

Obligatorio 2 - Aplicación de Diferencias Finitas de la Ecuación de Calor Unidimensional

Grupo 36

Guillermo Daniel Toyos Marfurt (5139879-9),
Juan José Mangado Esteves (5535227-0),
Martín Sebastián Piñeyro Olivera (5043446-7),
Ricardo Alejandro Fasanello Guillermo (4763979-9)

Tutor: Juan Piccini

Métodos Numéricos 2020

Instituto de Matemática y Estadística

Facultad de Ingeniería. Universidad de la República

Montevideo, Uruguay

Resumen

Se estudia el *problema de Cauchy Dirichlet para la ecuación del calor no homogénea* y con condición de borde nula en el caso unidimensional utilizando el método de integración θ -método. Para ello se presenta una solución discreta al problema y se demuestra su correctitud de forma analítica junto a su implementación. Adicionalmente se encuentra un criterio de estabilidad asintótica para el caso particular donde $\theta = 0$ y el sistema no tiene una fuente de generación calorífica. Lo cual asegura la convergencia del problema si la distancia finita entre cada instante del tiempo es menor el cuadrado de la diferencia entre cada punto del espacio sobre una constante. También se estudia experimentalmente el comportamiento del θ -método para distintos valores de θ y se presentan representaciones gráficas de la evolución temporal de la temperatura a lo largo de la barra junto a su diferencia con el comando *lsode* de *Octave* observándose una mejor aproximación con $\theta = \frac{1}{2}$. El informe finaliza con un análisis del error del θ -método en comparación con el comando *lsode* mostrando que la solución presentada genera resultados similares a los de *lsode* y su diferencia se puede reducir arbitrariamente reduciendo la distancia que se toma entre cada punto discreto del tiempo.

Keywords: Ecuaciones diferenciales en derivadas parciales, Ecuación del calor, Métodos de integración, θ -método

Índice

1. Introducción

Imagínese, lector, que usted se encuentra en un caluroso desierto y tiene una ardiente y fina barra de metal en sus manos. Ahora usted coloca en sus extremos dos cubos de hielo. ¿Que sucede? Este modelo físico puede modelarse mediante la siguiente ecuación en derivadas parciales conocida como la *ecuación del calor no homogénea*:

$$\frac{\partial u}{\partial t}(x, t) - \mu \frac{\partial^2 u}{\partial x^2}(x, t) = f(x, t)$$

La variable x representa la posición unidimensional en un segmento, t el tiempo y $u(x, t)$ la temperatura de la barra en el punto x e instante t . Supondremos que la barra es lo suficientemente fina y trataremos el problema en una dimensión, i.e. $x \in \mathbb{R}$. La función $f(x, t)$ modela una fuente de calor que calienta la barra uniformemente (el calor del desierto) y la constante μ es un coeficiente inherente al material. Como inicialmente la barra se encontraba caliente, tendremos una función $u_0(x)$ que modela la temperatura de la barra en el tiempo inicial del problema ($u(x, 0) = u_0(x)$). Por último, los cubos de hielo en los extremos de la barra pueden modelarse como una condición que debe cumplir u : En ambos extremos de la barra, la temperatura es 0 en todo instante. A este problema se lo conoce como *problema de Cauchy Dirichlet para la ecuación del calor no homogénea con contorno de Dirichlet homogéneo*.

Si bien es posible resolver el problema mencionado analíticamente, en éste informe trataremos de hallar una aproximación de $u(x, t)$ utilizando una cierta familia de métodos de integración conocida como θ -métodos.

El informe realiza los siguientes aportes:

- Se construye paso a paso una aproximación a la ecuación del calor en diferencias finitas y un esquema de implementación.
- Demostración de distintas propiedades de la solución planteada: Orden de los errores, existencia de solución y caracterización del método planteado.
- Obtención de un criterio de estabilidad asintótica para el caso $\theta = 0, f(x, t) = 0$.
- Un estudio experimental con distintas simulaciones para confirmar la correctitud del algoritmo comparándolo con el comando “*lsode*” de *Octave*.
- Comparación del error de la solución propuesta con el comando “*lsode*” para múltiples casos en una posición y tiempo dado.

El informe se organiza de la siguiente manera. La sección (??) trata sobre la metodología, planteo y análisis de la solución construida. La sección (??) presenta los resultados del

análisis experimental. Finalmente, la sección (??) presenta las conclusiones del trabajo.

2. Problema a resolver

Dado un intervalo $[a, b] \subset \mathbb{R}$, una constante $\mu > 0$, una función $u_0 : [a, b] \rightarrow \mathbb{R}$ y la función $u : [a, b] \times \mathbb{R} \rightarrow \mathbb{R}$ regida por la siguiente ecuación en derivadas parciales y condiciones iniciales:

$$\begin{cases} u \in C^4 \text{ en } (a, b) \times (0, +\infty) \text{ y continua en } [a, b] \times [0, +\infty) \\ \frac{\partial u}{\partial t}(x, t) - \mu \frac{\partial^2 u}{\partial x^2}(x, t) = f(x, t) \\ u(x, 0) = u_0(x), \quad \forall x \in [a, b] \\ u(a, t) = u(b, t) = 0, \quad \forall t \in \mathbb{R} \end{cases}$$

Se desea hallar una aproximación de los valores de $u(x, t)$ en un espacio discretizado. Es decir, transformando el problema infinitesimal en un problema de diferencias finitas, hallar las soluciones de $u(x_j, t^k)$ para $x_j = a + jh$ con $j = \{1 \dots N-1\}$, $h = \frac{(b-a)}{N}$ y $t^k = k\Delta t$, $k = \{1 \dots K\}$. Donde N y $(\Delta t, K)$ definen la diferencia entre cada punto del espacio y tiempo respectivamente.

3. Metodología

3.1. Estimador para $\frac{\partial^2 u}{\partial x^2}$

Dado un tiempo t fijo y suponiendo que u es C^4 según la variable espacial, se puede escribir el desarrollo de Taylor de u en (\bar{x}, t) como:

$$u(\bar{x} + h, t) = u(\bar{x}, t) + h \frac{\partial u}{\partial x}(\bar{x}, t) + \frac{h^2}{2!} \frac{\partial^2 u}{\partial x^2}(\bar{x}, t) + \frac{h^3}{3!} \frac{\partial^3 u}{\partial x^3}(\bar{x}, t) + R_3(\theta, t) \quad (1)$$

Donde R_3 es el resto de Lagrange de orden 4 [?, p. 348]

$$R_3(\theta, t) = \frac{h^4}{4!} \frac{\partial^4 u}{\partial x^4}(\theta, t), \quad \theta \in [\bar{x}, \bar{x} + h]$$

Realizando el cambio de variable $h := -h$ en (??) se obtiene:

$$u(\bar{x} - h, t) = u(\bar{x}, t) - h \frac{\partial u}{\partial x}(\bar{x}, t) + \frac{h^2}{2!} \frac{\partial^2 u}{\partial x^2}(\bar{x}, t) - \frac{h^3}{3!} \frac{\partial^3 u}{\partial x^3}(\bar{x}, t) + R_4(\theta^*, t), \quad \theta^* \in [\bar{x} - h, \bar{x}] \quad (2)$$

Sumando (??) y (??) se obtiene:

$$\underbrace{\frac{u(\bar{x} + h, t) + u(\bar{x} - h, t) - 2u(\bar{x}, t)}{h^2}}_{\text{Estimador de } \frac{\partial^2 u}{\partial x^2}} = \frac{\partial^2 u}{\partial x^2} + \underbrace{h^2 \left[\frac{1}{4!} \frac{\partial^4 u}{\partial x^4}(\theta, t) + \frac{1}{4!} \frac{\partial^4 u}{\partial x^4}(\theta^*, t) \right]}_{\text{Error } E_T}$$

Para h suficientemente pequeño, se puede suponer que $\theta = \theta^* = \bar{x}$. Así el error de la aproximación $E_T = h^2 \frac{1}{4!} \frac{\partial^4 u}{\partial x^4}(\bar{x}, t)$. Por lo que la aproximación es de segundo orden con respecto a h .

Utilizando este estimador se puede *semidiscretizar* el problema. i.e. u se aproxima en N puntos a distancia h cada uno. Dejando la aproximación de u en función solamente del tiempo: (Denotamos $u_j : \mathbb{R} \rightarrow \mathbb{R} / u_j(t) = u(x_j, t)$)

$$\begin{cases} \frac{du_j}{dt}(t) - \frac{\mu}{h^2}(u_{j-1}(t) - 2u_j(t) + u_{j+1}(t)) = f_j(t), & j \in \{1 \dots N-1\} \\ u_0(t) = u_N(t) = 0 \end{cases} \quad (3)$$

Este sistema se puede escribir matricialmente mediante los siguientes vectores y matrices:

$$\mathbf{u}(t) = \begin{pmatrix} u_1(t) \\ \vdots \\ u_{N-1}(t) \end{pmatrix} \quad \mathbf{f}(t) = \begin{pmatrix} f_1(t) \\ \vdots \\ f_{N-1}(t) \end{pmatrix} \quad \mathbf{u}_0 = \begin{pmatrix} u_0(x_1) \\ \vdots \\ u_0(x_{N-1}) \end{pmatrix} \quad \mathbf{A} = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix}$$

Específicamente, la matriz $\mathbf{A} \in \mathcal{M}^{(N-1) \times (N-1)}$ es una matriz tri-diagonal de Topelitz [?] cuyas entradas no nulas son:

$$a_{i,i} = 2 \quad (4)$$

$$a_{i+1,i} = -1 \quad \forall i \in \{1, \dots, N-1\} \quad (5)$$

$$a_{i,i+1} = -1 \quad (6)$$

Como $u_0(t) = u_N(t) = 0$. El sistema ?? es equivalente a:

$$\begin{cases} \frac{du}{dt}(t) = -\frac{\mu}{h^2} A u(t) + f(t) \\ u(0) = u_0 \end{cases} \quad (7)$$

3.2. Autovalores y autovectores de la matriz A

A continuación se procede a estudiar el espectro de la matriz A . Lo cual será útil para probar la correctitud de los métodos de resolución del sistema que se presentarán más adelante.

Proposición 1 Los valores propios λ_j , y los vectores propios q_j , $j \in \{1 \dots N-1\}$, de A son de la forma:

$$\lambda_j = 2 \left(1 - \cos \left(j \frac{\pi}{N} \right) \right)$$

$$q_j = \left(\sin \left(j \frac{\pi}{N} \right), \sin \left(2j \frac{\pi}{N} \right), \dots, \sin \left((N-1)j \frac{\pi}{N} \right) \right)^T$$

Demostración Basta con verificar que λ_j y q_j cumplen la definición de valores y vectores propios:

$$(\forall j \in \{1, \dots, N-1\})(A \cdot q_j = \lambda_j \cdot q_j)$$

Tomando un j genérico:

$$\lambda_j q_j = 2 \left(1 - \cos \left(j \frac{\pi}{N} \right) \right) \cdot \begin{pmatrix} \sin \left(j \frac{\pi}{N} \right) \\ \sin \left(2j \frac{\pi}{N} \right) \\ \vdots \\ \sin \left((N-1)j \frac{\pi}{N} \right) \end{pmatrix} = \begin{pmatrix} 2 \sin \left(j \frac{\pi}{N} \right) - 2 \cos \left(j \frac{\pi}{N} \right) \sin \left(j \frac{\pi}{N} \right) \\ 2 \sin \left(2j \frac{\pi}{N} \right) - 2 \cos \left(j \frac{\pi}{N} \right) \sin \left(2j \frac{\pi}{N} \right) \\ \vdots \\ 2 \sin \left((N-1)j \frac{\pi}{N} \right) - 2 \cos \left(j \frac{\pi}{N} \right) \sin \left((N-1)j \frac{\pi}{N} \right) \end{pmatrix}$$

Sea k una entrada cualquiera del vector $\lambda_j q_j$, y usando la identidad trigonométrica $2 \cos(b) \sin(a) = \sin(a+b) + \sin(a-b)$, se tiene que:

$$(\lambda_j q_j)_k = 2 \sin \left(k j \frac{\pi}{N} \right) - \left(\sin \left((k+1)j \frac{\pi}{N} \right) + \sin \left((k-1)j \frac{\pi}{N} \right) \right), \forall j, k \in \{1, \dots, N-1\}$$

Por otro lado:

$$A \cdot q_j = \begin{pmatrix} 2 & -1 & 0 & \dots & 0 \\ -1 & 2 & -1 & \dots & 0 \\ \vdots & & \ddots & & \vdots \\ 0 & \dots & -1 & 2 & -1 \\ 0 & \dots & 0 & -1 & 2 \end{pmatrix} \cdot \begin{pmatrix} \sin\left(j\frac{\pi}{N}\right) \\ \sin\left(2j\frac{\pi}{N}\right) \\ \vdots \\ \sin\left((N-1)j\frac{\pi}{N}\right) \end{pmatrix}$$

Para $k \in \{2, \dots, N-2\}$ se cumple que:

$$(A \cdot q_j)_k = 2 \sin\left(kj\frac{\pi}{N}\right) - \left(\sin\left((k+1)j\frac{\pi}{N}\right) + \sin\left((k-1)j\frac{\pi}{N}\right) \right)$$

Resta ver los casos $k = 1$ y $k = N-1$:

$$\left. \begin{aligned} (A \cdot q_j)_1 &= 2 \sin\left(j\frac{\pi}{N}\right) - \sin\left(2j\frac{\pi}{N}\right) \\ (\lambda_j \cdot q_j)_1 &= 2 \sin\left(j\frac{\pi}{N}\right) - \left(\sin\left(2j\frac{\pi}{N}\right) + \sin\left((1-1)j\frac{\pi}{N}\right) \right) \end{aligned} \right\} \xRightarrow[\sin(0)=0]{} (A \cdot q_j)_1 = (\lambda_j \cdot q_j)_1$$

$$\left. \begin{aligned} (A \cdot q_j)_{N-1} &= 2 \sin\left((N-1)j\frac{\pi}{N}\right) - \sin\left((N-2)j\frac{\pi}{N}\right) \\ (\lambda_j \cdot q_j)_{N-1} &= 2 \sin\left((N-1)j\frac{\pi}{N}\right) - \left(\sin\left(\frac{N}{N}j\pi\right) + \sin\left((N-2)j\frac{\pi}{N}\right) \right) \end{aligned} \right\} \xRightarrow[j \in \mathbb{N} \rightarrow \sin(j\pi)=0]{} (A \cdot q_j)_{N-1} = (\lambda_j \cdot q_j)_{N-1}$$

Se concluye que: $A \cdot q_j = \lambda_j \cdot q_j, \quad \forall j \in \{1, \dots, N-1\}.$ □

3.3. θ -método

El θ -método es un esquema de integración, el cual se basa en aproximar una función mediante una combinación conexas de las pendientes en dos puntos discretos. La forma general del θ -método, para el problema $\frac{dy}{dt} = f(t, y), y(x_0) = t_0$, es la siguiente [?, p.36]:

$$y_{n+1} = y_n + h \left[(1 - \theta) f(t_n, y_n) + \theta f(t_{n+1}, y_{n+1}) \right], \quad \theta \in [0, 1] \quad (8)$$

La idea del mismo se ilustra en el diagrama de la Figura ???. La cual muestra que el parámetro θ es el que define la cantidad de área azul que se va a tomar para calcular el área debajo de la curva en el intervalo $[x_j, x_{j+1}]$. Este hecho muestra que el θ -método es una

generalización de varios métodos de integración:

- $\theta = 0$: “Método de Euler hacia adelante”
- $\theta = \frac{1}{2}$: “Método del trapecio”
- $\theta = 1$: “Método de Euler hacia atrás”

Comprobemos el caso particular donde $\theta = \frac{1}{2}$. En forma general, el Método del Trapecio se define como [?, p.111]:

$$y_{n+1} = y_n + \frac{h}{2} [f(t_n, y_n) + f(t_{n+1}, y_{n+1})]$$

Es claro que cuando $\theta = \frac{1}{2}$, la expresión anterior coincide con ??.

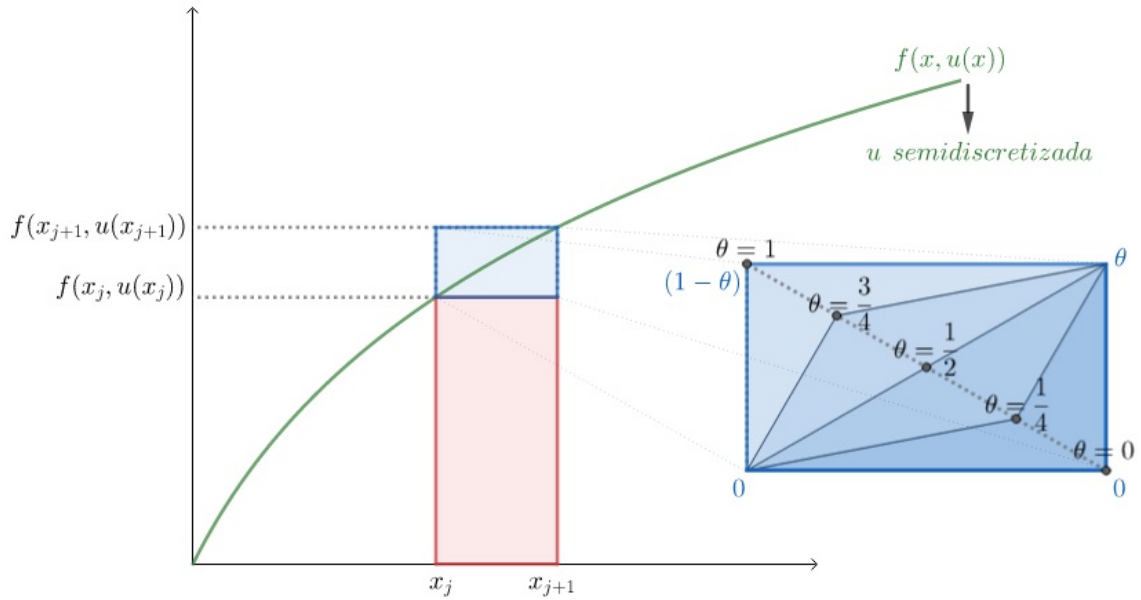


Figura 1. Ilustración del θ -método. Notar que en la imagen $f(x, t) = \frac{-\mu}{h^2} A \cdot u(t) + f^*(x, t)$, donde f^* es la función de generación calorífica del sistema.

Discretizando ahora el tiempo, sea $\Delta t > 0$ un paso del tiempo constante, se denotan los vectores como \mathbf{v}^k para referirse a los mismos evaluados en un tiempo $t^k = k\Delta t$. El siguiente sistema resulta de la aplicación del θ -método para la ecuación ??:

$$\begin{cases} \frac{\mathbf{u}^{k+1} - \mathbf{u}^k}{\Delta t} = -\frac{\mu}{h^2} \mathbf{A}(\theta \mathbf{u}^{k+1} + (1 - \theta) \mathbf{u}^k) + \theta \mathbf{f}^{k+1} + (1 - \theta) \mathbf{f}^k, k = 0, 1, \dots, K \\ \mathbf{u}^0 = \mathbf{u}_0 \end{cases} \quad (9)$$

Expresando dicho sistema en notación matricial:

$$\begin{cases} \mathbf{D} \mathbf{u}^{k+1} = (I - \frac{\mu}{h^2} \Delta t (1 - \theta) \mathbf{A}) \mathbf{u}^k + \mathbf{g}^{k+1} \\ \mathbf{u}^0 = \mathbf{u}_0 \end{cases} \quad (10)$$

donde $\mathbf{g}^{k+1} = \Delta t (\theta \mathbf{f}^{k+1} + (1 - \theta) \mathbf{f}^k)$, y $\mathbf{D} = (I + \frac{\mu}{h^2} \theta \Delta t \mathbf{A})$

Observar que el problema de aproximar la solución de una ecuación en derivadas parciales se transformó en resolver una recurrencia de sistemas lineales. Para poder resolver los mismos, es necesario que la matriz \mathbf{D} sea invertible.

Proposición 2 *La matriz \mathbf{D} es invertible.*

Demostración Para probar que la matriz \mathbf{D} es invertible, se usarán las siguientes propiedades y equivalencias:

1. \mathbf{D} es invertible $\Leftrightarrow \det(\mathbf{D}) \neq 0 \Leftrightarrow |\lambda| > 0 \quad \forall \lambda$ autovalor de \mathbf{D}
2. λ es vap de $\mathbf{A} \Rightarrow \alpha \lambda$ es vap de $\alpha \mathbf{A}$

Por definición, λ es autovalor de $\mathbf{D} \Leftrightarrow \det(\mathbf{D} - \lambda I) = 0$

$$\Leftrightarrow \det(I + \frac{\mu}{h^2} \theta \Delta t \mathbf{A} - \lambda I) = 0$$

$$\Leftrightarrow$$

$$\det(\frac{\mu}{h^2} \theta \Delta t \mathbf{A} - (\lambda - 1)I) = 0$$

$$\xLeftrightarrow{\lambda^* = \lambda^{-1}}$$

$$\det(\frac{\mu}{h^2} \theta \Delta t \mathbf{A} - (\lambda^*)I) = 0$$

Observar que estos son, por definición, los autovalores de la matriz $\frac{\mu}{h^2} \theta \Delta t \mathbf{A}$. Utilizando

(ii) y dado el hecho que se conocen los autovalores de \mathbf{A} (Prop ??):

$$\lambda^* = \frac{\mu}{h^2} \theta \Delta t 2(1 - \cos(j \frac{\pi}{N}))$$

$$\xLeftrightarrow{\lambda^* = \lambda - 1}$$

$$\lambda = \frac{\mu}{h^2} \theta \Delta t 2(1 - \cos(j \frac{\pi}{N})) + 1$$

Como $\mu, h^2, \theta, \Delta t$ son positivos (> 0) y $(1 - \cos(j \frac{\pi}{N})) \in (0, 2) \quad \forall j \in \{1 \dots N-1\}$, tenemos que $\lambda \geq 1 \Rightarrow |\lambda| > 0 \quad \forall \lambda$ autovalor de $\mathbf{D} \stackrel{(i)}{\Leftrightarrow} \mathbf{D}$ es invertible. \square

3.4. Caso $\theta = 0, f = 0$

Para este caso particular, donde $f^k = 0 \quad \forall k$. Se tiene a partir de la ecuación ?? la siguiente recurrencia:

$$\mathbf{u}^k = \mathbf{E}^k \mathbf{u}^0 \quad \text{con } k = \{1 \dots K\} \quad \text{y } \mathbf{E} = (I - \frac{\mu}{h^2} \Delta t \mathbf{A}) \quad (11)$$

Donde \mathbf{E}^k es la k -ésima potencia de \mathbf{E} y $\mathbf{u}^0 = \mathbf{u}_0$.

Para demostrar la estabilidad de este método, se prueba la siguiente proposición:

Proposición 3 $\rho(\mathbf{E}) < 1 \iff \Delta t < \frac{h^2}{2\mu}$, donde ρ es el radio espectral.

Demostración Para poder realizar la demostración, será de utilidad la siguiente propiedad:

1. Sean $A, I \in \mathcal{M}^{n \times n}$, donde A es una matriz diagonalizable cualquiera e I la identidad. Sean $\lambda_1, \dots, \lambda_n$ los autovalores de A , entonces $A + I$ tiene autovalores de la forma $\lambda_1 + 1, \dots, \lambda_n + 1$

Sea λ un vap de \mathbf{A} , y λ^* un vap de \mathbf{E} genérico. Por la propiedad mencionada anteriormente y (ii) de la proposición ?? , tenemos que:

$$\lambda^* = 1 + \frac{-\mu}{h^2} \Delta t \lambda \stackrel{\text{Prop ??}}{\iff} \lambda_j^* = 1 - \frac{\mu}{h^2} \Delta t 2(1 - \cos(j \frac{\pi}{N}))$$

$$\implies$$

$$p(\mathbf{E}) < 1 \iff \left| 1 - \frac{2\mu}{h^2} \Delta t (1 - \cos(j \frac{\pi}{N})) \right| < 1 \quad \forall j$$

Tomando un j genérico:

$$-2 < -\frac{\mu}{h^2} \Delta t 2(1 - \cos(j \frac{\pi}{N})) < 0$$

Por un lado:

$$-\frac{\mu}{h^2}\Delta t 2(1 - \cos(j\frac{\pi}{N})) < 0 \iff -\frac{\mu}{h^2}\Delta t < 0. \quad (12)$$

Lo cual es cierto. Pues el término $(1 - \cos(j\frac{\pi}{N})) \in (0, 2)$, por lo que no afecta al signo de la expresión y $\mu, h, \Delta t$ son positivos.

Por el otro:

$$-2 < -\frac{\mu}{h^2}\Delta t 2(1 - \cos(j\frac{\pi}{N})) \iff \frac{\mu}{h^2}\Delta t (1 - \cos(j\frac{\pi}{N})) < 1 \xrightarrow{N:=\frac{b-a}{h}} \frac{\mu}{h^2}\Delta t (1 - \cos(\frac{j\pi h}{b-a})) < 1$$

Entonces basta probar que:

$$\frac{\mu}{h^2}\Delta t (1 - \cos(\frac{j\pi h}{b-a})) < 1 \iff \Delta t < \frac{h^2}{2\mu} \quad (13)$$

(\Rightarrow) Suponiendo que $\frac{\mu}{h^2}\Delta t (1 - \cos(\frac{j\pi h}{b-a})) < 1$:

$$\frac{\mu}{h^2}\Delta t (1 - \cos(\frac{j\pi h}{b-a})) < 1 \Rightarrow \Delta t < \frac{h^2}{\mu} \cdot \frac{1}{(1 - \cos(\frac{j\pi h}{b-a}))} \stackrel{(*)}{\Rightarrow} \Delta t < \frac{h^2}{2\mu} \quad (14)$$

(*) Como $\inf_{j \in \{1 \dots N-1\}} \left\{ \frac{1}{(1 - \cos(\frac{j\pi h}{b-a}))} \right\} = \frac{1}{2}$, se cumple particularmente que $\Delta t < \frac{h^2}{2\mu}$.

(\Leftarrow) Suponiendo por hipótesis que $\Delta t < \frac{h^2}{2\mu}$:

$$\frac{\mu}{h^2}\Delta t (1 - \cos(\frac{j\pi h}{b-a})) \stackrel{(H)}{<} \frac{1}{2} (1 - \cos(\frac{j\pi h}{b-a})) \stackrel{(**)}{<} 1 \Rightarrow \frac{\mu}{h^2}\Delta t (1 - \cos(\frac{j\pi h}{b-a})) < 1 \quad (15)$$

(**) Como la expresión $(1 - \cos(\frac{j\pi h}{b-a}))$ está acotada entre $(0, 1)$ se cumple particularmente que $(1 - \cos(\frac{j\pi h}{b-a})) < 1$. \square

Utilizaremos la proposición anterior para demostrar que sí $\Delta t < \frac{h^2}{2\mu}$, el sistema es asintóticamente estable.

Proposición 4

$$\Delta t < \frac{h^2}{2\mu} \Rightarrow \lim_{k \rightarrow \infty} \mathbf{u}^k = \mathbf{0}$$

Demostración Recordando la siguiente propiedad del álgebra lineal: $E\mathbf{v} = \lambda\mathbf{v} \Rightarrow E^k\mathbf{v} = \lambda^k\mathbf{v}$.

Dado que todos los vaps de E son distintos entre sí (Prop ??), E es diagonalizable. i.e $\exists P \in \mathcal{M}^{(N-1) \times (N-1)} / E = PDP^{-1}$.

Por prop ??, tenemos que $\rho(E) < 1 \Rightarrow |\lambda_j| < 1 \quad \forall j \in \{1 \dots N-1\} \Rightarrow \lim_{k \rightarrow \infty} D^k = \mathbf{0}$.
 Como $E^k = PDP^{-1}PDP^{-1} \dots = PD^kP^{-1} \Rightarrow \lim_{k \rightarrow \infty} E^k = \lim_{k \rightarrow \infty} PD^kP^{-1} = \mathbf{0}$.
 Entonces $\lim_{k \rightarrow \infty} u^k = \lim_{k \rightarrow \infty} E^k u^0 = \mathbf{0}$ □

3.5. Error analítico del θ -método

Se procede a estudiar analíticamente el error del θ -método respecto al paso del tiempo Δt con x y t fijos.

Despejando u^{k+1} de la ecuación ??

$$u^{k+1} = u^k - \Delta t \left((1-\theta) \underbrace{\left(-\frac{\mu}{h^2} A u^k + f^k\right)}_{= \frac{du^k}{dt} \text{ (??)}} + \theta \underbrace{\left(-\frac{\mu}{h^2} A u^{k+1} + f^{k+1}\right)}_{= \frac{du^{k+1}}{dt} \text{ (??)}} \right) \quad (16)$$

Igualando la expresión anterior a su desarrollo de Taylor para u^{k+1} y reemplazando $\frac{du^{k+1}}{dt}$ también por su respectivo desarrollo de Taylor se tiene que:

$$\begin{aligned} u^k + \Delta t \frac{du^k}{dt} + \frac{1}{2} \Delta t^2 \frac{d^2 u^k}{dt^2} + \frac{1}{6} \Delta t^3 \frac{d^3 u^k}{dt^3} + \frac{1}{4!} \Delta t^4 \frac{d^4 u(\theta)}{dt^4} = \\ u^k + \Delta t(1-\theta) \frac{du^k}{dt} + \Delta t \theta \left(\frac{du^k}{dt} + \Delta t \frac{d^2 u^k}{dt^2} + \frac{1}{2} \Delta t^2 \theta \frac{d^3 u^k}{dt^3} + \frac{1}{6} \Delta t^3 \frac{d^4 u(\bar{\theta})}{dt^4} \right) \\ \theta, \bar{\theta} \in [t^k, t^{k+1}] \end{aligned}$$

Desarrollando la ecuación se obtiene la siguiente expresión para el error de truncamiento:

$$E_T = \left(\theta - \frac{1}{2}\right) \Delta t^2 \frac{d^2 u^k}{dt^2} + \left(\frac{1}{6} - \frac{\theta}{2}\right) \Delta t^3 \frac{d^3 u^k}{dt^3} + \frac{1}{4!} \Delta t^4 \frac{d^4 u(\theta^*)}{dt^4} + \frac{1}{6} \Delta t^4 \frac{d^4 u(\bar{\theta})}{dt^4} \quad (17)$$

Observar que si $\theta = \frac{1}{2}$ el término Δt^2 desaparece. Por lo tanto, el orden del error del θ -método respecto a Δt es 2 si $\theta = \frac{1}{2}$ y 1 si $\theta \neq \frac{1}{2}$. Es decir, el error del método será de orden 2 cuando el mismo es equivalente al método del trapecio.

3.6. Implementación

En esta última subsección se muestra una implementación del θ -método, utilizando el esquema ??, en pseudocódigo (Algo ??). Se pasan como parámetros $\theta, N, \Delta t, K$. Y se retorna una matriz $M \in \mathcal{M}^{(N+1) \times (K+1)}$ donde $M_{i,j} = u(x_i, t^j) : x_i = a + i.h$ y $t^j = j.\Delta t$ donde u es la aproximación de la solución provista por el método. Adicionalmente se pasan los parámetros a, b, u_0, f, μ que describen al problema.

Algorithm 1 θ -método

Require: $\theta \in [0, 1]$, $b > a$, $f : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$, $u_0 : \mathbb{R} \rightarrow \mathbb{R}$
1: $h \leftarrow \frac{b-a}{N}$
2: $A \leftarrow \{\mathcal{M}^{(N-1) \times (N-1)} : \text{cumple las reglas ??}\}$
3: $D \leftarrow I_{N-1} + \frac{\mu}{h^2} \cdot \theta \cdot \Delta t \cdot A$ ▷ I matriz identidad
4: $F \in \mathcal{M}^{(N+1) \times (K+1)}$
5: **for** $n = 1 : N + 1$ **do**
6: **for** $k = 0 : K$ **do**
7: $F(n, k + 1) \leftarrow f(a + n \cdot h, \Delta t \cdot k)$
8: **end for**
9: **end for**
10: $g \in \mathcal{M}^{(N+1) \times (K+1)}$
11: **for** $n = 1 : N + 1$ **do**
12: **for** $k = 1 : K$ **do**
13: $g(n, k) \leftarrow \Delta t \left(\theta \cdot F(a + h \cdot n, (k + 1)) \cdot \Delta t + (1 - \theta) F(a + h \cdot n, k \cdot \Delta t) \right)$
14: **end for**
15: **end for**
16: $M \in \mathcal{M}^{(N+1) \times (K+1)} : M = \mathbf{0}^{(N+1) \times (K+1)}$
17: **for** $n = 1 : N$ **do**
18: $M(a + h \cdot n, 1) \leftarrow u_0(a + h \cdot n)$
19: **end for**
20: $b \in \mathcal{M}^{(N-1) \times 1}$
21: **for** $k = 2 : K + 1$ **do**
22: $b \leftarrow \left(I_{N-1} - \frac{\mu}{h^2} \Delta t (1 - \theta) A \right) \cdot M(2 : N, k - 1) + g(2 : N, k)$
23: $M(2 : N, k) \leftarrow x : D \cdot x = b$
24: **end for**
25: **return** M

Para el caso particular ?? es posible implementar un algoritmo más simple mediante el esquema ?? (Algo ??). En este algoritmo se toman como parámetros $N, \Delta t, K$ más u_0, a, b, μ relacionados con el problema:

Algorithm 2 θ -método en el caso $\theta = 0, f(x, t) = 0 \forall x, t$

Require: $N \Delta t K a b \mu u_0$ $\triangleright b > a, u_0 : \mathbb{R} \rightarrow \mathbb{R}$

$h \leftarrow \frac{b-a}{N}$

2: $A \leftarrow \{\mathcal{M}^{(N-1) \times (N-1)} : \text{cumple las reglas ??}\}$

$M \in \mathcal{M}^{(N+1) \times (K+1)} : M = \mathbf{0}^{(N+1) \times (K+1)}$

4: **for** $n = 1 : N$ **do**

$M(a + h.n, 1) \leftarrow u_0(a + h.n)$

6: **end for**

$E \leftarrow I_{N-1} - \frac{\mu}{h^2 \Delta t} \cdot A$

8: $M \in \mathcal{M}^{(N+1) \times (K+1)} : M = \mathbf{0}^{(N+1) \times (K+1)}$

for $k = 2 : K + 1$ **do**

10: $M(2 : N, k) \leftarrow E^k \cdot M(2 : N, 1)$

end for

12: **return** M

4. Análisis Experimental

Para el análisis experimental se realizó una implementación del theta método en *Octave*. La simulación consta de una barra de longitud 1 ($[a, b] = [0, 1]$) con coeficiente de conductividad térmica $\mu = 1$. Una función $u_0(x) = 1000x(1-x)(1 + \frac{3}{2}x^3)$ que define la temperatura en t^0 . Y una función $f(x, t)$ que representa una fuente de generación calorífica. Para la comparación de resultados resolveremos también el problema utilizando la función *lsode* de *Octave*, la cual es un solucionador de ecuaciones diferenciales ordinarias. [?]

4.1. Simulación del θ -método con $f(x, t) = 1000 \sqrt{1-t}$

Para este caso particular, el sistema de ecuaciones diferenciales ?? resulta:

$$\frac{du}{dt}(t) = -\frac{1}{h^2} A \cdot u(t) + \begin{pmatrix} 1000 \cdot \sqrt{1-t} \\ \vdots \\ 1000 \cdot \sqrt{1-t} \end{pmatrix} \quad (18)$$

Podemos representar la barra como un segmento al cual cada punto del mismo le corresponde una temperatura. Para el problema presentado, el perfil de temperaturas en $t = 0$ puede representarse de la siguiente manera (La referencia de los colores se encuentra en la figura ??):

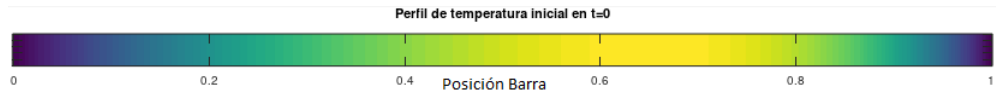


Figura 2.

Utilizando el comando *lsode* de Octave, podemos realizar una simulación de la evolución de la temperatura de la barra en el tiempo. La misma se puede visualizar tanto como un *heatmap* o como una gráfica tridimensional:

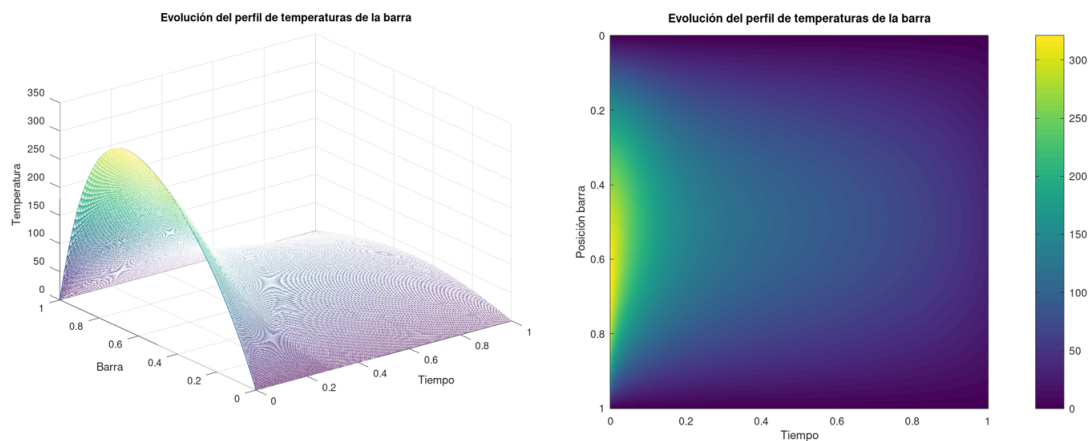


Figura 3. Representación de los resultados del comando *lsode* como *heatmap* y gráfica-3D

Simulando ahora con el θ -método con los parámetros: $\theta = \{\frac{3}{4}, \frac{1}{2}\}$, $N = 500$, $\Delta t = 2 \times 10^{-3}$, $K = 500$. Se obtienen las siguientes representaciones:

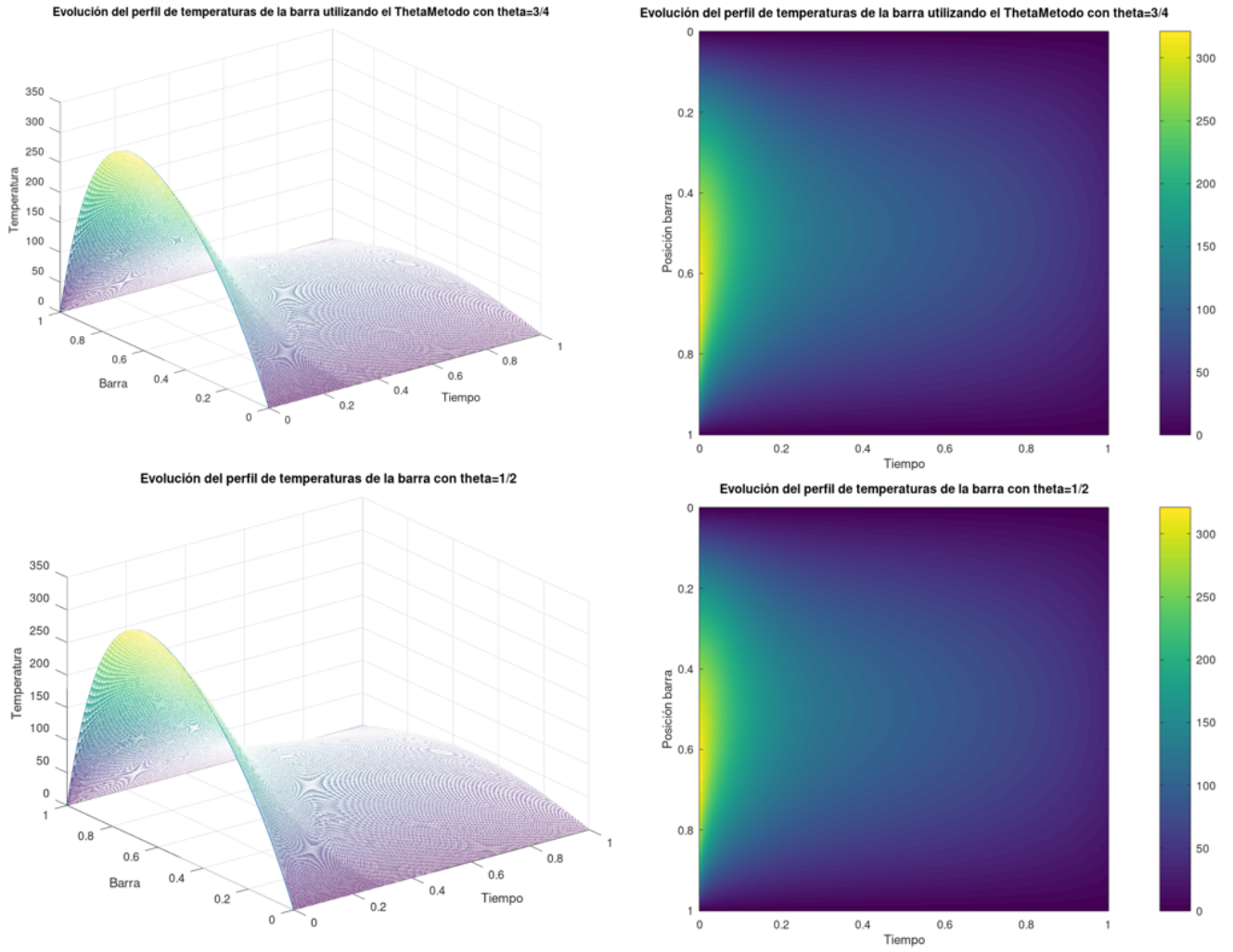


Figura 4. Representación de los resultados del θ -método para $\theta = \frac{3}{4}$ (arriba) y $\theta = \frac{1}{2}$ (abajo) como *heatmap* y gráfica-3D

Observamos que los resultados obtenidos (Fig ?? y Fig. ??) son prácticamente iguales a simple vista. Para observar con más detalle la diferencia se construye un *heatmap* adicional definido como el valor absoluto de la diferencia entre el θ -método y *lsode* en cada punto (x_j, t^k) :

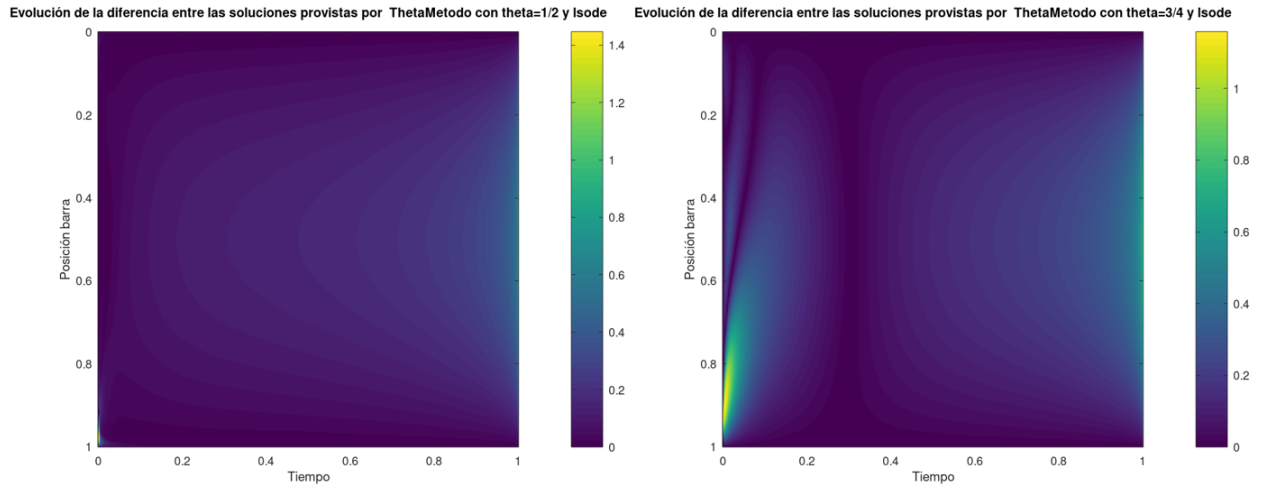


Figura 5. Representación de la diferencia entre θ -método con $\theta = \frac{3}{4}$ (derecha) y $\theta = \frac{1}{2}$ (izquierda) contra *Isode* como *heatmap*

Se aprecia un error de no más de una unidad para ambos valores de θ exceptuando un entorno particular en el tiempo alrededor de cero en la posición cercana a 1 para $\theta = \frac{3}{4}$. También se observa que el error es cada vez mayor cuanto más al centro de la barra se está y se va avanzando en t . Una posible interpretación es: dado que en los bordes por definición es $0 \forall t$, el error en los mismos es 0. Por lo que el error se amplificará al alejarnos de los bordes de la barra. Ya que la temperatura se calcula en base a ella misma y a su vecinos en el tiempo anterior (??).

Por el otro lado, con $\theta = \frac{3}{4}$ se observa un comportamiento más errático. Recordando el diagrama ?? conjeturamos que $\theta = \frac{3}{4}$ aproxima la integral más acertadamente en ciertos puntos y en otros comete mayor error que el método del trapecio. De hecho, sabemos que la función $1000 \sqrt{1-t}$ tiene derivada decreciente y la misma tiende a $-\infty$ a medida que nos acercamos al tiempo $t = 1$. Por lo que, considerando el diagrama ??, la recta entre dos puntos $(x_j, u(x_j))$ y $(x_{j+1}, u(x_{j+1}))$ tendrá una pendiente negativa de considerable magnitud en $t \approx 1$. Entonces la aproximación será peor cerca de este tiempo. Más aún, podemos notar que el error con $\theta = \frac{3}{4}$ es ligeramente mayor. Lo cual tiene sentido dado que se está tomando como aproximación un trapecio que aproxima aún peor el área debajo de la gráfica.

Considerando que el orden de la temperatura es de 10^2 , ambos métodos aproximan bastante bien a u con los parámetros dados. Pero para este caso el método del trapecio lo hace mejor.

4.2. Simulación del θ -método con $f(x, t) = 0$

Esta situación, donde la barra se encuentra desprovista de una fuente calorífica, será también simulada utilizando *lsode* como referencia y el θ -método con $\theta = 0$ con el esquema ???. Se utilizan los parámetros: $\theta = 0, N = 500, \Delta t = 2 \times 10^{-3}, K = 500$. No obstante, **no** fue posible realizar la simulación del θ -método en estas condiciones. Esto era esperable ya que en este caso $\frac{h^2}{2\mu} = 2 \times 10^{-6}$ por lo que no podemos asegurar la convergencia de la simulación debido a que la proposición ??? no se aplica. La siguiente imagen muestra el comportamiento de u utilizando *lsode*:

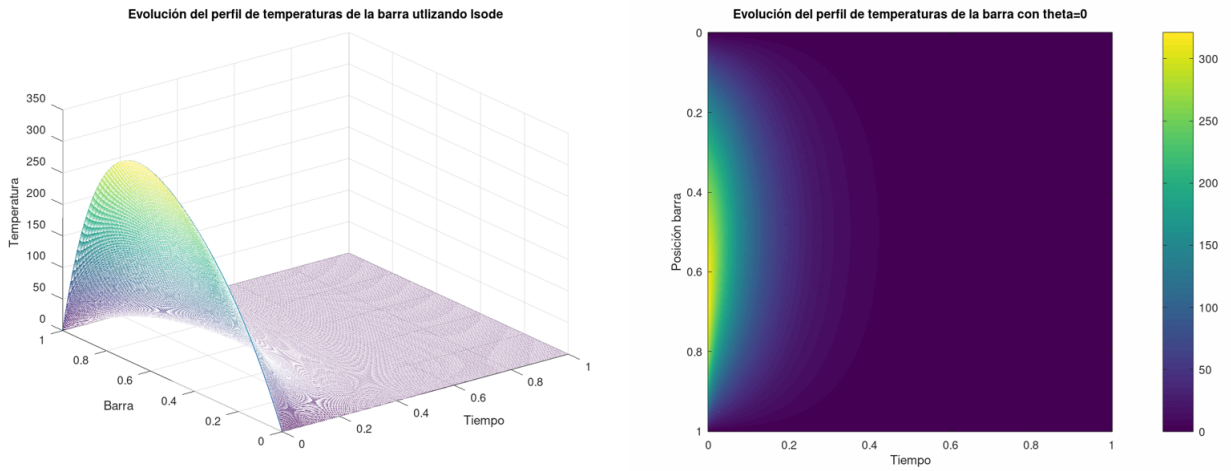


Figura 6. Representación de los resultados del comando *lsode* con $f(x, t) = 0$ mediante un *heatmap* y gráfica-3D

Obsérvese que el calor global de la barra decrece más rápido con $f(x, t) = 0$ que en el problema anterior con $f(x, t) = \sqrt{1 - t}$. Esto tiene sentido dado que en el problema anterior se le estaba dando calor al sistema en cada instante t . Pero en este caso ya no se le está proporcionando calor a la barra, y como la misma en sus 2 bordes tiene temperatura 0 en todo t , la misma se va a estar enfriando hasta que la barra entera alcance temperatura nula.

4.3. Error experimental del θ -método

En las condiciones del problema ???. Se estudia el error del θ -método comparándolo con la solución provista por el comando *lsode*.

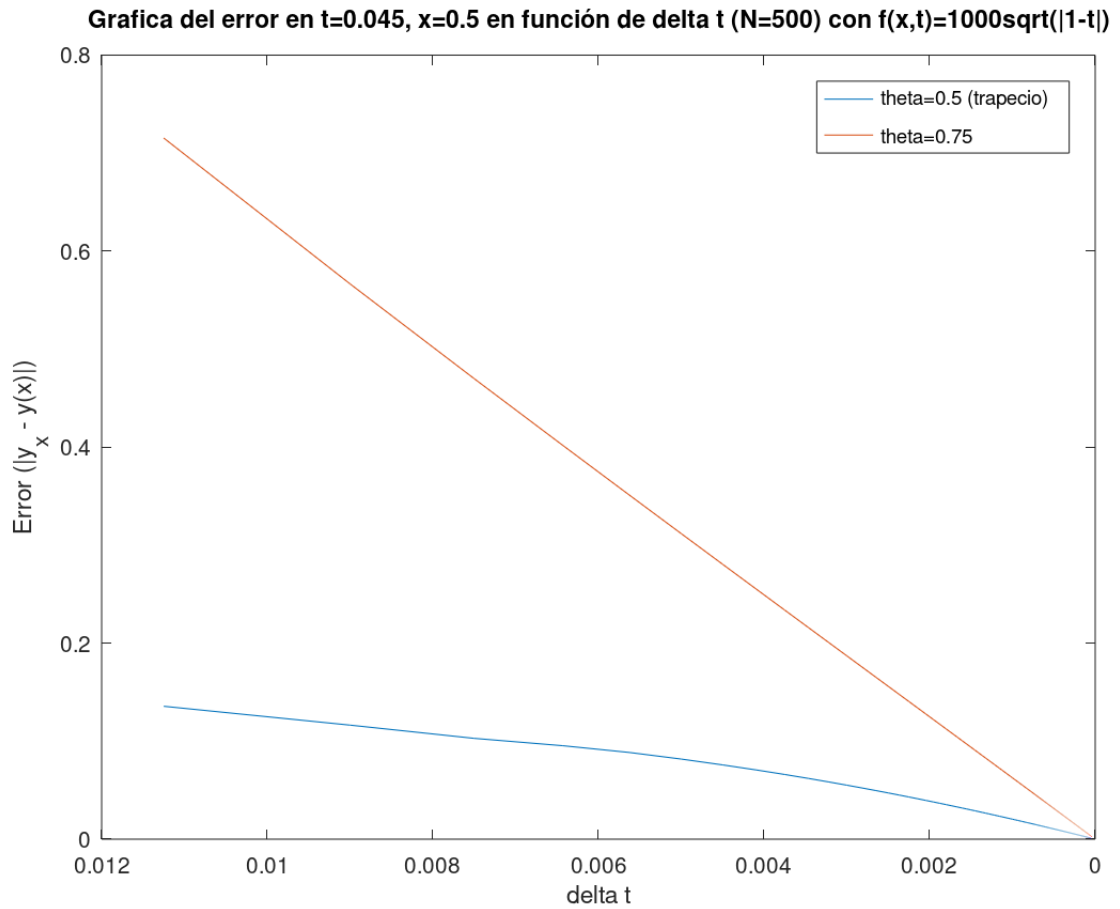


Figura 7. Valor absoluto de la diferencia entre la solución provista por *lsode* y el θ -método en función de Δt .

Podemos observar que el error de la aproximación cuando $\theta = \frac{1}{2}$ es menor que cuando $\theta = \frac{3}{4}$. Observar que el caso $\theta = \frac{1}{2}$ tiene un orden de decrecimiento mayor. Es decir, experimentalmente, el error del método del trapecio tiende a 0 más rápido que el θ -método con $\theta = \frac{3}{4}$. Estos resultados son acordes al orden analítico estudiado en la sección ??.

Por el otro lado, se estudia el error para el problema para el caso $f(x, t) = 0$. Notar que para que el método tenga solución, necesitamos que $\Delta t < \frac{h^2}{2\mu}$:

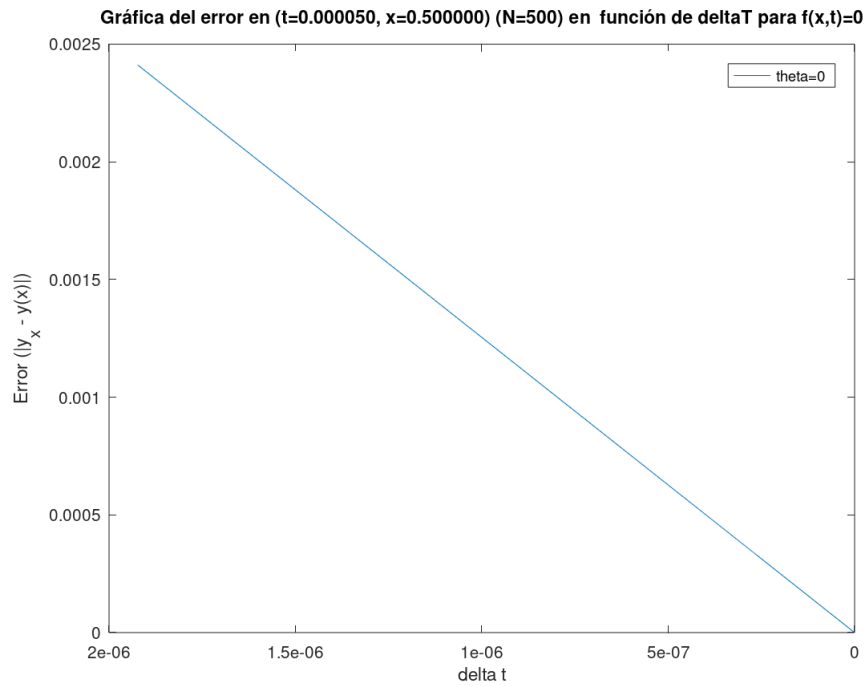


Figura 8. Valor absoluto de la diferencia entre la solución provista por *lsode* y el θ -método en función del Δt

Se observa que la evolución del error con $\theta = 0$ es similar a la de $\theta = \frac{3}{4}$. Lo cual es lo esperado dado que el orden del error del θ -método es el mismo $\forall \theta \neq \frac{1}{2}$

5. Conclusiones

Tras ahondar en la metodología que utiliza el θ -método y confirmar experimentalmente su funcionamiento, hemos conseguido implementar el algoritmo y estudiar su comportamiento satisfactoriamente en las simulaciones planteadas. También demostramos un criterio de estabilidad asintótica para el caso particular $\theta = 0$, $f(x, t) = 0$. Concluimos que el θ -método proporciona una aproximación satisfactoria de $u(x, t)$ con un error arbitrariamente pequeño definido por los parámetros $\Delta t, N$. También confirmamos experimentalmente que la evolución del error según Δt es la esperada para los casos $\theta = \{0, \frac{1}{2}, \frac{3}{4}\}$. Concluimos que para el problema planteado ($f(x, t) = 1000 \sqrt{1-t}$), el θ -método con $\theta = \frac{1}{2}$ resulta superior.

Si bien las ecuaciones en derivadas parciales se utilizan para modelar una infinidad de fenómenos físicos, como es el caso de la evolución de la distribución del calor en una barra, resolverlas no es un problema sencillo. En la sección de metodología se realiza un extenso trabajo matemático para transformar este problema a uno de ecuaciones lineales. La demostración de la correctitud de los esquemas de resolución planteados para luego realizar su simulación es lo que consideramos como el logro principal de este informe.

Una futura línea de trabajo podría ser realizar pruebas experimentales más generales. Trabajando no solo en función de Δt sino también en función de el tiempo y posición de la barra (x, t). Aunque lo anterior puede traer dificultades debido al elevado tiempo de cómputo necesario para realizar las simulaciones.