$), que el problema original. Además, las discontinuidades de los coeficientes del nuevo problema conforman una red uniforme en el intervalo $(0,1)$, de paso $\frac{1}{l+1}$. Esto justifica que, sin perdida de generalidad, el análisis numérico se realice solo para problemas similares a (9)-(11).

Observación 1 *Una vez obtenida la solución $u_1(x)$ de (9)-(11), es posible recuperar la solución $u(\xi)$ de (6)-(8) a través de la relación $u(\xi) = u_1(g^{-1}(\xi))$, donde $g^{-1}(\xi)$ representa la función inversa de $g(x)$, la cual, como fue comentado con anterioridad, existe y al ser una función lineal por tramo, es bastante sencilla de determinar.*

Observación 2 *La simplificación realizada es equivalente a considerar en el problema original, en cada intervalo (ξ_i, ξ_{i+1}), un paso distinto $h_i = \frac{\xi_{i+1}-\xi_i}{n}$, $n \in \mathbb{N}$ al elegir la malla.*

1.3 Selección de la malla y obtención del esquema en diferencias finitas

Sean $\{x_i\}_{i=1}^l$ los puntos de discontinuidad de los coeficientes del problema (9)-(11), los cuales conforman una red uniforme de paso $\frac{1}{l+1}$. Resulta conveniente entonces elegir una malla uniforme que contenga a los puntos x_i entre sus elementos. Para ello basta con elegir un paso $h = \frac{1}{n(l+1)}$, con $n \in \mathbb{N}$. De modo general, la expresión para los elementos de la malla es

$$\eta_j = \frac{j}{n(l+1)},$$

así, cuando $j = kn$, $k = 1, \ldots, l$, entonces η_j será el k-ésimo punto de discontinuidad. Como en dichos puntos la solución tendrá valores distintos a la derecha y a la izquierda, es conveniente entonces, para estos valores de j, definir 2 nodos, en vez de 1, denotados como η_j^- y η_j^+. Esta selección de h, junto con las consideraciones adicionales hechas para los puntos de discontinuidad, dará como resultado una cantidad de nodos igual a $m = n(l+1) + l - 1$ (sin incluir los extremos del intervalo $\eta_0 = 0$ y $\eta_{n(l-1)} = 1$). Estos nodos en lo adelante serán llamados nodos principales. Para simplificar la notación, en lo adelante se considerará $u_j = u_1(\eta_j)$, si $j \neq kn$, $u_j^- = u_1(\eta_j^-)$ y $u_j^+ = u_1(\eta_j^+)$ si $j = kn$ para $k = 1, \ldots, l$. Para abordar numéricamente el problema se definirán además los

nodos auxiliares $\eta_{j+1/2}$, como los puntos medios de los nodos definidos con anterioridad, es decir

$$\eta_{j+1/2} = \frac{\eta_j + \eta_{j+1}}{2}, j = 0, \ldots, n(l+1) - 1.$$

Estos nodos son meramente simbólicos, ya que no estarán incluidos en las ecuaciones que serán obtenidas. Su inclusión se debe a que servirán de apoyo para calcular las aproximaciones de las derivadas.
La ecuación (9) puede ser reescrita como

$$-(J(u_1)(x))' = f_1(x), x \in \Omega^*,$$

donde $J(u_1)(x)$ es el flujo asociado a $u_1(x)$, definido como

$$J(u_1)(x) = a_1(x)u_1'(x).$$

Para abreviar, en lo adelante se considerará

$$J(u_1)(\eta_{j+1/2}) = J_{j+1/2}.$$

Esto permite aproximar la derivada del flujo en los puntos η_j, si $j \neq kn$, a través de la fórmula en diferencias centrales para la derivada:

$$u'(x) = \frac{u(x+h) - u(x-h)}{2h} + \frac{h^2}{12}(u'''(x+\theta_2) - u'''(x-\theta_1)), \tag{12}$$

con $0 < \theta_1, \theta_2 < h$, y con las fórmulas en diferencias hacia adelante y hacia atrás para las derivadas en los puntos η_j^- y η_j^+ respectivamente, cuando $j = kn$, $k = 1, \ldots, l$:

$$u'(x^-) = \frac{u(x^-) - u(x-h)}{h} - \frac{h}{2}u''(x-\theta_1), \tag{13}$$

$$u'(x^+) = \frac{u(x+h) - u(x^+)}{h} + \frac{h}{2}u''(x+\theta_2). \tag{14}$$

De este modo

$$-(J(u_1)(\eta_j))' \approx -\frac{J_{j+1/2} - J_{j-1/2}}{h}, \text{ si } j \neq kn, \tag{15}$$

$$-(J(u_1)(\eta_j^-))' \approx -\frac{J(u_1)(\eta_j^-) - J_{j-1/2}}{h/2}, \text{ si } j = kn, \tag{16}$$

$$-(J(u_1)(\eta_j^+))' \approx -\frac{J_{j+1/2} - J(u_1)(\eta_j^+)}{h/2}, \text{ si } j = kn. \tag{17}$$

Utilizando nuevamente la expresión (12) para aproximar el flujo $J_{j+1/2}$, y las condiciones (10) y (11) para $J(u_1)(\eta_j^-)$ y $J(u_1)(\eta_j^+)$ respectivamente se tiene que

$$J_{j+1/2} = a_1(\eta_{j+1/2})u_1'(\eta_{j+1/2}) \approx a_{j+1/2}\frac{u_{j+1} - u_j}{h}, \tag{18}$$

$$J(u_1)(\eta_j^-) = J(u_1)(\eta_j^+) = a_1(\eta_j^-)u_1'(\eta_j^-) = \beta_k(u_j^+ - u_j^-). \tag{19}$$

Sustituyendo (18) y (19) en (15), (16) y (17) se llega a las ecuaciones lineales buscadas

$$h^{-2}(-a_{j-\frac{1}{2}}u_{j-1} + (a_{j-\frac{1}{2}} + a_{j+\frac{1}{2}})u_j - a_{j+\frac{1}{2}}u_{j+1}) = f(\eta_j), \tag{20}$$

$$-\frac{a_{j-\frac{1}{2}}}{h^2}u_{j-1} + (\frac{a_{j-\frac{1}{2}}}{h^2} + \frac{\beta_k}{h})u_j^- - \frac{\beta_k}{h}u_j^+ = f(\eta_j^-)/2, \tag{21}$$

$$-\frac{\beta_k}{h}u_j^- + (\frac{a_{j+\frac{1}{2}}}{h^2} + \frac{\beta_k}{h})u_j^+ - \frac{a_{j+\frac{1}{2}}}{h^2}u_{j+1} = f(\eta_j^+)/2. \tag{22}$$

La matriz asociada a este sistema es tridiagonal, simétrica y en la próxima sección será visto que también es definida positiva. Nótese que los nodos auxiliares introducidos solo aparecen para evaluar la función $a_1(x)$, la cual es conocida. Las componentes del vector independiente $b^T = (b_1, \ldots, b_{n(l+1)})$ son

$$\begin{aligned} b_1 &= f(\eta_1) + h^{-2}a_{1/2}t_0, \\ b_j &= f(\eta_j), \text{ si } j \neq kn, \\ b_j^- &= f(\eta_j^-)/2, \text{ si } j = kn, \\ b_j^+ &= f(\eta_j^+)/2, \text{ si } j = kn, \\ b_{n(l+1)} &= f(\eta_{n(l+1)}) + h^{-2}a_{n(l+1)+1/2}t_1. \end{aligned}$$

1.4 Un ejemplo

A continuación serán aplicados los resultados de la subsección anterior a un ejemplo específico. Considérese la función $a(x)$ definida como

$$a(x) = \begin{cases} a_1 & \text{si} \quad 0 < x < 1/2, \\ a_2 & \text{si} \quad 1/2 < x < 1, \end{cases}$$

con $a_1, a_2 > 0$. El problema a analizar numéricamente es

$$\begin{aligned} -\frac{d}{dx}\left(a(x)\frac{du}{dx}\right) &= 0, \ x \in (0,1) \setminus \{1/2\}, \\ a(x)\frac{du}{dx}\Big|_{(1/2)^-} &= \beta[\![u(x)]\!]_{1/2}, \\ [\![a(x)\frac{du}{dx}]\!]_{1/2} &= 0, \end{aligned}$$

donde $\beta > 0$. Además se considerarán las condiciones de frontera $u(0) = t_0$ y $u(1) = t_1$. La malla utilizada para dar solución al problema tendrá paso $h = 1/4$. De este modo, dicha malla estará conformada por los valores $(0, 1/4, (1/2)^-, (1/2)^+, 3/4, 1)$. Los nodos auxiliares son los puntos

$$\eta_{1/2} = 1/8, \ \eta_{3/2} = 3/8, \ \eta_{5/2} = 5/8, \ \eta_{7/2} = 7/8,$$

de modo que

$$\begin{aligned} a_{1/2} &= a(1/8) = a_1, \\ a_{3/2} &= a(3/8) = a_1, \\ a_{5/2} &= a(5/8) = a_2, \\ a_{7/2} &= a(7/8) = a_2. \end{aligned}$$

De acuerdo a esta información, el sistema de ecuaciones lineales derivado de aplicar el método de las diferencias finitas es

$$\begin{pmatrix} 8a_1 & -4a_1 & 0 & 0 \\ -4a_1 & 4a_1+\beta & -\beta & 0 \\ 0 & -\beta & 4a_2+\beta & -4a_2 \\ 0 & 0 & -4a_2 & 8a_2 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{pmatrix} = \begin{pmatrix} 4a_1t_0 \\ 0 \\ 0 \\ 4a_2t_1 \end{pmatrix}$$

Nótese que la primera y la última ecuación podrían ser divididas entre $4a_1$ y $4a_2$ respectivamente, pero no se ha realizado esta simplificación para conservar la simetría de la matriz. Los cálculos necesarios para dar solución a este sistema son bastante extensos, no obstante conducen a la solución

$$\begin{aligned} u_1 &= \frac{c}{4a_1} + t_0, \\ u_2 &= \frac{c}{2a_1} + t_0, \\ u_3 &= \frac{c}{2a_2} + c\Big(\frac{1}{\beta} + (\frac{1}{a_1} - \frac{1}{a_2})\frac{1}{2}\Big) + t_0, \\ u_4 &= \frac{3c}{4a_2} + c\Big(\frac{1}{\beta} + (\frac{1}{a_1} - \frac{1}{a_2})\frac{1}{2}\Big) + t_0, \end{aligned}$$

donde

$$c = \frac{t_1 - t_0}{\frac{1}{2}(a_1^{-1} + a_2^{-1}) + \beta^{-1}}.$$

Aun más, es posible comprobar que estos valores coinciden con los de la solución exacta, la cual es

$$u(x) = \begin{cases} \frac{c}{a_1}x + t_0 & \text{si} \quad 0 < x < 1/2, \\ \frac{c}{a_2}x + c\Big(\frac{1}{\beta} + (\frac{1}{a_1} - \frac{1}{a_2})\frac{1}{2}\Big) + t_0 & \text{si} \quad 1/2 < x < 1. \end{cases}$$

La coincidencia se debe a que la solución exacta está conformada por rectas, por lo que las aproximaciones para sus derivadas coinciden con las derivadas reales, haciendo que todos los valores obtenidos sean también exactos.

2. Consistencia, estabilidad y convergencia

En lo adelante, el problema (9)-(11) será representado como

$$L(u_1, f_1) = -(J(u_1)(x))' - f_1(x) = 0,$$

mientras que el problema obtenido por el método de las diferencias finitas, para un determinado valor de h será denotado como

$$L_h(u_h, f_h) = -A_{df}u_h - f_h = 0, \tag{23}$$

donde $-A_{df}$ es la matriz del sistema y f_h es el vector independiente. Esta nueva notación ayudará en el análisis de la consistencia y la estabilidad del método de las diferencias finitas aplicado al problema de valores en la frontera con condiciones de discontinuidad.

2.1 Estabilidad

En el segundo capítulo de [1] es dado el siguiente concepto de *estabilidad*

Definición 1 *Un método numérico está bien planteado, (o es estable) si para cualquier h, existe una única solución, dependiente continuamente de los datos.*

Según esta definición, para demostrar la estabilidad del método aquí utilizado, es necesario demostrar que la solución u_h de (23) es única, y además que depende continuamente de los datos del problema. Nótese que la matriz del sistema, $-A_{df}$, es tridiagonal, simétrica, $a_{ii} > 0$ para todo i,

$$a_{i(i-1)} = a_{(i-1)i} < 0$$

para todo $i \geqslant 2$ y $a_{(i+1)i} + a_{ii} + a_{i(i+1)} = 0$ para $i = 2, \ldots, m-1$, donde $m = n(l+1) + l - 1$ es la dimensión de la matriz $-A_{df}$. Además $a_{11} + a_{12} > 0$ y $a_{m(m-1)} + a_{mm} > 0$. Será demostrado a continuación que una matriz con estas condiciones, es definida positiva, es decir,

$$(v_h, -A_{df}v_h) > 0, \ \forall v_h \neq 0,$$

donde $(\cdot,\cdot)$ es el producto escalar usual en $\mathbb{R}^m$ y v_h una función definida sobre los nodos de la malla del problema. En efecto, al ser A_{df} tridiagonal y simétrica, entonces

$$(v_h, -A_{df}v_h) = \sum_{i=1}^{m} a_{ii}v_i^2 + \sum_{i=1}^{m-1} 2a_{i(i+1)}v_iv_{i+1}. \tag{24}$$

Utilizando la propiedad $a_{(i+1)i} + a_{ii} + a_{i(i+1)} = 0$ para $i = 2, \ldots, m-1$, el miembro derecho de (24) puede reescribirse como

$$\begin{aligned} &(v_h, -A_{df}v_h) \\ &\quad = (a_{11} + a_{12})v_1^2 + (a_{m(m-1)} + a_{mm})v_m^2 \\ &\qquad + \sum_{i=1}^{m-1} (-a_{i(i+1)}v_i^2 + 2a_{i(i+1)}v_iv_{i+1} - a_{i(i+1)}x_{i+1}^2) \\ &\quad = (a_{11} + a_{12})v_1^2 + (a_{m(m-1)} + a_{mm})v_m^2 \\ &\qquad - \sum_{i=1}^{m-1} a_{i(i+1)}(v_i - v_{i+1})^2. \end{aligned} \tag{25}$$

Finalmente, usando las condiciones $a_{i(i+1)} < 0$, $a_{11} + a_{12} > 0$ y $a_{m(m-1)} + a_{mm} > 0$ se puede concluir que $-A_{df}$ es definida positiva. Esto implica que es inversible, por lo tanto la solución de (23) existe y es única para cada valor de h. Además, sustituyendo los valores de a_{11}, a_{mm} y $a_{i(i+1)}$ en (25) se obtiene la siguiente expresión

$$\begin{aligned} (v_h, -A_{df}v_h) = &h^{-2}a_{1/2}v_1^2 + h^{-2}a_{m+1/2}v_m^2 \\ &+ h^{-2}\sum_{i\in I}^{m-1} a_{i+1/2}(v_i - v_{i+1})^2 \\ &+ h^{-1}\sum_{k=1}^{l} \beta_k(v_{kn}^+ - v_{kn}^-)^2, \end{aligned} \tag{26}$$

donde $I = \{i : i \neq kn \wedge x_i = x_{nk}^+\}$. La expresión (26) permite definir una norma en V_h, el espacio de las funciones definidas sobre los nodos de la malla. Dicha norma será denotada en lo adelante como $\|\cdot\|_A$ y viene dada por la expresión

$$\|v_h\|_A^2 = (v_h, -A_{df}v_h).$$

Considérese además la siguiente norma en V_h

$$\|v_h\|_h^2 = (v_h, v_h) = \sum_{i=1}^{m} v_i^2.$$

Entonces existe una constante c, independiente de h, tal que

$$\|v_h\|_h \leqslant c\|v_h\|_A, \ \ \forall v_h \in V_h.$$

En efecto, teniendo en cuenta que $a_{j+1/2} = a_1(\eta_{j+1/2})$, que $a_1(x)$ es mayor que una constante positiva α_0 para cualquier valor de x, que las constantes β_k son positivas y por lo tanto tienen un mínimo β_0, entonces, denotando $c_0 = (\text{mín}\{\alpha_0, \beta_0\})^{1/2}$, se tiene que, para cualquier valor de i

$$\begin{aligned} c_0 v_j =& c_0\Big(v_1 + \sum_{i=1}^{j-1}(v_{i+1} - v_i)\Big) \\ =& c_0\Big(v_1 + \sum_{i\in I'}(v_{i+1} - v_i) + \sum_{i\in K'}(v_i^+ - v_i^-)\Big), \end{aligned}$$

donde $I' = \{i \in I, i \leqslant j-1\}$ y $K' = \{i = kn, \leqslant j-1\}$. Elevando al cuadrado ambos miembros y utilizando la desigualdad de Minkowski

$$\Big(\sum_{i=1}^{d} p_i\Big)^2 \leqslant d\sum_{i=1}^{d} p_i^2,$$

la cual se cumple para cualquier entero $d \geqslant 1$ y cualquier suceción de números reales $\{p_i\}_{i=1}^d$, se tiene que

$$\begin{aligned} & c_0^2 v_j^2 \\ & \leqslant 3\Big(c_0^2 v_1^2 + j\sum_{i\in I'} c_0^2 (v_{i+1} - v_i)^2 + l\sum_{k=1}^{l}\beta_k (v_{kn}^+ - v_{kn}^-)^2\Big) \\ & \leqslant 3\Big(a_{1/2}v_1^2 + j\sum_{i\in I} a_{i+1/2}(v_{i+1} - v_i)^2 + l\sum_{k=1}^{l}\beta_k (v_{kn}^+ - v_{kn}^-)^2\Big). \end{aligned}$$

Sumando desde $j = 1$ hasta m

$$\begin{aligned} c_0^2\sum_{j=1}^{m} v_j^2 \leqslant& 3\Big(ma_{1/2}v_1^2 + \sum_{i=1}^{m} j\sum_{i\in I} a_{i+1/2}(v_i - v_{i+1})^2 \\ & + ml\sum_{k=1}^{l}\beta_k (v_{kn}^+ - v_{kn}^-)^2\Big) \\ =& 3\Big(ma_{1/2}v_1^2 + \frac{m(m+1)}{2}\sum_{i\in I} a_{i+1/2}(v_i - v_{i+1})^2 \\ & + ml\sum_{k=1}^{l}\beta_k (v_{kn}^+ - v_{kn}^-)^2\Big). \end{aligned}$$

Teniendo en cuenta que $m = n(l+1) + l - 1$ y que $h = \frac{1}{n(l+1)}$ y por lo tanto $hm = 1 + \frac{l-1}{n(l+1)} \leqslant 2$, para todo $n \in \mathbb{N}$

$$\begin{aligned} & c_0^2\sum_{j=1}^{m} v_j^2 \\ & \leqslant 12l\Big(h^{-1}a_{1/2}v_1^2 + h^{-2}\sum_{i\in I} a_{i+1/2}(v_i - v_{i+1})^2 \\ & \quad + h^{-1}\sum_{k=1}^{l}\beta_k (v_{kn}^+ - v_{kn}^-)^2\Big) \\ & \leqslant 12l\Big(h^{-2}a_{1/2}v_1^2 + h^{-2}a_{m+1/2}v_m^2 \\ & \quad + h^{-2}\sum_{i\in I} a_{i+1/2}(v_i - v_{i+1})^2 + h^{-1}\sum_{k=1}^{l}\beta_k (v_{kn}^+ - v_{kn}^-)^2\Big). \end{aligned}$$

Consecuentemente, para toda $v_h \in V_h$

$$\|v_h\|_h^2 \leqslant c\|v_h\|_A^2, \tag{27}$$

con $c = \frac{12l}{\text{mín}\{\alpha_0, \beta_0\}}$, independiente de h. Finalmente, para demostrar la estabilidad del método, basta con tomar la ecuación

$$-A_{df}u_h = f_h,$$

y multiplicarla escalarmente por u_h. De este modo se obtiene la relación

$$\|u_h\|_A^2 = (u_h, -A_{df}u_h) = (f_h, u_h).$$

Utilizando la desigualdad (27) y la desigualdad de Cauchy-Schwarz, se obtiene

$$\|u_h\|_h^2 \leqslant c\|f_h\|_h\|u_h\|_h,$$

y luego de simplificar un factor $\|u_h\|_h$

$$\|u_h\|_h \leqslant c\|f_h\|_h,$$

por lo tanto el método es estable, ya que la solución además de ser única, depende continuamente de los datos del problema.

2.2 Consistencia

Para demostrar la consistencia del método, es necesario probar que

$$L_h(u_1, f_1) = L_h(u_1, f_1) - L(u_1, f_1) \longrightarrow 0, \text{ si } h \to 0.$$

A partir de la aproximación (12), se tiene que

$$\begin{aligned} u_1'(x) =& \frac{u_1(x+h/2) - u_1(x-h/2)}{h} \\ & + \frac{h^2}{48}(u_1'''(x+\theta_2) + u_1'''(x-\theta_1)) \end{aligned}$$

con $0 < \theta_1, \theta_2 < h/2$, por lo tanto, la acotación de la solución exacta del problema y de sus derivadas implican que

$$\frac{u_1(x+h/2) - u_1(x-h/2)}{h} = u_1'(x) + h^2 R^1,$$

con $|R^1| \leqslant \frac{1}{24}\sup_{[0,1]}|u_1'''(x)|$. Sea $j \neq kn$, entonces

$$\begin{aligned}
&L_h(u_1,f_1)(\eta_j)\\
&=-\frac{a_{j+\frac{1}{2}}\frac{u_1(\eta_{j+1})-u_1(\eta_j)}{h}-a_{j-\frac{1}{2}}\frac{u_1(\eta_j)-u_1(\eta_{j-1})}{h}}{h}-f(\eta_j)\\
&=-\frac{a_{j+\frac{1}{2}}u_1'(\eta_{j+\frac{1}{2}})-a_{j-\frac{1}{2}}u_1'(\eta_{j-\frac{1}{2}})}{h}-f(\eta_j)\\
&\quad-h(a_{j+\frac{1}{2}}R^1_{j+1}-a_{j-\frac{1}{2}}R^1_j)\\
&=-\frac{J(u_1)(\eta_{j+\frac{1}{2}})-J(u_1)(\eta_{j-\frac{1}{2}})}{h}-f(\eta_j)\\
&\quad-h(a_{j+\frac{1}{2}}R^1_{j+1}-a_{j-\frac{1}{2}}R^1_j).
\end{aligned}$$

Utilizando nuevamente la aproximación (12), esta vez para el flujo

$$\begin{aligned}
&L_h(u_1,f_1)(\eta_j)\\
&=-\frac{J(u_1)(\eta_{j+\frac{1}{2}})-J(u_1)(\eta_{j-\frac{1}{2}})}{h}-f(\eta_j)\\
&\quad-h(a_{j+\frac{1}{2}}R^1_{j+1}-a_{j-\frac{1}{2}}R^1_j)\\
&=-\Big(J(u_1)(\eta_j)\Big)'-f(\eta_j)-h^2R^2_j-h(a_{j+\frac{1}{2}}R^1_{j+1}-a_{j-\frac{1}{2}}R^1_j)\\
&=-h^2R^2_j-h(a_{j+\frac{1}{2}}R^1_{j+1}-a_{j-\frac{1}{2}}R^1_j),
\end{aligned}$$

con

$$|R^2_j| \leqslant \frac{1}{24}\sup_{[0,1]}|\Big(J(u_1)\Big)'''| = \frac{1}{24}\sup_{[0,1]}|f''|,$$

de modo que

$$|L_h(u_1,f_1)(\eta_j)| \leqslant \frac{h^2}{24}\sup_{[0,1]}|f''| + \frac{h\alpha_1}{12}\sup_{[0,1]}|u_1'''(x)|,$$

o equivalentemente, para h suficientemente pequeño

$$|L_h(u_1,f_1)(\eta_j)| \leqslant \frac{h\alpha_1}{6}\sup_{[0,1]}|u_1'''(x)|,$$

donde α_1 es el supremo de la función $a(x)$.
De igual manera, las aproximaciones (13) y (14) implican que

$$\begin{aligned}
\frac{u_1(x^-)-u_1(x-h/2)}{h/2} &= u_1'(x^-)+hR^3,\\
\frac{u_1(x+h/2)-u_1(x^+)}{h/2} &= u_1'(x^+)-hR^3.
\end{aligned}$$

con

$$|R^3| \leqslant \frac{1}{4}\sup_{[0,1]}|u_1''(x)|.$$

Luego, para $j = kn$

$$\begin{aligned}
&2L_h(u_1,f_1)(\eta_j^-)\\
&=-\frac{\beta_k(u_1(\eta_j^+)-u_1(\eta_j^-))-a_{j-1/2}\frac{u_1(\eta_j^-)-u_1(\eta_{j-1})}{h}}{h/2}-f(\eta_j^-)\\
&=-\frac{J(u_1)(\eta_j^-)-J(u_1)(\eta_{j-1/2})-a_{j-1/2}h^2R^1_j}{h/2}-f(\eta_j^-)\\
&=-\frac{J(u_1)(\eta_j^-)-J(u_1)(\eta_{j-1/2})}{h/2}-f(\eta_j^-)+2a_{j-1/2}hR^1_j\\
&=-\Big(J(u_1)(\eta_j^-)\Big)'-f(\eta_j^-)-2hR^3_j-2a_{j-1/2}hR^1_j\\
&=-2hR^3_j-2a_{j-1/2}hR^1_j.
\end{aligned}$$

Una aproximación similar puede hacerse para $L_h(u_1,f_1)(\eta_j^+)$, asegurando que

$$\begin{aligned}
|L_h(u_1,f_1)(\eta_j^-)| &\leqslant h(\frac{1}{4}\sup_{[0,1]}|u_1''(x)|+\frac{\alpha_1}{24}\sup_{[0,1]}|f''|),\\
|L_h(u_1,f_1)(\eta_j^+)| &\leqslant h(\frac{1}{4}\sup_{[0,1]}|u_1''(x)|+\frac{\alpha_1}{24}\sup_{[0,1]}|f''|).
\end{aligned}$$

Por lo tanto, se ha demostrado el resultado requerido, ya que, tomando la norma $\|\cdot\|_h$ definida en el análisis de la estabilidad

$$\|v_h\|_h^2 = \sum_{i=1}^m v_i^2,$$

se tiene que

$$\begin{aligned}
&\|L_h(u_1,f_1)\|_h^2\\
&=\sum_{i=1}^m (L_h(u_1,f_1)(\eta_i))^2\\
&=\sum_{j\neq nk}\Big(L_h(u_1,f_1)(\eta_j)\Big)^2\\
&\quad+\sum_{j=nk}\Big(L_h(u_1,f_1)(\eta_j^-)\Big)^2+\Big(L_h(u_1,f_1)(\eta_j^+)\Big)^2\\
&\leqslant m\frac{h^2\alpha_1^2}{36}\sup_{[0,1]}|u_1'''(x)|^2\\
&\quad+4lh^2(\frac{1}{16}\sup_{[0,1]}|u_1''(x)|^2+\frac{\alpha_1^2}{576}\sup_{[0,1]}|f''|^2).
\end{aligned}$$

Nuevamente, teniendo en cuenta que $mh \leqslant 2$,

$$\begin{aligned}
\|L_h(u_1,f_1)\|_h^2 \leqslant &\frac{h\alpha_1^2}{18}\sup_{[0,1]}|u_1'''(x)|^2\\
&+4lh^2\Big(\frac{1}{16}\sup_{[0,1]}|u_1''(x)|^2+\frac{\alpha_1^2}{576}\sup_{[0,1]}|f''|^2\Big),
\end{aligned}$$

o equivalentemente para h suficientemente pequeño

$$\|L_h(u_1,f_1)\|_h^2 \leqslant \frac{h\alpha_1^2}{9}\sup_{[0,1]}|u_1'''(x)|^2.$$

Por lo tanto, el método es consistente, ya que.

$$\|L_h(u_1,f_1)\|_h \longrightarrow 0 \text{ si } h\to 0.$$

2.3 Convergencia

Introduciendo la función de error global e_h, definida sobre los nodos de la malla como

$$e(\eta_j) = u_1(\eta_j) - u_h(\eta_j),$$

se puede apreciar que dicha función satisface las relaciones

$$-A_{df}e_h = -A_{df}u_1 - (-A_{df}u_h) = -A_{df}u_1 - f_h = L_h(u_1, f_1).$$

Teniendo en cuenta que el método es estable, la función e_h, también cumple que

$$\|e_h\|_h \leqslant c\|L_h(u_1, f_1)\|_h.$$

Finalmente la consistencia del método y la independencia de h de la constante c implican

$$\|e_h\|_h^2 \leqslant c^2 \frac{h\alpha_1^2}{9} \sup_{[0,1]} |u_1'''(x)|^2 \longrightarrow 0 \text{ si } h \to 0.$$

Concluyendo, la solución u_h converge hacia la solución exacta u_1 cuando h tiende a cero, siempre y cuando esta última sea al menos tres veces continuamente diferenciable por tramos en el intervalo $(0,1)$.

3. Algunos ejemplos de la aplicación del método

Considérese la función

$$a(x) = \begin{cases} 1 & \text{si} \quad 0 < x < 1/8, \\ x^2 & \text{si} \quad 1/8 < x < 3/5, \\ x^3 & \text{si} \quad 3/5 < x < 1. \end{cases}$$

Entonces, el problema

$$-\frac{d}{dx}\left(a(x)\frac{du}{dx}\right) = 0,\ x \in \Omega^* = (0,1) \setminus \{1/8, 3/5\}, \tag{28}$$

$$a(x)\frac{du}{dx}\Big|_{x_i^-} = [\![u(x)]\!]_{x_i}, \tag{29}$$

$$[\![a(x)\frac{du}{dx}]\!]_{x_i} = 0, \tag{30}$$

con las condiciones de frontera $u(0) = 1$ y $u(1) = 2$, posee como solución exacta a

$$u(x) = \begin{cases} \frac{72}{673}x + 1 & \text{si} \quad 0 < x < 1/8, \\ -\frac{72}{673}\frac{1}{x} + \frac{1330}{673} & \text{si} \quad 1/8 < x < 3/5, \\ -\frac{36}{673}\frac{1}{x^2} + \frac{1382}{673} & \text{si} \quad 3/5 < x < 1. \end{cases}$$

La Figura 1 muestra el gráfico de dicha solución.

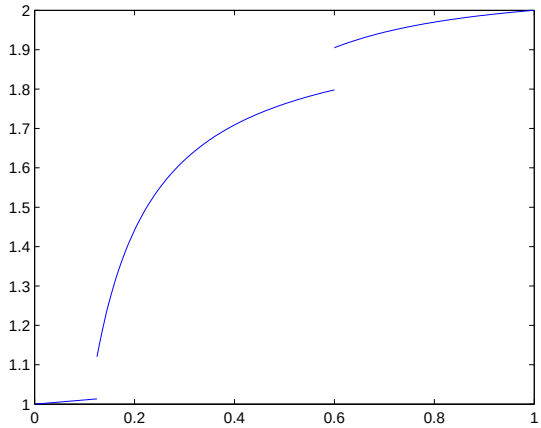

Figura 1. Gráfico de la solución exacta del problema (28)-(30).

Por otro lado, la Figura 2 muestra la comparación de $u(x)$ (en rojo) con las soluciones numéricas (en azul para $n = 2$, $n = 5$, $n = 10$ y $n = 100$, es decir, para un tamaño de paso, en el problema simplificado, $h = 0{,}25$, $h = 0{,}1$, $h = 0{,}05$ y $h = 0{,}005$.

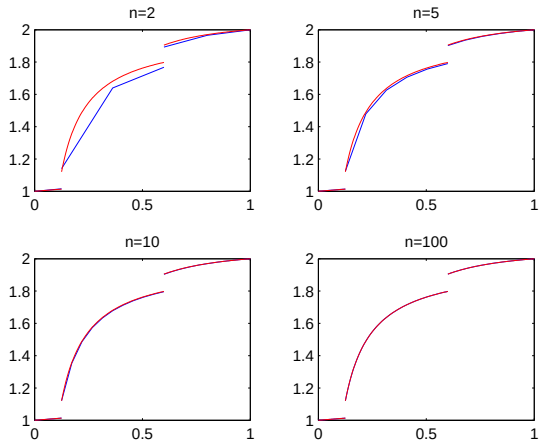

Figura 2. Comparación entre las soluciones numéricas para distintos valores de n y la solución exacta.

La convergencia de las soluciones numéricas hacia la solución exacta a medida que crece n es evidente.
Considérese ahora el mismo problema (28)-(30), pero con $f(x) = \sin x + x\cos x = (x\sin x)'$. En este caso en la fórmula de la solución general dada en la introducción, aparecerán términos del tipo

$$\int_{1/8}^{x} \frac{\sin x}{x},$$

que no pueden ser expresados en términos de funciones elementales. La Figura 3 muestra las distintas aproximaciones de la solución para este nuevo problema, a partir de distintos valores de n.

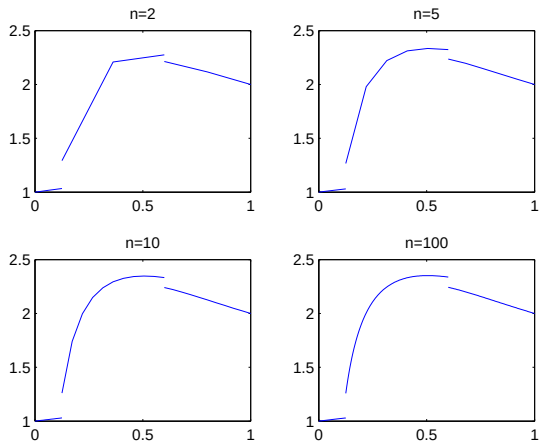

Figura 3. Soluciones numéricas para distintos valores de n del problema (28)-(30) modificado.

4. Conclusiones

El problema de conducción del calor con condiciones de contacto resulta de especial interés en la teoría de la homogeneización, ya sea simple o reiterada, donde aparecen problemas de este tipo, que al depender de un parámetro rapídamente oscilante, dificulta su tratamiento numérico. Sin embargo, durante el proceso de homogeneización aparecen los llamados problemas locales, además del problema homogeneizado, los cuales presentan una estructura similar al problema original, pero son independientes del parámetro pequeño, por lo tanto si resulta conveniente tratarlos numéricamente.

En el presente informe, luego de simplificar el problema original para que cumpliera determinadas condiciones, fue aplicado el método de las diferencias finitas a la ya mencionada ecuación, provando finalmente, la estabilidad y consistencia del método y por lo tanto la convergencia de las soluciones numéricas a la solución exacta. las funciones de MATLAB implementadas fueron discutidas e incluidas en el anexo y algunos ejemplos fueron discutidos.

Otra método numérico factible para abordar la ecuación aquí estudiada es el método de elementos finitos, puesto que es posible encontrar una formulación variacional para el problema de conducción del calor con condiciones de discontinuidad.

Agradecimientos

Los autores desean agradecer al Proyecto Nacional de Ciencias Básicas de Cuba PNCB-UH-57/14 “Modelación matemática de fenómenos dependientes de varias escalas microestructurales”; y a la Cátedra IIMAS-UNAM 2016.

Referencias

[1] A. Quarteroni, R. Sacco, F. Saleri, Numerical Mathematics, Springer-Verlag New York, ISBN 0-387-98959-5, 2007.

[2] L. Beilina, Hybrid Discontinuous Finite Element/Finite Difference Method for Maxwell’s Equations, AIP Conference Proceedings 1281, 324 (2010);

[3] De-Kang Mao, A Treatment of Discontinuities for Finite Differences Methods, Journal of Computational Physics 103,359-369 (1992)

[4] De-Kang Mao, A Treatment of Discontinuities for Finite Differences Methods in the Two-Dimensional Case, Journal of Computational Physics 104,377-397 (1993)

[5] Jirásek M., Comparative study on fnite elements with embedded discontinuities, Computer Methods in Applied Mechanics and Engineering, 188(1-3), pp. 307-330, 2000.

[6] Juarez-Luna G., Ayala A., Elmentos Finitos con Discontinuidades Interiores Mejorados para Simular el Daño en Sólidos, Concreto y Cemento. Investigación y Desarrollo, Vol. 6 Núm. 1, 15-35, 2014.

Vol. 27, 1976

[7] S. Larsson and V. Thomhée, Partial Differential Equations with Numerical Methods, Springer-Verlag Berlin Heidelberg, 2009.