

**ANÁLISE COMPARATIVA DE MÉTODOS
NUMÉRICOS PARA EQUAÇÕES
DIFERENCIAIS PARCIAIS PARABÓLICAS**

Graziela Marchi Tiago

Dissertação de Mestrado
Pós-Graduação em Matemática Aplicada
MAP - 056

Análise Comparativa de Métodos Numéricos para Equações Diferenciais Parciais Parabólicas

Graziela Marchi Tiago

Dissertação apresentada ao Instituto de Biociências, Letras e Ciências Exatas da Universidade Estadual Paulista “Júlio de Mesquita Filho”, Câmpus de São José do Rio Preto, São Paulo, para a obtenção do título de Mestre em Matemática Aplicada.

Orientadora: Profa. Dra. Heloisa H. Marino Silva

São José do Rio Preto

2001

Resumo

O objetivo desta dissertação é estudar e comparar Fórmulas Trapezoidais Estendida e Generalizada e a Regra de Simpson Estendida quando aplicadas na resolução numérica de Equações Diferenciais Parciais Parabólicas em uma dimensão, mais especificamente para as Equações de Difusão. O Erro de Truncamento e a Estabilidade dos métodos também são analisados algebricamente. Outro método utilizado na comparação numérica é o conhecido método de Crank-Nicolson, que se baseia na Fórmula Trapezoidal Clássica para integração no tempo.

Palavras-chave: Equação de difusão, Equação de convecção-difusão, Sistema de diferenças finitas.

Abstract

The purpose of this work is to study and compare Extended and Generalized Trapezoidal Formulas and Extended Simpson Rules for the numerical solution of one dimension partial differential equations, more specifically for Diffusion Equations. Also, the truncation error and the properties of stability of the methods are investigated algebraically. Another method used for comparison is the well known Crank-Nicolson scheme based on the Classical Trapezoidal Formula for integration in time.

Keywords: Diffusion equation, Convection-diffusion equation, Finite-difference schemes.

“Se podemos sonhar,
podemos realizar. ”
(Walt Disney)

Aos meus pais,
Wilson Barbosa Tiago e
Maria Thereza Marchi Tiago.

Aos meus irmãos
Fabrício e Daniela.

Dedico

Agradecimentos

À minha família que sempre me apóia e são meus melhores amigos. Em especial, a minha mãe e ao meu pai.

À professora Heloisa, pela orientação deste trabalho.

A todos os professores, em especial Brandão, Roberto Andreani e Ângela, pelas importantes contribuições para minha formação acadêmica.

A todos os colegas da pós-graduação que estiveram comigo no decorrer deste curso pela amizade e ajuda prestada.

A todos os professores e funcionários que de alguma forma colaboraram para a realização deste trabalho.

A Deus.

À CAPES, e aos meus pais pelo auxílio financeiro.

Sumário

Introdução	1
1 Preliminares	3
1.1 Definição das Equações Diferenciais Parciais Parabólicas	3
1.2 Notação	4
1.3 Operadores Lineares	4
1.4 Definição dos Métodos Numéricos	5
1.4.1 Fórmula Trapezoidal Estendida	6
1.4.2 Fórmula Trapezoidal Generalizada	7
1.4.3 Regra de Simpson Estendida	8
1.5 Estabilidade	8
2 Equação de Difusão	11
2.1 Introdução	11
2.2 Fórmula Trapezoidal Estendida - Sistema de Diferenças Finitas (ETF-FDS)	12
2.2.1 Erro de Truncamento Local (ETF-FDS)	14
2.2.2 Estabilidade do sistema ETF-FDS	15
2.2.3 Condições de Contorno de Newman	16
2.3 Fórmula Trapezoidal Generalizada - Sistema de Diferenças Finitas (GTF(α)-FDS)	19
2.3.1 Erro de Truncamento Local (GTF(α)-FDS)	20

2.3.2	Estabilidade do sistema $GTF(\alpha)$ -FDS	21
2.3.3	Condições de Contorno de Newman	22
3	Equação de Convecção-Difusão	24
3.1	Introdução	24
3.2	Fórmula Trapezoidal Estendida - Sistema de Diferenças Finitas ($ETF(\beta_0)$ -FDS)	26
3.2.1	Erro de Truncamento Local ($ETF(\beta_0)$ -FDS)	28
3.2.2	Estabilidade do sistema $ETF(\beta_0)$ -FDS	29
3.2.3	Sistemas $ETF(\beta_0)$ -FDS diferenciados pelo valor de β_0	31
3.2.4	Condições de Contorno de Newman	32
3.3	Regra de Simpson Estendida ($ESR(\alpha_0)$)	34
3.3.1	Erro de Truncamento Local ($ESR(\alpha_0)$)	35
3.3.2	Estabilidade do método $ESR(\alpha_0)$	36
3.3.3	Sistemas $ESR(\alpha_0)$ diferenciados pelo valor de α_0	38
3.3.4	Condições de Contorno de Newman	39
3.4	Fórmula Trapezoidal Generalizada - Sistema de Diferenças Finitas($GTF(\alpha)$ -FDS)	40
3.4.1	Erro de Truncamento Local ($GTF(\alpha)$ -FDS)	41
3.4.2	Estabilidade do sistema $GTF(\alpha)$ -FDS	42
3.4.3	Transformação da Equação de Convecção-Difusão para a Equação de Difusão	43
4	Resultados Numéricos	45
4.1	Métodos Avaliados	45
4.1.1	Equação de Difusão	45
4.1.2	Equação de Convecção-Difusão	46
4.2	Implementação dos Métodos	47
4.3	Resultados Numéricos	48
4.4	Conclusões	58

Apêndice

60

Referências Bibliográficas

73

Introdução

Neste trabalho, consideramos a solução numérica de equações diferenciais parciais parabólicas em uma dimensão, mais especificamente equações de difusão. Um método de diferenças finitas bastante utilizado na obtenção da solução aproximada para estas equações é o método de Crank-Nicolson [8] que baseia-se na Fórmula Trapezoidal Clássica para integração na variável temporal. É conhecido que a Fórmula Trapezoidal possui a propriedade da A-estabilidade, mas não da L-estabilidade. Para problemas com descontinuidade nas condições de contorno e condição inicial, o método de Crank-Nicolson produz, em geral, oscilações indesejáveis na solução numérica próxima ao contorno. Um método que tem mostrado ser mais adequado para este tipo de problemas é o denominado método implícito. Este método emprega a Fórmula de Euler Regressiva para integração na variável temporal, a qual é L-estável.

O objetivo principal desta dissertação é analisar alguns métodos que tem sido propostos na literatura, para a solução de equações de difusão. Estes métodos utilizam para a integração na variável temporal as Fórmulas Trapezoidais Estendida [11] e Generalizada [6], e a Regra de Simpson Estendida [7]. Verificamos, por meio de experimentos numéricos, que estes métodos tem um desempenho superior ao Crank-Nicolson, especialmente para problemas com descontinuidade nas condições inicial e de contorno.

A Fórmula Trapezoidal Estendida foi proposta por Usmani e Agarwal [14], para solução numérica de problemas de valor inicial. Ela é obtida pelo acoplamento de dois métodos lineares de passo múltiplo, tem ordem 3 e é A-estável. Posteriormente, Jacques [11], apresentou uma modificação desta fórmula e obteve uma família de métodos estendidos de passo um, convergentes de ordem 3 que dependem de um parâmetro. Analisando a estabilidade desta família de métodos, ele obteve para diferentes valores do parâmetro, fórmulas A-estáveis e L-estáveis.

Chawla et al [6], apresentaram uma família de Fórmulas Trapezoidais Generalizadas para a integração de problemas do tipo “stiff”. Dependendo da escolha do parâmetro α , fórmulas A-estáveis e L-estáveis podem ser obtidas.

A Regra de Simpson Estendida, proposta por Chawla et al [7], é uma família de métodos que dependem de um parâmetro. Analisando a estabilidade desta família de métodos, para diferentes valores do parâmetro, métodos A-estáveis e L-estáveis podem ser obtidos.

Este trabalho está organizado em capítulos, sendo que no capítulo 1 apresentamos alguns conceitos básicos e os métodos que serão utilizados para a solução de equações diferenciais parciais parabólicas.

No capítulo 2, trabalhamos com a equação de difusão pura para a qual aplicaremos as Fórmulas Trapezoidais Estendida e Generalizada. O erro de truncamento local e a estabilidade dos métodos são estudados algebricamente.

No capítulo 3, trabalhamos com a equação de convecção-difusão para a qual aplicaremos as Fórmulas Trapezoidais Estendida e Generalizada, e a Regra de Simpson Estendida. O erro de truncamento local e a estabilidade também são estudados algebricamente.

No capítulo 4, fornecemos os resultados obtidos da implementação de todos os métodos envolvidos, utilizando o Fortran 77 e o software Mathematica para plotarmos os gráficos, e em seguida, apresentamos as considerações finais fornecendo uma análise comparativa dos métodos, e um apêndice com o exemplo de um dos programas implementados.

Capítulo 1

Preliminares

Neste capítulo, apresentamos alguns conceitos elementares e os métodos numéricos que serão utilizados no decorrer deste trabalho.

1.1 Definição das Equações Diferenciais Parciais Parabólicas

As equações estudadas neste trabalho, são a equação de difusão pura e a equação de convecção-difusão, definidas respectivamente a seguir:

$$\frac{\partial u(x, t)}{\partial t} = v \frac{\partial^2 u(x, t)}{\partial x^2} \quad (1.1.1)$$

e

$$\frac{\partial u(x, t)}{\partial t} + c \frac{\partial u(x, t)}{\partial x} = v \frac{\partial^2 u(x, t)}{\partial x^2} \quad (1.1.2)$$

sujeitas à condição inicial:

$$u(x, 0) = f(x), \quad 0 \leq x \leq l$$

e às condições de contorno:

$$u(0, t) = a(t), \quad u(l, t) = b(t), \quad t \geq 0$$

onde x é a variável espacial, t a variável temporal, c e v são constantes.

Estaremos preocupados com a obtenção de aproximações para a solução das equações (1.1.1) e (1.1.2) através de técnicas numéricas. Desta forma, supomos sempre que os problemas que serão resolvidas possuem solução.

1.2 Notação

Como estamos interessados em determinar aproximações para a solução de equações diferenciais parciais parabólicas através da substituição das derivadas espaciais por meio de diferenças finitas, estaremos, regularmente, usando vetores, matrizes e derivadas parciais de vetores. Assim, fornecemos a notação que será utilizada para denotarmos estes elementos.

Os vetores serão representados em negrito. Tomemos, por exemplo, o vetor \mathbf{u} de dimensão $n \times 1$, como

$$\mathbf{u}_j = \mathbf{u}(t_j) = \begin{bmatrix} u(x_1, t_j) \\ \vdots \\ u(x_N, t_j) \end{bmatrix}$$

onde x é a variável espacial e t é a variável temporal.

As matrizes serão representadas por letras maiúsculas e as derivadas dos vetores como $(\mathbf{u}_t)_j = \frac{\partial \mathbf{u}(t_j)}{\partial t}$. Além disso, a posição de cada elemento do vetor será indicada por $u_{i,j} = u(x_i, t_j)$.

1.3 Operadores Lineares

Para aplicarmos os métodos numéricos na solução das equações diferenciais parciais parabólicas, usaremos os seguintes operadores de diferenças finitas: o operador de diferença centrada, o operador deslocamento, os operadores de diferença progressiva e regressiva que serão denotados, respectivamente, por δ_x , E_t^+ , Δ , ∇ e definidos como:

Operadores de diferenças centradas de segunda e primeira ordem, respectivamente:

$$\begin{aligned}\delta_x^2 u_{i,j} &= \frac{1}{h^2}[u(x_{i-1}, t_j) - 2u(x_i, t_j) + u(x_{i+1}, t_j)] \simeq \frac{\partial^2 u(x_i, t_j)}{\partial x^2} \\ \delta_x u_{i,j} &= \frac{1}{2h}[u(x_{i+1}, t_j) - u(x_{i-1}, t_j)] \simeq \frac{\partial u(x_i, t_j)}{\partial x}, \quad i = 1, \dots, N\end{aligned}$$

Operador deslocamento na variável temporal:

$$E_t^+ u_{i,j} = u_{i,j+1}$$

Operador de diferença progressiva na variável espacial:

$$\Delta u_{i,j} = u_{i+1,j} - u_{i,j}$$

Operador de diferença regressiva na variável espacial:

$$\nabla u_{i,j} = u_{i,j} - u_{i-1,j}$$

1.4 Definição dos Métodos Numéricos

Para obter a solução numérica das equações (1.1.1) e (1.1.2), utilizamos discretização por diferenças finitas, mais especificamente, usamos diferenças centradas de segunda ordem na variável espacial para aproximar a derivada de segunda ordem. As equações obtidas, juntamente com as condições de fronteira e condição inicial, constituem um problema de valor inicial que será resolvido pelos métodos numéricos que descrevemos mais adiante.

Esses métodos são divididos quanto às propriedades da A-estabilidade e da L-estabilidade. Definimos estes conceitos aplicando os métodos numéricos à uma equação teste

$$y' = \lambda y, \quad \lambda \in C, \quad \text{Re}(\lambda) < 0,$$

cuja solução $y(x) = k.exp(\lambda x)$, onde k é uma constante arbitrária, tende a zero quando x tender para o infinito.

Obtemos, então, uma equação de diferenças da forma

$$y_{n+1} = R(\hat{h})y_n$$

onde $\hat{y} = \lambda h$ e $R(\hat{h})$ é a função de estabilidade do método. Claramente, $y_n \rightarrow 0$ quando $n \rightarrow \infty$ se, e somente se, $|R(\hat{h})| < 1$ e o método é absolutamente estável para os valores de \hat{h} que satisfazem esta desigualdade.

Assim definimos:

Definição 1.4.1 Denominamos a região R_A do plano complexo, definida por $R_A = \{\hat{h} \in C : |R(\hat{h})| < 1\}$, de região de estabilidade absoluta do método.

Definição 1.4.2 Um método é A-estável se $R_A \supseteq \{\hat{h} : Re(\hat{h}) < 0\}$.

Definição 1.4.3 Um método é L-estável se for A-estável e além disso, $|R(\hat{h})| \rightarrow 0$ quando $Re(\hat{h}) \rightarrow -\infty$.

Apresentamos, agora, os métodos numéricos que serão utilizados na resolução das equações de difusão e de convecção-difusão.

1.4.1 Fórmula Trapezoidal Estendida

A Fórmula Trapezoidal Estendida proposta por Usmani e Agarwal [14], para a solução do problema de valor inicial

$$y' = f(t, y), \quad y(a) = \eta, \quad a \leq t \leq b \quad (1.4.3)$$

é dada pelas equações:

$$\hat{y}_{j+2} = 5y_j - 4y_{j+1} + h(2f_j + 4f_{j+1})$$

$$\hat{f}_{j+2} = f(t_{j+2}, \hat{y}_{j+2}),$$

$$y_{j+1} = y_j + \frac{h}{12}(5f_j + 8f_{j+1} - \hat{f}_{j+2}). \quad (1.4.4)$$

Esta fórmula é A-estável e tem ordem 3. Posteriormente, Jacques [11] apresentou uma modificação desta fórmula e obteve uma família de métodos estendidos de passo um, convergentes de ordem 3 que dependem de um parâmetro β_0 , definida por:

$$\hat{y}_{j+2} = (1 + 2\beta_0)y_j - 2\beta_0y_{j+1} + h[\beta_0f_j + (2 + \beta_0)f_{j+1}],$$

$$\hat{f}_{j+2} = f(t_{j+2}, \hat{y}_{j+2}),$$

$$y_{j+1} = y_j + \frac{h}{12}[5f_j + 8f_{j+1} - \hat{f}_{j+2}]. \quad (1.4.5)$$

As Fórmulas Trapezoidais Estendidas, denotadas por ETF(β_0), são A-estáveis $\forall \beta_0 \geq -1$, e L-estáveis para $\beta_0 = 0$.

Uma classe geral de métodos de passo um estendidos foi apresentada por Chawla et al [6].

1.4.2 Fórmula Trapezoidal Generalizada

Apresentamos aqui a Fórmula Trapezoidal Generalizada, proposta por Chawla et al [6] para solução do p.v.i. (1.4.3), dada pelas equações:

$$y_{j+1} = y_j + \frac{h}{2}[(1 - \alpha)f_j + \alpha\hat{f}_j + f_{j+1}],$$

$$\hat{f}_j = f(t_j, \hat{y}_j),$$

$$\hat{y}_j = y_{j+1} - hy'_{j+1} \quad (1.4.6)$$

para o parâmetro $\alpha \in [0, 1]$. Denotaremos o método por GTF(α) o qual é L-estável para $\alpha \in (0, 1]$ e A-estável para $\alpha = 0$, que resulta na Fórmula Trapezoidal Clássica ou média aritmética, ou seja,

$$y_{j+1} = y_j + \frac{h}{2}[f_j + f_{j+1}].$$

1.4.3 Regra de Simpson Estendida

Para a solução numérica do p.v.i. (1.4.3), Chawla et al [7] forneceram uma família de um parâmetro, α_0 , que denominaram de Regras de Simpson Estendidas, descritas como:

$$\begin{aligned}\tilde{y}_{j+\frac{1}{2}} &= \frac{1}{4}y_j + \frac{3}{4}y_{j+1} - \frac{h}{4}f_{j+1}, \\ \tilde{f}_{j+\frac{1}{2}} &= f(t_{j+\frac{1}{2}}, \tilde{y}_{j+\frac{1}{2}}), \\ \hat{y}_{j+\frac{1}{2}} &= \alpha_0 y_j + (1 - \alpha_0)y_{j+1} + \frac{h}{24}[(1 + 4\alpha_0)f_j + 8(2\alpha_0 - 1)\tilde{f}_{j+\frac{1}{2}} + (4\alpha_0 - 5)f_{j+1}], \\ \hat{f}_{j+\frac{1}{2}} &= f(t_{j+\frac{1}{2}}, \hat{y}_{j+\frac{1}{2}}), \\ y_{j+1} &= y_j + \frac{h}{6}[f_j + f_{j+1} + 4\hat{f}_{j+\frac{1}{2}}].\end{aligned}\tag{1.4.7}$$

As Regras de Simpson Estendidas, denotadas por $\text{ESR}(\alpha_0)$, são A-estáveis $\forall \alpha_0 \leq \frac{1}{2}$, e L-estáveis para $\alpha_0 < \frac{1}{2}$.

1.5 Estabilidade

Para estudarmos a estabilidade dos métodos numéricos envolvidos neste trabalho, aplicamo-os às equações (1.1.1) e (1.1.2) com condições de contorno homogêneas e obtivemos um sistema da forma:

$$\mathbf{u}_{j+1} = Q\mathbf{u}_j, \quad j = 0, 1, 2, \dots$$

onde Q é a matriz de amplificação do método.

Assim, o método é estável [13] se os auto-valores de Q são em módulo menores ou iguais a um, ou seja,

$$|\lambda(Q)| \leq 1.\tag{1.5.8}$$

Definição 1.5.1 *Um método será incondicionalmente estável quando a desigualdade (1.5.8) for satisfeita $\forall h, k > 0$, ou seja, quando não temos restrições quanto ao tamanho dos passos. Caso contrário, o método será condicionalmente estável.*

Enunciaremos, a seguir, dois teoremas e algumas propriedades, necessários para demonstrarmos a estabilidade dos métodos.

Teorema 1.5.1 *Os auto-valores da matriz $N \times N$*

$$A = \begin{bmatrix} a & b & & & \\ c & a & b & & \\ & \ddots & \ddots & \ddots & \\ & & c & a & b \\ & & & c & a \end{bmatrix}$$

são:

$$\lambda(A) = a + 2\sqrt{bc} \cos\left(\frac{s\pi}{N+1}\right), \quad s = 1, \dots, N,$$

onde a, b, c podem ser reais ou complexos.

A demonstração deste teorema pode ser encontrada em [13, pg154].

Teorema 1.5.2 (Teorema de Gerschgorin) [12] *Seja A uma matriz $N \times N$ e seja D_i o disco do plano complexo com centro em a_{ii} e raio $r_i = \sum_{\substack{j=1 \\ j \neq i}}^N |a_{ij}|$, ou seja,*

$$D_i = \{z \in \mathbb{C} / |z - a_{ii}| \leq \sum_{j=1}^N |a_{ij}|\}, \quad j \neq i.$$

Os autovalores da matriz A estão contidos em $D = \cup D_i$, $i = 1, \dots, N$. Além disso, a união de k discos que não interceptam os $(N - k)$ discos restantes contem exatamente k autovalores (contando os múltiplos).

Propriedades 1.5.1 (Encontram-se em [12] e [13])

Suponha A uma matriz com autovalores λ_i . Então,

- (i) A matriz kA tem autovalores $k\lambda_i$;
- (ii) A matriz A^p , onde p é um número inteiro positivo, tem autovalores λ_i^p ;
- (iii) Se A for não singular, A^{-1} tem autovalores $\frac{1}{\lambda_i}$;
- (iv) A matriz $A + kI$ tem autovalores $\lambda_i + k$;

No caso dos sistemas envolvidos neste trabalho, a matriz de amplificação Q sempre depende de uma matriz cujos autovalores são conhecidos. Desta forma, podemos calcular os autovalores de Q usando os resultados a seguir [13, pg58]:

(v) Seja x um autovetor correspondente ao autovalor λ de uma matriz A . Então, $Ax = \lambda x$. Assim, $A(Ax) = A^2x = \lambda Ax = \lambda^2x$, mostrando que a matriz A^2 tem autovalor λ^2 correspondendo ao autovetor x . Analogamente, $A^p = \lambda^p x$, $p = 3, 4, \dots$

Se $f(A) = a_p A^p + a_{p-1} A^{p-1} + \dots + a_0 I$ é um polinômio em A com coeficientes a_p, \dots, a_0 , então $f(A)x = (a_p \lambda^p + \dots + a_0)x = f(\lambda)x$, mostrando que $f(A)$ tem autovalor $f(\lambda)$ correspondente ao autovetor x .

(vi) O autovalor de $[f_1(A)]^{-1} f_2(A)$ correspondente ao autovetor x é $f_2(\lambda)/f_1(\lambda)$, onde $f_1(A)$ e $f_2(A)$ são polinômios em A .

Capítulo 2

Equação de Difusão

Neste capítulo, apresentamos a equação de difusão com condições de contorno de Dirichlet para a qual aplicaremos as Fórmulas Trapezoidais Estendida e Generalizada, analisando o erro de truncamento local e a estabilidade dos dois métodos. Os resultados são também estendidos para a equação com condições de contorno de Newman.

2.1 Introdução

Consideremos a seguinte equação de difusão:

$$\frac{\partial u(x, t)}{\partial t} = v \frac{\partial^2 u(x, t)}{\partial x^2}, \quad 0 < x < l, \quad t > 0, \quad (2.1.1)$$

sujeita à condição inicial

$$u(x, 0) = f(x), \quad 0 \leq x \leq l, \quad (2.1.2)$$

e às condições de contorno de Dirichlet

$$u(0, t) = a(t), \quad u(l, t) = b(t), \quad t \geq 0. \quad (2.1.3)$$

Para um número inteiro positivo N , consideramos a grade retangular (x_i, t_j) , $x_i = ih$, $i = 0, \dots, N+1$, $t_j = jk$, com incremento espacial $h = \frac{l}{N+1}$ e incremento temporal $k > 0$. Ainda, definimos a constante $r = \frac{vk}{h^2}$.

Para a integração numérica de (2.1.1), o método mais utilizado é o de Crank-Nicolson que consiste na “média aritmética” dos métodos implícito e explícito para equações diferenciais parciais parabólicas, dados respectivamente por:

$$-ru_{i-1,j} + (1 + 2r)u_{i,j} - ru_{i+1,j} = u_{i,j-1}, \quad i = 1, \dots, N - 1, \quad j = 1, 2, \dots$$

$$u_{i,j+1} = u_{i,j} + r(u_{i-1,j} - 2u_{i,j} + u_{i+1,j}).$$

Assim, o método de Crank-Nicolson pode ser escrito como:

$$-ru_{i-1,j+1} + (2 + 2r)u_{i,j+1} - ru_{i+1,j+1} = ru_{i-1,j} + (2 - 2r)u_{i,j} + ru_{i+1,j}.$$

Este método é incondicionalmente estável $\forall r > 0$. Experimentos mostram que, em geral, este método produz numa região próxima do contorno, oscilações na solução numérica quando as condições inicial e de contorno são descontínuas.

Para se contornar este problema, outros métodos tem sido propostos. Assim, na seção seguinte, fazemos uma análise da Fórmula Trapezoidal Estendida para a equação de difusão (2.1.1) e na seção 2.3, analisamos a Fórmula Trapezoidal Generalizada para a mesma equação. Estes métodos se mostraram mais adequados na solução numérica dessa equação.

2.2 Fórmula Trapezoidal Estendida - Sistema de Diferenças Finitas (ETF-FDS)

Substituindo a derivada espacial da equação (2.1.1) pela fórmula de diferenças centradas, obtemos:

$$\frac{\partial u(x_i, t)}{\partial t} = \frac{v}{h^2}[u(x_{i-1}, t) - 2u(x_i, t) + u(x_{i+1}, t)], \quad i = 1, \dots, N. \quad (2.2.4)$$

Considerando

$$\mathbf{u}(t) = \begin{bmatrix} u(x_1, t) \\ u(x_2, t) \\ \vdots \\ u(x_{N-1}, t) \\ u(x_N, t) \end{bmatrix}, \quad J = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix},$$

$$\mathbf{c}(t) = \begin{bmatrix} a(t) \\ 0 \\ \vdots \\ 0 \\ b(t) \end{bmatrix}$$

e usando as condições de contorno (2.1.3), podemos escrever o sistema (2.2.4) como

$$\frac{\partial \mathbf{u}(t)}{\partial t} = \frac{v}{h^2} [\mathbf{c}(t) - J\mathbf{u}(t)] \quad (2.2.5)$$

com a condição inicial $\mathbf{u}(0) = [f(x_1), \dots, f(x_N)]^t$.

Agora, aplicando a Fórmula Trapezoidal Estendida ($\beta_0 = 0$) (1.4.5) em (2.2.5) e usando as notações definidas no capítulo 1, temos

$$\begin{aligned} \hat{\mathbf{u}}_{j+2} &= \mathbf{u}_j + 2k\mathbf{u}'_{j+1} = \mathbf{u}_j + 2k\left\{\frac{v}{h^2}[\mathbf{c}_{j+1} - J\mathbf{u}_{j+1}]\right\} \Rightarrow \\ \hat{\mathbf{u}}_{j+2} &= \mathbf{u}_j + 2r[\mathbf{c}_{j+1} - J\mathbf{u}_{j+1}] \end{aligned} \quad (2.2.6)$$

e

$$\begin{aligned} \mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{k}{12}\{5\mathbf{u}'_j + 8\mathbf{u}'_{j+1} - \hat{\mathbf{u}}'_{j+2}\} \Rightarrow \\ \mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{k}{12}\left\{5\frac{v}{h^2}[\mathbf{c}_j - J\mathbf{u}_j] + 8\frac{v}{h^2}[\mathbf{c}_{j+1} - J\mathbf{u}_{j+1}] - \frac{v}{h^2}[\mathbf{c}_{j+2} - J\hat{\mathbf{u}}_{j+2}]\right\} \end{aligned} \quad (2.2.7)$$

Agora, substituindo $\hat{\mathbf{u}}_{j+2}$ dado por (2.2.6) em (2.2.7), obtemos

$$(I + \frac{2}{3}rJ + \frac{1}{6}r^2J^2)\mathbf{u}_{j+1} = (I - \frac{1}{3}rJ)\mathbf{u}_j + \frac{r}{12}[5\mathbf{c}_j + 2(4I + rJ)\mathbf{c}_{j+1} - \mathbf{c}_{j+2}] \quad (2.2.8)$$

onde I denota a matriz identidade.

Chamamos (2.2.8) de Fórmula Trapezoidal Estendida - Sistema de Diferenças Finitas (ETF-FDS) para a equação de difusão (2.1.1). Vale notar que este sistema é pentadiagonal.

2.2.1 Erro de Truncamento Local (ETF-FDS)

Para obtermos o erro de truncamento local do sistema ETF-FDS (2.2.8), primeiramente vamos escrever todas as equações envolvidas nele usando os operadores do cálculo de diferenças finitas. Assim, obtemos

$$\begin{aligned} \hat{u}_{i,j+2} &= u_{i,j} + 2\frac{kv}{h^2}[u_{i-1,j+1} - 2u_{i,j+1} + u_{i+1,j+1}] = \\ &u_{i,j} + 2\frac{kv}{h^2}\delta_x^2 u_{i,j+1} \end{aligned} \quad (2.2.9)$$

e então, a discretização da Fórmula Trapezoidal Estendida para a equação (2.1.1) é dada por

$$\begin{aligned} u_{i,j+1} &= u_{i,j} + \frac{kv}{12h^2}[5(u_{i-1,j} - 2u_{i,j} + u_{i+1,j}) + 8(u_{i-1,j+1} - 2u_{i,j+1} + u_{i+1,j+1}) - \\ &(\hat{u}_{i-1,j+2} - 2\hat{u}_{i,j+2} + \hat{u}_{i+1,j+2})] \\ \Rightarrow \frac{1}{k}[u_{i,j+1} - u_{i,j}] &= \frac{v}{12h^2}[(5u_{i-1,j} + 8u_{i-1,j+1} - \hat{u}_{i-1,j+2}) - \\ &2(5u_{i,j} + 8u_{i,j+1} - \hat{u}_{i,j+2}) + (5u_{i+1,j} + 8u_{i+1,j+1} - \hat{u}_{i+1,j+2})] \\ \Rightarrow \frac{1}{k}[u_{i,j+1} - u_{i,j}] &= \frac{v}{12h^2}\delta_x^2(5u_{i,j} + 8u_{i,j+1} - \hat{u}_{i,j+2}). \end{aligned} \quad (2.2.10)$$

Agora, substituindo $\hat{u}_{i,j+2}$, dado por (2.2.9), em (2.2.10), a ETF-FDS é descrita por

$$\frac{1}{k}[u_{i,j+1} - u_{i,j}] = \frac{v}{6h^2}\delta_x^2[2u_{i,j} + (4 - \frac{kv}{h^2}\delta_x^2)u_{i,j+1}]. \quad (2.2.11)$$

Utilizando o operador deslocamento E_t^+ , o operador de diferenças finitas que define o método (2.2.11) resulta em

$$L = \frac{1}{k}(E_t^+ - 1) - \frac{v}{6h^2}\delta_x^2[2 + (4 - \frac{kv}{h^2}\delta_x^2)E_t^+]. \quad (2.2.12)$$

Vamos obter, então, o erro de truncamento local para o operador L .

Seja $D_x = \frac{d}{dx}, \dots$, e note que $D_t = vD_x^2$. Aplicando Série de Taylor em todos os operadores envolvidos em (2.2.12), temos:

$$\frac{1}{h^2}\delta_x^2 = D_x^2 + \frac{h^2}{12}D_x^4 + \frac{h^4}{360}D_x^6 + \dots$$

e

$$E_t^+ = I + hD_x + \frac{h^2}{2}D_x^2 + \dots \quad (2.2.13)$$

Substituindo-os obtemos

$$2 + \left(4 - \frac{kv}{h^2}\delta_x^2\right)E_t^+ = 6 + 3kD_t + k^2D_t^2 + \frac{1}{6}k^3D_t^3 - \frac{1}{12}kh^2D_tD_x^2 - \frac{1}{12}k^2h^2D_t^2D_x^2 + \dots$$

e, consequentemente

$$\begin{aligned} \frac{v}{6h^2}\delta_x^2[2 + \left(4 - \frac{kv}{h^2}\delta_x^2\right)E_t^+] &= D_t + \frac{1}{2}kD_t^2 + \frac{1}{6}k^2D_t^3 + \frac{1}{36}k^3D_t^4 + \frac{1}{12}h^2D_tD_x^2 + \\ &+ \frac{1}{360}h^4D_tD_x^4 + \frac{1}{36}kh^2D_t^2D_x^2 + \dots \end{aligned}$$

Logo, temos que

$$L = \frac{1}{72}k^3D_t^4 - \frac{1}{12}h^2D_tD_x^2 - \frac{1}{36}kh^2D_t^2D_x^2 - \frac{1}{360}h^4D_tD_x^4 + \dots \quad (2.2.14)$$

mostrando com isso que o sistema ETF-FDS tem ordem $O(h^2) + O(k^3)$, ou seja, possui ordem 3 na variável temporal e ordem 2 na variável espacial.

2.2.2 Estabilidade do sistema ETF-FDS

Para condições de contorno homogêneas, o sistema ETF-FDS (2.2.8) pode ser escrito como

$$\mathbf{u}_{j+1} = Q\mathbf{u}_j, \quad j = 0, 1, 2, \dots \quad (2.2.15)$$

com a matriz de amplificação para o sistema de diferenças dada por

$$Q = \left(I + \frac{2}{3}rJ + \frac{1}{6}r^2J^2\right)^{-1}\left(I - \frac{1}{3}rJ\right).$$

Pela definição 1.5.1, para mostrarmos que o sistema ETF-FDS é incondicionalmente estável, devemos mostrar que os auto-valores da matriz Q são em

módulo menores ou iguais a um. Também pelo Teorema 1.5.1, sabe-se que os autovalores da matriz J são dados por

$$\lambda_s(J) = 4 \cos^2(\theta_s), \quad \theta_s = \frac{s\pi}{2(N+1)}, \quad s = 1, \dots, N.$$

Assim, os autovalores da matriz Q (Propriedades 1.5.1) são dados por

$$\lambda_s(Q) = \frac{3 - 4r \cos^2(\theta_s)}{3 + 8r \cos^2(\theta_s) + 8r^2 \cos^4(\theta_s)}, \quad s = 1, \dots, N.$$

Agora,

$$\begin{aligned} \left| \frac{3 - 4r \cos^2(\theta_s)}{3 + 8r \cos^2(\theta_s) + 8r^2 \cos^4(\theta_s)} \right| \leq 1 &\Leftrightarrow -1 \leq \frac{3 - 4r \cos^2(\theta_s)}{3 + 8r \cos^2(\theta_s) + 8r^2 \cos^4(\theta_s)} \leq 1 \\ &\Leftrightarrow r > 0 \quad \text{e} \quad \left(r \leq \frac{-3}{2 \cos^2(\theta_s)} \quad \text{ou} \quad r \geq 0 \right). \end{aligned}$$

Percebemos então, que o sistema ETF-FDS é incondicionalmente estável $\forall r > 0$, ou seja, não temos restrições sobre h e k .

2.2.3 Condições de Contorno de Newman

Aplicamos, aqui, a Fórmula Trapezoidal Estendida - Sistema de Diferenças Finitas para a equação de difusão (2.1.1), com a mesma condição inicial (2.1.2), mas com condições de contorno de Newman da forma:

$$\frac{\partial u(0, t)}{\partial x} = a_1 u(t) - b_1 \quad \text{e} \quad \frac{\partial u(l, t)}{\partial x} = -a_2 u(t) + b_2, \quad (2.2.16)$$

onde a_1, a_2, b_1, b_2 são constantes não-negativas. Primeiramente, consideramos a discretização das condições de contorno para $x = 0$. Introduzimos o ponto fictício x_{-1} , e substituindo a derivada espacial pelo operador de diferenças centradas de primeira ordem, obtemos

$$\begin{aligned} \frac{\partial u(0, t)}{\partial x} &= a_1 u_0(t) - b_1 \quad \text{e} \quad \frac{\partial u(x_i, t)}{\partial x} = \frac{1}{2h} [u(x_{i+1}, t) - u(x_{i-1}, t)] \\ \Rightarrow a_1 u_0(t) - b_1 &= \frac{1}{2h} [u_1(t) - u_{-1}(t)] \end{aligned} \quad (2.2.17)$$

Usando a equação (2.2.17) e eliminando u_{-1} da equação (2.2.4), para $i = 0$, obtemos:

$$\begin{aligned}\frac{\partial u_0(t)}{\partial t} &= \frac{v}{h^2}[u_{-1}(t) - 2u_0(t) + u_1(t)] = \\ &= \frac{v}{h^2}[-(a_1 u_0(t) - b_1)2h + u_1(t) - 2u_0(t) + u_1(t)] \Rightarrow \\ \frac{\partial u_0(t)}{\partial t} &= \frac{v}{h^2}[2u_1(t) - (2 + 2ha_1)u_0(t)] + \frac{2v}{h}b_1.\end{aligned}\quad (2.2.18)$$

De maneira análoga, para condições de contorno em $x = l$, temos:

$$\begin{aligned}\frac{\partial u(l, t)}{\partial x} &= -a_2 u_{N+1}(t) + b_2 \quad e \quad \frac{\partial u(x_i, t)}{\partial x} = \frac{1}{2h}[u(x_{i+1}, t) - u(x_{i-1}, t)] \\ \Rightarrow -a_2 u_{N+1}(t) + b_2 &= \frac{1}{2h}[u_{N+2}(t) - u_N(t)]\end{aligned}$$

e eliminando u_{N+2} da equação 2.2.4, para $i = N + 1$, obtemos então

$$\frac{\partial u_{N+1}(t)}{\partial t} = \frac{v}{h^2}[2u_N(t) - (2 + 2ha_2)u_{N+1}(t)] + \frac{2v}{h}b_2.\quad (2.2.19)$$

Definindo

$$\mathbf{u}(t) = \begin{bmatrix} u(x_0, t) \\ u(x_1, t) \\ \vdots \\ u(x_N, t) \\ u(x_{N+1}, t) \end{bmatrix}, \quad A = \begin{bmatrix} 2 + 2ha_1 & -2 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -2 & 2 + 2ha_2 \end{bmatrix},$$

$$\mathbf{b} = \begin{bmatrix} b_1 \\ 0 \\ \vdots \\ 0 \\ b_2 \end{bmatrix}$$

então, podemos escrever (2.2.18), (2.2.4) (para $i = 1, \dots, N$) e (2.2.19) como

$$\frac{\partial \mathbf{u}(t)}{\partial t} = -\frac{v}{h^2}A\mathbf{u}(t) + \frac{2v}{h}\mathbf{b}.\quad (2.2.20)$$

Agora, aplicando a Fórmula Trapezoidal Estendida ($\beta_0 = 0$) (1.4.5) para (2.2.20) temos

$$\begin{aligned}\hat{\mathbf{u}}_{j+2} &= \mathbf{u}_j + 2k\mathbf{u}'_{j+1} \Rightarrow \\ \hat{\mathbf{u}}_{j+2} &= \mathbf{u}_j - 2rA\mathbf{u}_{j+1} + 4rh\mathbf{b}\end{aligned}\quad (2.2.21)$$

e

$$\begin{aligned}\mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{k}{12}[5\mathbf{u}'_j + 8\mathbf{u}'_{j+1} - \hat{\mathbf{u}}'_{j+2}] \Rightarrow \\ \mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{r}{12}(24h\mathbf{b} - 5A\mathbf{u}_j - 8A\mathbf{u}_{j+1} + A\hat{\mathbf{u}}_{j+2}).\end{aligned}\quad (2.2.22)$$

Substituindo $\hat{\mathbf{u}}_{j+2}$ dado por (2.2.21) em (2.2.22), obtemos

$$(I + \frac{2}{3}rA + \frac{1}{6}r^2A^2)\mathbf{u}_{j+1} = (I - \frac{1}{3}rA)\mathbf{u}_j + \frac{1}{3}rh(6I + rA)\mathbf{b}.\quad (2.2.23)$$

Para mostrar a estabilidade do sistema (2.2.23), primeiro notamos que com a ajuda do Teorema de Gerschgorin, os auto-valores da matriz A , denotados por μ , são todos não-negativos (Teorema 1.5.2). Agora, os auto-valores da matriz de amplificação

$$Q^* = (I + \frac{2}{3}rA + \frac{1}{6}r^2A^2)^{-1}(I - \frac{1}{3}rA),$$

do sistema (2.2.23) são dados por (Propriedades 1.5.1)

$$\lambda(Q^*) = \frac{1 - \frac{1}{3}r\mu}{1 + \frac{2}{3}r\mu + \frac{1}{6}r^2\mu^2}.$$

Assim, analisando a região onde $|\lambda(Q^*)| \leq 1$, obtemos que o sistema (2.2.23) será incondicionalmente estável $\forall r > 0$.

2.3 Fórmula Trapezoidal Generalizada - Sistema de Diferenças Finitas (GTF(α)-FDS)

Analogamente à Seção 2.2, aplicamos a Fórmula Trapezoidal Generalizada (1.4.6) em (2.2.5) e usando as notações definidas no capítulo 1, temos

$$\hat{\mathbf{u}}_j = \mathbf{u}_{j+1} - \frac{kv}{h^2}[\mathbf{c}_{j+1} - J\mathbf{u}_{j+1}] \Rightarrow \hat{\mathbf{u}}_j = (I + rJ)\mathbf{u}_{j+1} - r\mathbf{c}_{j+1} \quad (2.3.24)$$

e

$$\begin{aligned} \mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{k}{2}[(1 - \alpha)\mathbf{u}'_j + \alpha\hat{\mathbf{u}}'_j + \mathbf{u}'_{j+1}], \quad \alpha \in [0, 1] \Rightarrow \\ \mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{r}{2}[\mathbf{c}_j + \mathbf{c}_{j+1} - J\{(1 - \alpha)\mathbf{u}_j + \alpha\hat{\mathbf{u}}_j + \mathbf{u}_{j+1}\}]. \end{aligned} \quad (2.3.25)$$

Agora, substituindo $\hat{\mathbf{u}}_{j+2}$ dado por (2.3.24) em (2.3.25), obtemos

$$(I + \frac{1}{2}(1 + \alpha)rJ + \frac{1}{2}\alpha r^2 J^2)\mathbf{u}_{j+1} = (I - \frac{1}{2}(1 - \alpha)rJ)\mathbf{u}_j + \frac{r}{2}\mathbf{c}_j + \frac{r}{2}[I + \alpha rJ]\mathbf{c}_{j+1}, \quad (2.3.26)$$

onde I denota a matriz identidade.

Chamamos (2.3.26) de Fórmula Trapezoidal Generalizada - Sistema de Diferenças Finitas (GTF(α)-FDS) para a equação de difusão (2.1.1).

O sistema (2.3.26) requer, em geral, a solução de um sistema linear pentadiagonal com a matriz dos coeficientes

$$M = I + \frac{1}{2}(1 + \alpha)rJ + \frac{1}{2}\alpha r^2 J^2.$$

Em [3], os autores sugerem que para N grande podemos calcular a solução de (2.3.26) resolvendo o sistema linear pentadiagonal pelo produto de duas matrizes tridiagonais P , onde $M = P^2$ e

$$P = I + \frac{1}{4}(1 + \alpha)rJ, \quad \text{com} \quad \alpha = 3 - 2\sqrt{2}.$$

2.3.1 Erro de Truncamento Local (GTF(α)-FDS)

Para obtermos o erro de truncamento local do sistema GTF(α)-FDS (2.3.26), adotamos o mesmo procedimento da seção 2.2.1 e usando as notações definidas no capítulo 1, temos

$$\begin{aligned}\hat{u}_{i,j} &= u_{i,j+1} - \frac{kv}{h^2}[u_{i-1,j+1} - 2u_{i,j+1} + u_{i+1,j+1}] = \\ &= \left(1 - \frac{kv}{h^2}\delta_x^2\right)u_{i,j+1}\end{aligned}\quad (2.3.27)$$

e então, a discretização da Fórmula Trapezoidal Generalizada (1.4.6) para a equação (2.1.1) é dada por

$$\frac{1}{k}[u_{i,j+1} - u_{i,j}] = \frac{1}{2}\frac{v}{h^2}\delta_x^2((1-\alpha)u_{i,j} + u_{i,j+1} + \alpha\hat{u}_{i,j}).\quad (2.3.28)$$

Agora, substituindo $\hat{u}_{i,j}$ dado por (2.3.27) em (2.3.28), a GTF(α)-FDS é descrita por

$$\frac{1}{k}[u_{i,j+1} - u_{i,j}] = \frac{v}{2h^2}\delta_x^2[(1-\alpha)u_{i,j} + \alpha(1 - \frac{kv}{h^2}\delta_x^2)u_{i,j+1} + u_{i,j+1}].\quad (2.3.29)$$

Utilizando o operador deslocamento E_t^+ , o operador de diferenças finitas do método (2.3.29) resulta em

$$L = \frac{1}{k}(E_t^+ - 1) - \frac{v}{2h^2}\delta_x^2[(1-\alpha) + \{\alpha(1 - \frac{kv}{h^2}\delta_x^2) + 1\}E_t^+].\quad (2.3.30)$$

Passemos então, ao erro de truncamento local para o operador L .

Novamente, seja $D_x = \frac{d}{dx}, \dots$, e note que $D_t = vD_x^2$. Utilizando (2.2.13), obtemos

$$\begin{aligned}(1-\alpha) + \{\alpha(1 - \frac{kv}{h^2}\delta_x^2) + 1\}E_t^+ &= 2 + kD_t + (1-\alpha)\frac{1}{2}k^2D_t^2 + \\ &+ (1-2\alpha)\frac{1}{6}k^3D_t^3 + (1-3\alpha)\frac{1}{24}k^4D_t^4 - \\ &+ \frac{1}{12}\alpha kh^2D_tD_x^2 - \frac{1}{12}\alpha k^2h^2D_x^2D_t^2 - \dots\end{aligned}$$

e

$$\begin{aligned} \frac{v}{2h^2}\delta_x^2[(1-\alpha) + \{\alpha(1 - \frac{kv}{h^2}\delta_x^2) + 1\}E_t^+] = D_t + \frac{1}{2}kD_t^2 + (1-\alpha)\frac{1}{4}K^2D_t^3 + \\ (1-2\alpha)\frac{1}{12}k^3D_t^4 + (1-3\alpha)\frac{1}{48}k^4D_t^5 + \frac{1}{24}(1-\alpha)kh^2D_t^2D_x^2 + \\ \frac{1}{48}k^2h^2D_t^3D_x^2 + \frac{1}{12}h^2D_x^2D_t + \frac{1}{360}h^4D_x^4D_t \dots \end{aligned}$$

Logo, temos que

$$\begin{aligned} L = \frac{1}{12}(3\alpha-1)k^2D_t^3 - \frac{1}{12}h^2D_tD_x^2 + \frac{1}{24}(4\alpha-1)k^3D_t^4 + \\ \frac{1}{24}(\alpha-1)kh^2D_x^2D_t^2 + \frac{1}{80}(5\alpha-1)k^4D_t^5 + \\ \frac{1}{48}(3\alpha-1)k^2h^2D_x^2D_t^3 - \frac{1}{360}h^4D_x^4D_t + \dots \end{aligned} \quad (2.3.31)$$

mostrando que o sistema GTF(α)-FDS tem ordem $O(h^2) + O(k^2)$ se $\alpha \neq \frac{1}{3}$. É interessante notar que para $\alpha = \frac{1}{3}$, o método GTF($\frac{1}{3}$)-FDS tem ordem 3 na variável temporal.

2.3.2 Estabilidade do sistema GTF(α)-FDS

Para condições de contorno homogêneas, o sistema GTF(α)-FDS (2.3.26) pode ser escrito como

$$\mathbf{u}_{j+1} = Q\mathbf{u}_j, \quad j = 0, 1, 2, \dots \quad (2.3.32)$$

com matriz de amplificação dada por

$$Q = (I + \frac{1}{2}(1+\alpha)rJ + \frac{1}{2}\alpha r^2 J^2)^{-1} (I - \frac{1}{2}(1-\alpha)rJ).$$

Pela definição 1.5.1, para provarmos que o sistema GTF(α)-FDS, $\alpha \in [0, 1]$, é incondicionalmente estável, devemos ter os auto-valores da matriz Q , em módulo, menores ou iguais a um. Também pelo Teorema 1.5.1, sabemos que os auto-valores da matriz J são dados por

$$\lambda_s(J) = 4 \cos^2(\theta_s), \quad \theta_s = \frac{s\pi}{2(N+1)}, \quad s = 1, \dots, N.$$

Assim, os auto-valores da matriz Q (Propriedades 1.5.1) são dados por

$$\lambda_s(Q) = \frac{1 - 2(1 - \alpha)r \cos^2(\theta_s)}{1 + 2(1 + \alpha)r \cos^2(\theta_s) + 8\alpha r^2 \cos^4(\theta_s)}, \quad s = 1, \dots, N.$$

Agora,

$$\begin{aligned} & \left| \frac{1 - 2(1 - \alpha)r \cos^2(\theta_s)}{1 + 2(1 + \alpha)r \cos^2(\theta_s) + 8\alpha r^2 \cos^4(\theta_s)} \right| \leq 1 \Leftrightarrow \\ -1 & \leq \frac{1 - 2(1 - \alpha)r \cos^2(\theta_s)}{1 + 2(1 + \alpha)r \cos^2(\theta_s) + 8\alpha r^2 \cos^4(\theta_s)} \leq 1 \\ \Leftrightarrow & r > 0 \quad \text{e} \quad r \geq \frac{-1}{2 \cos^2(\theta_s)} \end{aligned}$$

Percebemos então, que o sistema $GTF(\alpha)$ -FDS será incondicionalmente estável $\forall r > 0$.

2.3.3 Condições de Contorno de Newman

Analogamente à Seção 2.2.3, aplicamos a Fórmula Trapezoidal Generalizada (1.4.6) para (2.2.20) e obtivemos com isso

$$\begin{aligned} \hat{\mathbf{u}}_j &= \mathbf{u}_{j+1} - k\mathbf{u}'_{j+1} \Rightarrow \\ \hat{\mathbf{u}}_j &= (I + rA)\mathbf{u}_{j+1} - 2rh\mathbf{b} \end{aligned} \quad (2.3.33)$$

e

$$\begin{aligned} \mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{k}{2}[(1 - \alpha)\mathbf{u}'_j + \alpha\hat{\mathbf{u}}'_j + \mathbf{u}'_{j+1}] \Rightarrow \\ \mathbf{u}_{j+1} &= \mathbf{u}_j - \frac{r}{2}A[(1 - \alpha)\mathbf{u}_j + \alpha\hat{\mathbf{u}}_j + \mathbf{u}_{j+1}] + 2rh\mathbf{b}. \end{aligned} \quad (2.3.34)$$

Substituindo $\hat{\mathbf{u}}_{j+2}$ dado por (2.3.33) em (2.3.34), obtemos

$$(I + \frac{1}{2}(1 + \alpha)rA + \frac{1}{2}\alpha r^2 A^2)\mathbf{u}_{j+1} = (I - \frac{1}{2}(1 - \alpha)rA)\mathbf{u}_j + rh(2I + \alpha rA)\mathbf{b}. \quad (2.3.35)$$

Para mostrar a estabilidade do método (2.3.35), primeiro notamos que com a ajuda do Teorema de Gerschgorin, os auto-valores da matriz A , denotados por μ , são todos não-negativos (Teorema 1.5.2). Agora, os auto-valores da matriz de amplificação

$$Q^* = (I + \frac{1}{2}(1 + \alpha)rA + \frac{1}{2}\alpha r^2 A^2)^{-1}(I - \frac{1}{2}(1 - \alpha)rA),$$

do sistema (2.3.35) são dados por (Propriedades 1.5.1)

$$\lambda(Q^*) = \frac{1 - \frac{1}{2}(1 - \alpha)r\mu}{1 + \frac{1}{2}(1 + \alpha)r\mu + \frac{1}{2}\alpha r^2 \mu^2}.$$

Assim, analisando a região onde $|\lambda(Q^*)| \leq 1$, obtemos que o sistema (2.3.35) será incondicionalmente estável $\forall r > 0$.

Capítulo 3

Equação de Convecção-Difusão

Neste capítulo, apresentamos a equação de convecção-difusão com condições de contorno de Dirichlet para o qual aplicaremos as Fórmulas Trapezoidais Estendida e Generalizada, e a Regra de Simpson Estendida, analisando o erro de truncamento local e a estabilidade dos três métodos. Os resultados são também estendidos para a equação com condições de contorno de Newman.

3.1 Introdução

Consideremos a seguinte equação de convecção-difusão:

$$\frac{\partial u(x, t)}{\partial t} + c \frac{\partial u(x, t)}{\partial x} = v \frac{\partial^2 u(x, t)}{\partial x^2}, \quad 0 < x < l, \quad t > 0, \quad (3.1.1)$$

sujeita à condição inicial

$$u(x, 0) = f(x), \quad 0 \leq x \leq l, \quad (3.1.2)$$

e às condições de contorno de Dirichlet

$$u(0, t) = a(t), \quad u(l, t) = b(t), \quad t \geq 0. \quad (3.1.3)$$

Para um número inteiro positivo N , consideramos a grade retangular (x_i, t_j) , $x_i = ih$, $i = 0, \dots, N + 1$, $t_j = jk$, com incremento espacial $h = \frac{l}{N+1}$ e incremento temporal $k > 0$. Ainda, definimos as constantes $r = \frac{vk}{h^2}$ e $\rho = \frac{ck}{h}$.

Para a integração numérica de (3.1.1), os métodos mais utilizados são o explícito, o implícito e o Crank-Nicolson. Um dos métodos utilizado para resolver (3.1.1) é o explícito

$$u_{i,j+1} = u_{i,j} - \frac{\rho}{2}(u_{i+1,j} - u_{i-1,j}) + r(u_{i+1,j} - 2u_{i,j} + u_{i-1,j})$$

que emprega a Fórmula de Euler para integração no tempo.

Um outro método também utilizado para resolver (3.1.1) é o implícito

$$\left(-r + \frac{\rho}{2}\right)u_{i+1,j+1} + (1 + 2r)u_{i,j+1} + \left(-r - \frac{\rho}{2}\right)u_{i-1,j+1} = u_{i,j}$$

que emprega a Fórmula de Euler Regressiva para integração no tempo.

O método de Crank-Nicolson é a "média aritmética" dos métodos implícito e explícito para equações diferenciais parciais parabólicas e pode ser escrito como:

$$\begin{aligned} &\left(\frac{-r}{2} + \frac{\rho}{4}\right)u_{i+1,j+1} + (1 + r)u_{i,j+1} + \left(\frac{-r}{2} - \frac{\rho}{4}\right)u_{i-1,j+1} = \\ &\left(\frac{r}{2} - \frac{\rho}{4}\right)u_{i+1,j} + (1 - r)u_{i,j} + \left(\frac{r}{2} + \frac{\rho}{4}\right)u_{i-1,j}. \end{aligned}$$

Este método é incondicionalmente estável. Note que para $c = 0$, ele se reduz ao conhecido método de Crank-Nicolson para a equação de difusão $u_t = \nu u_{xx}$.

Experimentos mostram que, em geral, este método produz numa região próxima do contorno, oscilações na solução numérica quando as condições inicial e de contorno são descontínuas.

Para se contornar este problema outros métodos tem sido propostos, e na seção seguinte, fazemos uma análise da Fórmula Trapezoidal Estendida para a equação de convecção-difusão (3.1.1); na seção 3.3, analisamos a Regra de Simpson Estendida e na seção 3.4, analisamos a Fórmula Trapezoidal Generalizada para a mesma equação. Estes métodos se mostram mais adequados na solução numérica dessa equação.

3.2 Fórmula Trapezoidal Estendida - Sistema de Diferenças Finitas (ETF(β_0)-FDS)

Substituindo as derivadas espaciais da equação (3.1.1) pela fórmula de diferenças centradas, obtemos:

$$\frac{\partial u(x_i, t)}{\partial t} + \frac{c}{2h}[u(x_{i+1}, t) - u(x_{i-1}, t)] = \frac{v}{h^2}[u(x_{i-1}, t) - 2u(x_i, t) + u(x_{i+1}, t)], \quad i = 1, \dots, N. \quad (3.2.4)$$

Considerando

$$\mathbf{u}(t) = \begin{bmatrix} u(x_1, t) \\ u(x_2, t) \\ \vdots \\ u(x_{N-1}, t) \\ u(x_N, t) \end{bmatrix}, \quad J = \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & \dots \\ 0 & 0 & 0 & -1 & 2 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & \dots & \dots & 1 \\ 0 & 0 & 0 & -1 & 0 \end{bmatrix}, \quad \mathbf{a}(t) = \begin{bmatrix} a(t) \\ 0 \\ \vdots \\ 0 \\ -b(t) \end{bmatrix}, \quad \mathbf{b}(t) = \begin{bmatrix} a(t) \\ 0 \\ \vdots \\ 0 \\ b(t) \end{bmatrix}$$

e usando as condições de contorno (3.1.3), podemos escrever o sistema (3.2.4) como

$$\frac{\partial \mathbf{u}(t)}{\partial t} = \frac{c}{2h}[\mathbf{a}(t) - B\mathbf{u}(t)] + \frac{v}{h^2}[\mathbf{b}(t) - J\mathbf{u}(t)], \quad (3.2.5)$$

com a condição inicial $\mathbf{u}(0) = [f(x_1), \dots, f(x_N)]^t$.

Agora, aplicando a Fórmula Trapezoidal Estendida (1.4.5) em (3.2.5) e usando as notações definidas no capítulo 1, temos

$$\begin{aligned}\hat{\mathbf{u}}_{j+2} &= (1 + 2\beta_0)\mathbf{u}_j - 2\beta_0\mathbf{u}_{j+1} + k[\beta_0\mathbf{u}'_j + (2 + \beta_0)\mathbf{u}'_{j+1}] \Rightarrow \\ \hat{\mathbf{u}}_{j+2} &= (1 + 2\beta_0)\mathbf{u}_j - 2\beta_0\mathbf{u}_{j+1} + \frac{1}{2}[\beta_0(\mathbf{c}_j - C\mathbf{u}_j) + (2 + \beta_0)(\mathbf{c}_{j+1} - C\mathbf{u}_{j+1})]\end{aligned}\quad (3.2.6)$$

e

$$\begin{aligned}\mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{k}{12}[5\mathbf{u}'_j + 8\mathbf{u}'_{j+1} - \hat{\mathbf{u}}'_{j+2}] \Rightarrow \\ \mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{1}{24}[5(\mathbf{c}_j - C\mathbf{u}_j) + 8(\mathbf{c}_{j+1} - C\mathbf{u}_{j+1}) - (\mathbf{c}_{j+2} - C\hat{\mathbf{u}}_{j+2})],\end{aligned}\quad (3.2.7)$$

onde $C = \rho B + 2rJ$ e $\mathbf{c}_j = \rho\mathbf{a}_j + 2r\mathbf{b}_j$.

Agora, substituindo $\hat{\mathbf{u}}_{j+2}$ dado por (3.2.6) em (3.2.7), temos

$$\begin{aligned}(I + \frac{1}{12}(4 + \beta_0)C + \frac{1}{48}(2 + \beta_0)C^2)\mathbf{u}_{j+1} &= (I + \frac{1}{12}(\beta_0 - 2)C - \frac{1}{48}\beta_0C^2)\mathbf{u}_j + \\ &\frac{1}{48}[(10I + \beta_0C)\mathbf{c}_j + \{16I + (2 + \beta_0)C\}\mathbf{c}_{j+1} - 2\mathbf{c}_{j+2}],\end{aligned}\quad (3.2.8)$$

onde I denota a matriz identidade.

Chamamos (3.2.8) de Fórmula Trapezoidal Estendida - Sistema de Diferenças Finitas (ETF(β_0)-FDS) para a equação de convecção-difusão (3.1.1). Vale notar que o sistema (3.2.8) é pentadiagonal.

Diferenciaremos os sistemas (ETF(β_0)-FDS) pela estabilidade e erro de truncamento local, ou seja, dependendo do valor do parâmetro β_0 teremos um sistema particular, distinguindo subclasses do sistema (3.2.8). Para $\beta_0 \geq -1$, (3.2.8) será conhecido como a ETF(β_0)-FDS A-estável (AS-ETF(β_0)), e quando $\beta_0 = 0$, será conhecido como a ETF(0)-FDS L-estável (LS-ETF(0)) para a equação de convecção-difusão.

3.2.1 Erro de Truncamento Local (ETF(β_0)-FDS)

Novamente, seguindo o raciocínio do capítulo 2, obteremos o erro de truncamento local do sistema ETF(β_0)-FDS (3.2.8). Usamos as notações definidas no capítulo 1, e definimos uma união dos operadores de diferenças finitas dada por

$$\Omega = \frac{v}{h^2}\delta^2 - \frac{c}{2h}(\Delta + \nabla).$$

A matriz C pode ser representada pela discretização do operador de diferenças $\Omega : C \rightarrow -2k\Omega$, ou seja, se fizermos $-2k\Omega$ voltamos à matriz C , e com isto segue que o operador de diferenças para o sistema (3.2.8) é dado por

$$\begin{aligned} u_{i,j+1} &= u_{i,j} + \left[-\frac{1}{12}(4 + \beta_0)C - \frac{1}{48}(2 + \beta_0)C^2\right]u_{i,j+1} + \\ &\quad \left[\frac{1}{12}(\beta_0 - 2)C - \frac{1}{48}\beta_0 C^2\right]u_{i,j} \Rightarrow \\ u_{i,j+1} - u_{i,j} &= \left[-\frac{1}{12}(4 + \beta_0)(-2k\Omega) - \frac{1}{48}(2 + \beta_0)(-2k\Omega)^2\right]u_{i,j+1} + \\ &\quad \left[\frac{1}{12}(\beta_0 - 2)(-2k\Omega) - \frac{1}{48}\beta_0(-2k\Omega)^2\right]u_{i,j} \Rightarrow \\ \frac{1}{k}[u_{i,j+1} - u_{i,j}] &= \left[\frac{1}{6}(4 + \beta_0)(\Omega) - \frac{1}{12}(2 + \beta_0)k\Omega^2\right]u_{i,j+1} + \\ &\quad \left[-\frac{1}{6}(\beta_0 - 2)(\Omega) - \frac{1}{12}\beta_0 k\Omega^2\right]u_{i,j} \Rightarrow \\ L &= \frac{1}{k}(E_t^+ - 1) + \left[-\frac{1}{6}(4 + \beta_0)\Omega + \frac{1}{12}(2 + \beta_0)k\Omega^2\right]E_t^+ + \\ &\quad \left[\frac{1}{6}(\beta_0 - 2)\Omega + \frac{1}{12}\beta_0 k\Omega^2\right] \end{aligned} \quad (3.2.9)$$

Agora, da equação (3.1.1) temos

$$D_t + cD_x = vD_x^2$$

e usando (2.2.13) e aplicando Série de Taylor nos operadores Δ e ∇ temos

$$(\Delta + \nabla) = D_x + \frac{1}{6}h^2 D_x^3 + \dots,$$

Agora, substituindo os operadores de (2.2.13) e $\Delta + \nabla$ em Ω obtemos

$$\Omega = D_t^2 + \frac{h^2}{12}D_x^3(vD_x - 2c) + \frac{h^4}{360}D_x^5(vD_x - 3c) + \dots$$

e

$$\Omega^2 = D_t^2 + \frac{h^2}{6} D_x^3 (v D_x - 2c) D_t + \dots$$

Com estes resultados, obtemos

$$\begin{aligned} L = & \frac{1}{72}(1 + \beta_0)k^3 D_t^4 + \frac{1}{720}(6 + 5\beta_0)k^4 D_t^5 - \\ & \frac{1}{12}h^2 D_x^3 (v D_x - 2c) + \frac{1}{72}(\beta_0 - 2)kh^2 D_x^3 (v D_x - 2c) D_t + \\ & \frac{1}{144}\beta_0 k^2 h^2 D_x^3 (v D_x - 2c) D_t^2 - \frac{1}{360}h^4 D_x^5 (v D_x - 3c) + \dots \end{aligned} \quad (3.2.10)$$

Assim, o sistema ETF(β_0)-FDS (3.2.8) tem ordem 3 na variável temporal.

Note que para $\beta_0 = -1$, o método terá ordem 4 na variável temporal.

3.2.2 Estabilidade do sistema ETF(β_0)-FDS

Para condições de contorno homogêneas, o sistema ETF(β_0)-FDS (3.2.8) pode ser escrito como

$$\mathbf{u}_{j+1} = \Phi \mathbf{u}_j, \quad j = 0, 1, 2, \dots \quad (3.2.11)$$

com a matriz de amplificação para o sistema de diferenças dada por

$$\Phi = \left(I + \frac{4 + \beta_0}{12} C + \frac{2 + \beta_0}{48} C^2 \right)^{-1} \left(I + \frac{\beta_0 - 2}{12} C - \frac{\beta_0}{48} C^2 \right).$$

Pela definição 1.5.1, para provarmos que o sistema ETF(β_0)-FDS é incondicionalmente estável, devemos ter que os auto-valores da Φ são em módulo menores ou iguais a um. Desde que

$$C = \begin{bmatrix} 4r & \rho - 2r & 0 & 0 & 0 \\ -(\rho + 2r) & 4r & \rho - 2r & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & -(\rho + 2r) & 4r & \rho - 2r \\ 0 & 0 & 0 & -(\rho + 2r) & 4r \end{bmatrix}$$

e pelo Teorema 1.5.1, sabe-se que os auto-valores da matriz C são dados por

$$\lambda_s(C) = 2(2r - \sqrt{4r^2 - \rho^2}) \cos\left(\frac{s\pi}{N+1}\right), \quad \rho \leq 2r$$

ou

$$\lambda_s(C) = 2(2r + \sqrt{-1}\sqrt{\rho^2 - 4r^2}) \cos\left(\frac{s\pi}{N+1}\right), \quad \rho \geq 2r.$$

Assim, os auto-valores da matriz Φ (propriedades 1.5.1), são dados por

$$\lambda_s(\Phi) = \frac{1 + \frac{1}{12}(\beta_0 - 2)\lambda_s(C) - \frac{1}{48}\beta_0(\lambda_s(C))^2}{1 + \frac{1}{12}(4 + \beta_0)\lambda_s(C) + \frac{1}{48}(2 + \beta_0)(\lambda_s(C))^2}, \quad s = 1, \dots, N. \quad (3.2.12)$$

Escrevendo $\lambda_s(C) = x + \sqrt{-1}y$, com $x > 0$ e $\lambda_s(\Phi) = \frac{Num}{Denom}$, obtemos

$$\begin{aligned} |Denom|^2 &= \left[1 + \frac{1}{12}(4 + \beta_0)x + \frac{1}{48}(2 + \beta_0)(x^2 - y^2)\right]^2 + \\ &\quad \left[\frac{1}{12}(4 + \beta_0)y + \frac{1}{24}(2 + \beta_0)xy\right]^2 \end{aligned}$$

e

$$\begin{aligned} |Num|^2 &= \left[1 + \frac{1}{12}(\beta_0 - 2)x - \frac{1}{48}\beta_0(x^2 - y^2)\right]^2 + \\ &\quad \left[\frac{1}{12}(\beta_0 - 2)y - \frac{1}{24}\beta_0xy\right]^2. \end{aligned}$$

Assim, $|\lambda_s(\Phi)| \leq 1$ quando

$$|Denom|^2 - |Num|^2 \geq 0.$$

Calculamos então,

$$\begin{aligned} |Denom|^2 - |Num|^2 &= \frac{1}{576}[\{12 + (\beta_0 + 1)x\}\{48 + 8(\beta_0 + 1)x + x^2\} + \\ &\quad (\beta_0 + 1)y^4 + 2xy^2\{6 + 2(\beta_0 + 1)\} + \\ &\quad (\beta_0 + 1)x\}]. \end{aligned} \quad (3.2.13)$$

Segue que $|\lambda_s(\Phi)| \leq 1$, com $x > 0$ e $\forall \rho, r > 0$, e então o sistema ETF(β_0)-FDS será incondicionalmente estável $\forall \beta_0 \geq -1$.

3.2.3 Sistemas ETF(β_0)-FDS diferenciados pelo valor de β_0

Pela discussão anterior de erro de truncamento local e estabilidade, segue três métodos ETF(β_0)-FDS que merecem destaque especial.

(i) **LS-ETF(0)**: sistema incondicionalmente estável, de ordem 3 na variável temporal, com o método aplicado (ETF(0)) L-estável.

$$(I + \frac{1}{3}C + \frac{1}{24}C^2)\mathbf{u}_{j+1} = (I - \frac{1}{6}C)\mathbf{u}_j + \frac{1}{24}[5\mathbf{c}_j + (8I + C)\mathbf{c}_{j+1} - \mathbf{c}_{j+2}] \quad (3.2.14)$$

(ii) **AS-ETF(-1)**: sistema incondicionalmente estável, de ordem 4 na variável temporal, com o método aplicado (ETF(-1)) A-estável.

$$(I + \frac{1}{4}C + \frac{1}{48}C^2)\mathbf{u}_{j+1} = (I - \frac{1}{4}C + \frac{1}{48}C^2)\mathbf{u}_j + \frac{1}{48}[(10I - C)\mathbf{c}_j + (16I + C)\mathbf{c}_{j+1} - 2\mathbf{c}_{j+2}] \quad (3.2.15)$$

(iii) **AS-ETF($2(1 + \sqrt{3})$)** com a matriz dos coeficientes fatorada: a matriz pentadiagonal, no sistema (3.2.8) pode ser expressa como o quadrado de uma matriz tridiagonal para o valor de $\beta_0 = 2(1 + \sqrt{3})$, como segue:

$$I + \frac{1}{12}(4 + 2(1 + \sqrt{3}))C + \frac{1}{48}(2 + 2(1 + \sqrt{3}))C^2 = \\ [I + (\frac{6 + 2\sqrt{3}}{12})C + (\frac{4 + 2\sqrt{3}}{48})C^2] = [I + (\frac{3 + \sqrt{3}}{12})C]^2$$

Para este valor de β_0 , ETF(β_0) é A-estável e o sistema (3.2.8) é incondicionalmente estável. Então, neste caso, para o sistema (3.2.8) necessitamos resolver um sistema linear tridiagonal duas vezes, com a mesma matriz dos coeficientes, para cada passo de integração.

3.2.4 Condições de Contorno de Newman

Aplicamos, aqui, a Fórmula Trapezoidal Estendida - Sistema de Diferenças Finitas para a equação de convecção-difusão (3.1.1), com a mesma condição inicial (3.1.2), mas com condições de contorno de Newman da forma:

$$\frac{\partial u(0, t)}{\partial x} = a_1 u(t) - b_1 \quad \text{e} \quad \frac{\partial u(l, t)}{\partial x} = -a_2 u(t) + b_2, \quad (3.2.16)$$

onde a_1, a_2, b_1, b_2 são constantes não-negativas. Primeiramente, consideramos a discretização das condições de contorno para $x = 0$. Introduzimos o ponto fictício x_{-1} , e substituindo a derivada de primeira ordem pelo operador de diferenças centradas, obtemos

$$\begin{aligned} \frac{\partial u(0, t)}{\partial x} &= a_1 u_0(t) - b_1 \quad \text{e} \quad \frac{\partial u(x_i, t)}{\partial x} = \frac{1}{2h} [u(x_{i+1}, t) - u(x_{i-1}, t)] \Rightarrow \\ a_1 u_0(t) - b_1 &= \frac{1}{2h} [u_1(t) - u_{-1}(t)]. \end{aligned} \quad (3.2.17)$$

Usando a equação (3.2.17) e eliminando u_{-1} da equação (3.2.4), para $i = 0$, temos:

$$\begin{aligned} \frac{\partial u_0(t)}{\partial t} &= \frac{v}{h^2} [u_{-1}(t) - 2u_0(t) + u_1(t)] - \frac{c}{2h} [u_1(t) - u_{-1}(t)] \Rightarrow \\ \frac{\partial u_0(t)}{\partial t} &= \frac{v}{h^2} [2u_1(t) - (2 + 2ha_1)u_0(t) + 2hb_1] + \frac{c}{2h} [2hb_1 - 2ha_1 u_0(t)]. \end{aligned} \quad (3.2.18)$$

De maneira análoga, para condições de contorno em $x = l$, temos:

$$\frac{\partial u_{N+1}(t)}{\partial t} = \frac{c}{2h} [2ha_2 u_{N+1}(t) - 2hb_2] + \frac{v}{h^2} [2u_N(t) - (2 + 2ha_2)u_{N+1}(t) + 2hb_2]. \quad (3.2.19)$$

Consideremos

$$\mathbf{u}(t) = \begin{bmatrix} u(x_0, t) \\ u(x_1, t) \\ \vdots \\ u(x_N, t) \\ u(x_{N+1}, t) \end{bmatrix}, \quad P = \begin{bmatrix} 2ha_1 & 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & -1 & 0 & 1 \\ 0 & 0 & 0 & 0 & -2ha_2 \end{bmatrix},$$

$$Q = \begin{bmatrix} 2 + 2ha_1 & -2 & 0 & 0 & 0 \\ -1 & 2 & -1 & 0 & 0 \\ 0 & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & -1 & 2 & -1 \\ 0 & 0 & 0 & -2 & 2 + 2ha_2 \end{bmatrix},$$

$$\alpha = \begin{bmatrix} 2hb_1 \\ 0 \\ \vdots \\ 0 \\ -2hb_2 \end{bmatrix}, \quad \beta = \begin{bmatrix} 2hb_1 \\ 0 \\ \vdots \\ 0 \\ 2hb_2 \end{bmatrix}.$$

Então, podemos escrever (3.2.18), (3.2.4) (para $i = 1, \dots, N$) e (3.2.19) como

$$\frac{\partial \mathbf{u}(t)}{\partial t} = \frac{c}{2h}[\alpha - P\mathbf{u}(t)] + \frac{v}{h^2}[\beta - Q\mathbf{u}(t)]. \quad (3.2.20)$$

Agora, aplicando a Fórmula Trapezoidal Estendida (1.4.5) para (3.2.20) temos

$$\begin{aligned} \hat{\mathbf{u}}_{j+2} &= (1 + 2\beta_0)\mathbf{u}_j - 2\beta_0\mathbf{u}_{j+1} + k[\beta_0\mathbf{u}'_j + (2 + \beta_0)\mathbf{u}'_{j+1}] \Rightarrow \\ \hat{\mathbf{u}}_{j+2} &= [(1 + 2\beta_0)I - \frac{1}{2}\beta_0 M]\mathbf{u}_j - [2\beta_0 I + \frac{1}{2}(2 + \beta_0)M]\mathbf{u}_{j+1} + (1 + \beta_0)\gamma \end{aligned} \quad (3.2.21)$$

e

$$\begin{aligned} \mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{k}{12}[5\mathbf{u}'_j + 8\mathbf{u}'_{j+1} - \hat{\mathbf{u}}'_{j+2}] \Rightarrow \\ \mathbf{u}_{j+1} &= \mathbf{u}_j - \frac{1}{24}M(5\mathbf{u}_j + 8\mathbf{u}_{j+1} - \hat{\mathbf{u}}_{j+2}) + \frac{1}{2}\gamma, \end{aligned} \quad (3.2.22)$$

onde $M = \rho P + 2rQ$ e $\gamma = \rho\alpha + 2r\beta$.

Substituindo $\hat{\mathbf{u}}_{j+2}$ de (3.2.21) em (3.2.22), obtemos

$$\begin{aligned} (I + \frac{1}{12}(\beta_0 - 2)M + \frac{1}{48}(2 + \beta_0)M^2)\mathbf{u}_{j+1} &= (I + \frac{1}{12}(\beta_0 - 2)M - \frac{1}{48}\beta_0 M^2)\mathbf{u}_j + \\ &\quad \frac{1}{2}[I + \frac{1}{12}(1 + \beta_0)M]\gamma. \end{aligned} \quad (3.2.23)$$

3.3 Regra de Simpson Estendida (ESR(α_0))

Analogamente à Seção 3.2, aplicamos a Regra de Simpson Estendida (1.4.7) para (3.2.5) e usando as notações definidas no capítulo 1, temos

$$\begin{aligned}\tilde{\mathbf{u}}_{j+\frac{1}{2}} &= \frac{1}{4}\mathbf{u}_j + \frac{3}{4}\mathbf{u}_{j+1} - \frac{k}{4}\mathbf{u}'_{j+1} \Rightarrow \\ \tilde{\mathbf{u}}_{j+\frac{1}{2}} &= \frac{1}{4}\mathbf{u}_j + \frac{1}{8}(6I + C)\mathbf{u}_{j+1} - \frac{1}{8}\mathbf{c}_{j+1},\end{aligned}\quad (3.3.24)$$

$$\begin{aligned}\hat{\mathbf{u}}_{j+\frac{1}{2}} &= \alpha_0\mathbf{u}_j + (1 - \alpha_0)\mathbf{u}_{j+1} + \frac{k}{24}[(1 + 4\alpha_0)\mathbf{u}'_j + \\ &\quad 8(2\alpha_0 - 1)\tilde{\mathbf{u}}'_{j+\frac{1}{2}} + (4\alpha_0 - 5)\mathbf{u}'_{j+1}] \Rightarrow \\ \hat{\mathbf{u}}_{j+\frac{1}{2}} &= \alpha_0\mathbf{u}_j + (1 - \alpha_0)\mathbf{u}_{j+1} + \frac{1}{48}[(1 + 4\alpha_0)(\mathbf{c}_j - C\mathbf{u}_j) + \\ &\quad 8(2\alpha_0 - 1)(\mathbf{c}_{j+\frac{1}{2}} - C\tilde{\mathbf{u}}_{j+\frac{1}{2}}) + (4\alpha_0 - 5)(\mathbf{c}_{j+1} - C\mathbf{u}_{j+1})]\end{aligned}\quad (3.3.25)$$

e

$$\begin{aligned}\mathbf{u}_{j+1} &= \mathbf{u}_j - \frac{h}{6}[\mathbf{u}'_j + 4\hat{\mathbf{u}}'_{j+\frac{1}{2}} + \mathbf{u}'_{j+1}] \Rightarrow \\ \mathbf{u}_{j+1} &= \mathbf{u}_j - \frac{1}{12}C(\mathbf{u}_j + 4\hat{\mathbf{u}}_{j+\frac{1}{2}} + \mathbf{u}_{j+1}) + \frac{1}{12}(\mathbf{c}_j + 4\mathbf{c}_{j+\frac{1}{2}} + \mathbf{c}_{j+1}),\end{aligned}\quad (3.3.26)$$

onde $C = \rho B + 2rJ$ e $\mathbf{c}_j = \rho\mathbf{a}_j + 2r\mathbf{b}_j$.

Agora, substituindo $\tilde{\mathbf{u}}_{j+\frac{1}{2}}$ dado por (3.3.24) em (3.3.25) obtemos

$$\begin{aligned}\hat{\mathbf{u}}_{j+\frac{1}{2}} &= \frac{1}{48}(48\alpha_0I - (8\alpha_0 - 1)C)\mathbf{u}_j + \frac{1}{48}(48(1 - \alpha_0)I - \\ &\quad (16\alpha_0 - 11)C - (2\alpha_0 - 1)C^2)\mathbf{u}_{j+1} + \frac{1}{48}(1 + 4\alpha_0)\mathbf{c}_j + \\ &\quad \frac{1}{6}(2\alpha_0 - 1)\mathbf{c}_{j+\frac{1}{2}} + \frac{1}{48}((4\alpha_0 - 5)I + (2\alpha_0 - 1)\mathbf{c}_{j+1}).\end{aligned}\quad (3.3.27)$$

Então, substituindo $\hat{\mathbf{u}}_{j+\frac{1}{2}}$ dado por (3.3.27) em (3.3.26), obtemos

$$\begin{aligned}(I + \frac{1}{12}(5 - 4\alpha_0)C + \frac{1}{144}(11 - 16\alpha_0)C^2 + \frac{1}{144}(1 - 2\alpha_0)C^3)\mathbf{u}_{j+1} = \\ (I - \frac{1}{12}(1 + 4\alpha_0)C + \frac{1}{144}(8\alpha_0 - 1)C^2)\mathbf{u}_j + \frac{1}{12}[I - \frac{1}{12}(1 + 4\alpha_0)C]\mathbf{c}_j \\ + \frac{1}{3}[I + \frac{1}{6}(1 - 2\alpha_0)C]\mathbf{c}_{j+\frac{1}{2}} + \frac{1}{12}[I + \frac{1}{12}(5 - 4\alpha_0)C + \frac{1}{12}(1 - 2\alpha_0)C^2]\mathbf{c}_{j+1},\end{aligned}\quad (3.3.28)$$

onde I denota a matriz identidade.

Chamamos (3.3.28) de Regra de Simpson Estendida (ESR(α_0)) para a equação de convecção-difusão (3.1.1). Vale notar que o sistema (3.3.28) é septadiagonal.

Diferenciamos os sistemas ESR(α_0) pela estabilidade e erro de truncamento local, ou seja, dependendo do valor do parâmetro α_0 teremos um método particular, distinguindo subclasses do sistema (3.3.28). Para $\alpha_0 < \frac{1}{2}$, (3.3.28) será conhecido como a ESR(α_0) L-estável (LS-ESR(α_0)), e quando $\alpha_0 \leq \frac{1}{2}$, será conhecido como a ESR(α_0) A-estável (AS-ESR(α_0)) para a equação de convecção-difusão, em especial tomamos $\alpha_0 = \frac{1}{2}$.

3.3.1 Erro de Truncamento Local (ESR(α_0))

Novamente, seguindo raciocínio do capítulo 2, obteremos o Erro de Truncamento Local do sistema ESR(α_0) (3.3.28). Usamos as notações definidas no capítulo 1, e definimos uma união dos operadores de diferenças dada por

$$\Omega = \frac{v}{h^2}\delta^2 - \frac{c}{2h}(\Delta + \nabla).$$

A matriz C pode ser representada pela discretização do operador de diferenças $\Omega : C \rightarrow -2k\Omega$, ou seja, se fizermos $-2k\Omega$ voltamos à matriz C e assim segue que o operador de diferenças para o sistema (3.3.28) é dado por

$$\begin{aligned} L = & \frac{1}{k}(E_t^+ - 1) - \frac{1}{36}[6(5 - 4\alpha_0)\Omega - (11 - 16\alpha_0)k\Omega^2 + 2(1 - 2\alpha_0)k^2\Omega^3]E_t^+ \\ & - \frac{1}{36}[6(1 + 4\alpha_0)\Omega + (8\alpha_0 - 1)k\Omega^2] \end{aligned} \quad (3.3.29)$$

Agora, da equação (3.3.1) temos

$$D_t + cD_x = vD_x^2$$

e usando (2.2.13) e

$$(\Delta + \nabla) = D_x + \frac{1}{6}h^2D_x^3 + \dots,$$

temos que

$$\begin{aligned}\Omega &= D_t^2 + \frac{h^2}{12}D_x^3(vD_x - 2c) + \frac{h^4}{360}D_x^5(vD_x - 3c) + \dots, \\ \Omega^2 &= D_t^2 + \frac{h^2}{6}D_x^3(vD_x - 2c)D_t + \frac{h^4}{180}D_x^5(vD_x - 3c)D_t + \dots\end{aligned}$$

e

$$\Omega^3 = D_t^3 + \frac{h^2}{4}D_x^3(vD_x - 2c)D_t^2 + \frac{h^4}{120}D_x^5(vD_x - 3c)D_t^2 + \frac{h^4}{72}D_x^6(vD_x - 2c)^2D_t + \dots$$

Com estes resultados, obtemos

$$\begin{aligned}L &= \frac{1}{2160}(20\alpha_0 - 7)k^4D_t^5 + \frac{1}{180}(\alpha_0 - 1)k^5D_t^6 - \\ &\quad \frac{1}{12}h^2D_x^3(vD_x - 2c) - \frac{1}{72}(1 + 4\alpha_0)kh^2D_x^3(vD_x - 2c)D_t + \\ &\quad \frac{1}{432}(1 - 8\alpha_0)k^2h^2D_x^3(vD_x - 2c)D_t^2 - \frac{1}{360}h^4D_x^5(vD_x - 3c) + \dots\end{aligned}\tag{3.3.30}$$

Assim, o sistema $\text{ESR}(\alpha_0)$ (3.3.28) tem ordem 4 na variável temporal. Note que para $\alpha_0 = \frac{7}{20}$, o método terá ordem 5 na variável temporal.

3.3.2 Estabilidade do método $\text{ESR}(\alpha_0)$

Para condições de contorno homogêneas, o sistema ESR (3.3.28) pode ser escrito como

$$\mathbf{u}_{j+1} = \Phi \mathbf{u}_j, \quad j = 0, 1, 2, \dots\tag{3.3.31}$$

com a matriz de amplificação para o sistema de diferenças dada por

$$\Phi = \left(I + \frac{5 - 4\alpha_0}{12}C + \frac{11 - 16\alpha_0}{144}C^2 + \frac{1 - 2\alpha_0}{144}C^3\right)^{-1} \left(I - \frac{1 + 4\alpha_0}{12}C + \frac{8\alpha_0 - 1}{144}C^2\right).$$

Pela Definição 1.5.1, para provarmos que o sistema $\text{ESR}(\alpha_0)$ é incondicionalmente estável devemos ter que os auto-valores da matriz Φ são em módulo menores ou iguais a um. Desde que

$$C = \begin{bmatrix} 4r & \rho - 2r & 0 & 0 & 0 \\ -(\rho + 2r) & 4r & \rho - 2r & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & -(\rho + 2r) & 4r & \rho - 2r \\ 0 & 0 & 0 & -(\rho + 2r) & 4r \end{bmatrix}$$

e pelo Teorema 1.5.1, sabe-se que os auto-valores da matriz C são dados por

$$\lambda_s(C) = 2(2r - \sqrt{4r^2 - \rho^2}) \cos\left(\frac{s\pi}{N+1}\right), \quad \rho \leq 2r$$

ou

$$\lambda_s(C) = 2(2r + \sqrt{-1}\sqrt{\rho^2 - 4r^2}) \cos\left(\frac{s\pi}{N+1}\right), \quad \rho \geq 2r.$$

Assim, os auto-valores da matriz Φ (propriedade 1.5.1) são dados por

$$\lambda_s(\Phi) = \frac{144 - 12(1 + 4\alpha_0)\lambda_s(C) + (8\alpha_0 - 1)(\lambda_s(C))^2}{144 + 12(5 - 4\alpha_0)\lambda_s(C) + (11 - 16\alpha_0)(\lambda_s(C))^2 + (1 - 2\alpha_0)(\lambda_s(C))^3}, \quad s = 1, \dots, N. \quad (3.3.32)$$

Escrevendo $\lambda_s(C) = x + \sqrt{-1}y$ e $\lambda_s(\Phi) = \frac{Num}{Denom}$, obtemos

$$|Denom|^2 = [144 + 12(5 - 4\alpha_0)x + (11 - 16\alpha_0)(x^2 - y^2) + (x^3 - 3xy^2)(1 - 2\alpha_0)]^2 + [12(5 - 4\alpha_0)y + 2(11 - 16\alpha_0)xy + (1 - 2\alpha_0)(3x^2y - y^3)]^2$$

e

$$|Num|^2 = [144 - 12(1 + 4\alpha_0)x + (8\alpha_0 - 1)(x^2 - y^2)]^2 + [-12(1 + 4\alpha_0)y + 2(8\alpha_0 - 1)xy]^2.$$

Assim, $|\lambda_s(\Phi)| \leq 1$ quando

$$|Denom|^2 - |Num|^2 \geq 0.$$

Notamos que, para os dois casos dos $\lambda_s(C)$, com $x > 0$, segue que $|\lambda_s(\Phi)| \leq 1$, $\forall \rho, r > 0$, e então o método $ESR(\alpha_0)$ será incondicionalmente estável $\forall \alpha_0 \leq \frac{1}{2}$.

3.3.3 Sistemas ESR(α_0) diferenciados pelo valor de α_0

Pela discussão anterior de erro de truncamento local e estabilidade, segue dois métodos ESR(α_0) que merecem destaque especial.

(i) **LS-ESR**($\frac{7}{20}$): sistema incondicionalmente estável, de ordem 5 na variável temporal, com o método aplicado (ESR($\frac{7}{20}$)) L-estável.

$$\begin{aligned} (I + \frac{3}{10}C + \frac{3}{80}C^2 + \frac{1}{480}C^3)\mathbf{u}_{j+1} = & (I - \frac{1}{5}C + \frac{1}{80}C^2)\mathbf{u}_j \\ & + \frac{1}{12}[I - \frac{1}{5}C]\mathbf{c}_j + \frac{1}{3}[I + \frac{1}{20}C]\mathbf{c}_{j+\frac{1}{2}} + \frac{1}{12}[I + \frac{3}{10}C + \frac{1}{40}C^2]\mathbf{c}_{j+1}. \end{aligned} \quad (3.3.33)$$

Esta matriz septadiagonal do sistema (3.3.33), pode ser fatorada como

$$\begin{aligned} (I + \frac{3}{10}C + \frac{3}{80}C^2 + \frac{1}{480}C^3) = \\ [I + \frac{1}{60}(6 + 3s - s^2)C] \times [I + \frac{1}{60}(12 - 3s + s^2)C + \frac{1}{240}(3 - s + s^2)C^2] \end{aligned}$$

onde $s = \sqrt[3]{3}$. Desta forma, para o sistema (3.3.33) resolvemos um sistema linear tridiagonal e um pentadiagonal a cada tempo de integração.

(ii) **AS-ESR**($\frac{1}{2}$): sistema incondicionalmente estável, de ordem 4 na variável temporal, com o método aplicado (ESR($\frac{1}{2}$)) A-estável.

$$\begin{aligned} (I + \frac{1}{4}C + \frac{1}{48}C^2)\mathbf{u}_{j+1} = & (I - \frac{1}{4}C + \frac{1}{48}C^2)\mathbf{u}_j + \\ & \frac{1}{48}[(4I - C)\mathbf{c}_j + 16\mathbf{c}_{j+\frac{1}{2}} + (4I + C)\mathbf{c}_{j+1}] \end{aligned} \quad (3.3.34)$$

3.3.4 Condições de Contorno de Newman

Analogamente à Seção 3.2.4, aplicamos a Regra de Simpson Estendida (1.4.7) para (3.2.20) e obtemos

$$\begin{aligned}\tilde{\mathbf{u}}_{j+\frac{1}{2}} &= \frac{1}{4}\mathbf{u}_j + \frac{3}{4}\mathbf{u}_{j+1} - \frac{k}{4}\mathbf{u}'_{j+1} \Rightarrow \\ \tilde{\mathbf{u}}_{j+\frac{1}{2}} &= \frac{1}{4}\mathbf{u}_j + \frac{1}{8}(6I + M)\mathbf{u}_{j+1} - \frac{1}{8}\gamma,\end{aligned}\quad (3.3.35)$$

$$\begin{aligned}\hat{\mathbf{u}}_{j+\frac{1}{2}} &= \alpha_0\mathbf{u}_j + (1 - \alpha_0)\mathbf{u}_{j+1} + \frac{k}{24}[(1 + 4\alpha_0)\mathbf{u}'_j + \\ &\quad 8(2\alpha_0 - 1)\tilde{\mathbf{u}}'_{j+\frac{1}{2}} + (4\alpha_0 - 5)\mathbf{u}'_{j+1}] \Rightarrow \\ \hat{\mathbf{u}}_{j+\frac{1}{2}} &= \alpha_0\mathbf{u}_j + (1 - \alpha_0)\mathbf{u}_{j+1} - \frac{1}{48}M[(1 + 4\alpha_0)\mathbf{u}_j + \\ &\quad 8(2\alpha_0 - 1)\tilde{\mathbf{u}}_{j+\frac{1}{2}} + (4\alpha_0 - 5)\mathbf{u}_{j+1}] - \frac{1}{4}(1 - 2\alpha_0)\gamma\end{aligned}\quad (3.3.36)$$

e

$$\begin{aligned}\mathbf{u}_{j+1} &= \mathbf{u}_j + \frac{k}{6}[\mathbf{u}'_j + 4\hat{\mathbf{u}}'_{j+\frac{1}{2}} + \mathbf{u}'_{j+1}] \Rightarrow \\ \mathbf{u}_{j+1} &= \mathbf{u}_j - \frac{1}{12}M(\mathbf{u}_j + 4\hat{\mathbf{u}}_{j+\frac{1}{2}} + \mathbf{u}_{j+1}) + \frac{1}{2}\gamma,\end{aligned}\quad (3.3.37)$$

onde $M = \rho P + 2rQ$ e $\gamma = \rho\alpha + 2r\beta$.

Substituindo $\tilde{\mathbf{u}}_{j+\frac{1}{2}}$ dado por (3.3.35) em (3.3.36), obtemos

$$\begin{aligned}\hat{\mathbf{u}}_{j+\frac{1}{2}} &= (\alpha_0 I - \frac{1}{48}(8\alpha_0 - 1)M)\mathbf{u}_j + ((1 - \alpha_0)I - \frac{1}{48}(16\alpha_0 - 11)M - \\ &\quad \frac{1}{48}(2\alpha_0 - 1)M^2)\mathbf{u}_{j+1} + \frac{1}{48}(2\alpha_0 - 1)[12I + M]\gamma.\end{aligned}\quad (3.3.38)$$

Agora, substituindo $\hat{\mathbf{u}}_{j+\frac{1}{2}}$ de (3.3.38) em (3.3.37), obtemos

$$\begin{aligned}&(I + \frac{1}{12}(5 - 4\alpha_0)M + \frac{1}{144}(11 - 16\alpha_0)M^2 + \frac{1}{144}(1 - 2\alpha_0)M^3)\mathbf{u}_{j+1} \\ = &(I - \frac{1}{12}(1 + 4\alpha_0)M + \frac{1}{144}(8\alpha_0 - 1)M^2)\mathbf{u}_j + \\ &\frac{1}{2}[I + \frac{1}{6}(1 - 2\alpha_0)M + \frac{1}{72}(1 - 2\alpha_0)M^2]\gamma.\end{aligned}\quad (3.3.39)$$

3.4 Fórmula Trapezoidal Generalizada - Sistema de Diferenças Finitas(GTF(α)-FDS)

Analogamente à Seção 3.2, aplicamos a Fórmula Trapezoidal Generalizada (GTF(α)) (1.4.6) em (3.2.5) e usando as notações definidas no capítulo 1, temos

$$\hat{\mathbf{u}}_j = \mathbf{u}_{j+1} - k\mathbf{u}'_{j+1} \Rightarrow$$

$$\hat{\mathbf{u}}_j = \mathbf{u}_{j+1} - k\left[\frac{c}{2h}\{\mathbf{a}_{j+1} - B\mathbf{u}_{j+1}\} + \frac{v}{h^2}\{\mathbf{b}_{j+1} - J\mathbf{u}_{j+1}\}\right]$$

e

$$\mathbf{u}_{j+1} = \mathbf{u}_j + \frac{k}{2}[(1 - \alpha)\mathbf{u}'_j + \alpha\hat{\mathbf{u}}'_j + \mathbf{u}'_{j+1}] \Rightarrow$$

$$\mathbf{u}_{j+1} = \mathbf{u}_j + \frac{k}{2}[(1 - \alpha)\left\{\frac{c}{2h}(\mathbf{a}_j - B\mathbf{u}_{j+1}) + \frac{v}{h^2}(\mathbf{b}_j - J\mathbf{u}_j)\right\}$$

$$+ \alpha\left\{\frac{c}{2h}(\mathbf{a}_j - B\hat{\mathbf{u}}_j) + \frac{v}{h^2}(\mathbf{b}_j - J\hat{\mathbf{u}}_j)\right\}$$

$$+ \left\{\frac{c}{2h}(\mathbf{a}_{j+1} - B\mathbf{u}_{j+1}) + \frac{v}{h^2}(\mathbf{b}_{j+1} - J\mathbf{u}_{j+1})\right\}], \quad (3.4.40)$$

com I matriz identidade e $\alpha \in [0, 1]$. Seja $A = \frac{\rho}{2}B + rJ$ e $\mathbf{c}_j = \frac{\rho}{2}\mathbf{a}_j + r\mathbf{b}_j$, então

$$\hat{\mathbf{u}}_j = (I + A)\mathbf{u}_{j+1} - \mathbf{c}_{j+1} \quad (3.4.41)$$

e

$$\mathbf{u}_{j+1} = [I - \frac{1}{2}(1 - \alpha)A]\mathbf{u}_j - \frac{1}{2}A\mathbf{u}_{j+1} - \frac{1}{2}\alpha A\hat{\mathbf{u}}_j + \frac{1}{2}(\mathbf{c}_j + \mathbf{c}_{j+1}). \quad (3.4.42)$$

Agora, substituindo $\hat{\mathbf{u}}_j$ dado por (3.4.41) em (3.4.42) obtemos

$$(I + \frac{1}{2}(1 + \alpha)A + \frac{1}{2}\alpha A^2)\mathbf{u}_{j+1} = (I - \frac{1}{2}(1 - \alpha)A)\mathbf{u}_j + \frac{1}{2}[\mathbf{c}_j + (I + \alpha A)\mathbf{c}_{j+1}]. \quad (3.4.43)$$

Chamamos (3.4.43) de Fórmula Trapezoidal Generalizada - Sistema de Diferenças Finitas (GTF(α)-FDS) para a equação de convecção-difusão (3.1.1). Vale notar que para $\alpha = 0$, GTF(0)-FDS é a Fórmula Trapezoidal Clássica (AM-TF) para a equação (3.1.1), e é um sistema equivalente ao método de Crank-Nicolson; incluindo $\rho = 0$, torna-se o conhecido método de Crank-Nicolson para a equação (2.1.1).

3.4.1 Erro de Truncamento Local (GTF(α)-FDS)

Novamente, para obtermos o erro de truncamento local do sistema GTF(α)-FDS (3.4.43), usamos os mesmos procedimentos do capítulo 2. Definimos a união dos operadores de diferenças finitas dados por

$$M = \frac{v}{h^2}\delta^2 - \frac{c}{2h}(\Delta + \nabla).$$

Então, podemos escrever

$$\hat{u}_{i,j} = (1 - kM)u_{i,j+1}. \quad (3.4.44)$$

Assim, o operador de diferenças que define o sistema de diferenças finitas (3.4.43) é dado por

$$L = \frac{1}{k}(E_t^+ - 1) - \frac{1}{2}[(1 - \alpha)M + \{(1 + \alpha)M - \alpha kM^2\}E_t^+]. \quad (3.4.45)$$

Agora, da equação (3.1.1) temos

$$D_t + cD_x = vD_x^2$$

e usando (2.2.13) e

$$(\Delta + \nabla) = D_x + \frac{1}{6}h^2D_x^3 + \dots,$$

temos que

$$M = D_t + \frac{h^2}{12}D_x^3(vD_x - 2c) + \frac{h^4}{360}D_x^5(vD_x - 3c) + \dots$$

e

$$M^2 = D_t^2 + \frac{h^2}{6}D_x^3(vD_x - 2c)D_t + \dots$$

Com estes resultados, obtemos

$$\begin{aligned} L = & \frac{1}{12}(3\alpha - 1)k^2D_t^3 + \frac{1}{24}(4\alpha - 1)k^3D_t^4 - \\ & \frac{1}{12}h^2D_x^3(vD_x - 2c) - \frac{1}{24}(1 - \alpha)kh^2D_x^3(vD_x - 2c)D_t + \\ & \frac{1}{48}(3\alpha - 1)k^2h^2D_x^3(vD_x - 2c)D_t^2 - \frac{1}{360}h^4D_x^5(vD_x - 3c) + \dots \end{aligned} \quad (3.4.46)$$

Assim, o sistema GTF(α)-FDS (3.4.43) tem ordem 2 em ambas variáveis temporal e espacial. Note que para $\alpha = \frac{1}{3}$, o sistema terá ordem 3 na variável temporal.

3.4.2 Estabilidade do sistema GTF(α)-FDS

Como nas Seções anteriores, para condições de contorno homogêneas, o sistema GTF(α)-FDS (3.4.43) pode ser escrito como

$$\mathbf{u}_{j+1} = Q\mathbf{u}_j, \quad j = 0, 1, 2, \dots \quad (3.4.47)$$

com a matriz de amplificação para o método de diferenças dada por

$$Q = (I + \frac{1}{2}(1 + \alpha)A + \frac{1}{2}\alpha A^2)^{-1}(I - \frac{1}{2}(1 - \alpha)A).$$

Pela definição 1.5.1, para provarmos que o método GTF(α)-FDS é incondicionalmente estável devemos ter que os auto-valores da matriz Q são em módulo menores ou iguais a um. Desde que

$$A = \begin{bmatrix} 2r & -r + \frac{\rho}{2} & 0 & 0 & 0 \\ -r - \frac{\rho}{2} & 2r & -r + \frac{\rho}{2} & 0 & 0 \\ 0 & \dots & \dots & \dots & 0 \\ 0 & 0 & -r - \frac{\rho}{2} & 2r & -r + \frac{\rho}{2} \\ 0 & 0 & 0 & -r - \frac{\rho}{2} & 2r \end{bmatrix}$$

e pelo Teorema 1.5.1, sabe-se que os auto-valores da matriz A são dados por

$$\lambda_s(A) = 2r - \sqrt{4r^2 - \rho^2} \cos\left(\frac{s\pi}{N+1}\right), \quad \rho \leq 2r$$

ou

$$\lambda_s(A) = 2r + \sqrt{-1}\sqrt{\rho^2 - 4r^2} \cos\left(\frac{s\pi}{N+1}\right), \quad \rho \geq 2r.$$

Assim, os auto-valores da matriz Q (Propriedades 1.5.1) são dados por

$$\lambda_s(Q) = \frac{2 - (1 - \alpha)\lambda_s(A)}{2 + (1 + \alpha)\lambda_s(A) + \alpha(\lambda_s(A))^2}, \quad s = 1, \dots, N. \quad (3.4.48)$$

Escrevendo $\lambda_s(A) = -z$, com $z = x + y\sqrt{-1}$, $Re(z) < 0$, e $\lambda_s(Q) = \frac{Num}{Denom}$, obtemos

$$|Denom|^2 \geq [2 + (1 + \alpha)x + \alpha x^2]^2 + [(1 - \alpha)y^2]$$

e

$$|Num|^2 = [2 - (1 - \alpha)x]^2 + [(1 - \alpha_0)^2 y^2].$$

Assim, $|\lambda_s(\Phi)| \leq 1$ quando

$$|Denom|^2 - |Num|^2 \geq 0.$$

Notamos que para os dois casos dos $\lambda_s(A)$, $Re(z) < 0$, segue que $|\lambda_s(Q)| \leq 1$, $\forall \rho, r > 0$, e $\forall \alpha \in [0, 1]$. A igualdade só é válida para $\alpha = 0$ e $r = 0$. Então, o método GTF(α)-FDS será incondicionalmente estável $\forall r, \rho > 0$.

3.4.3 Transformação da Equação de Convecção-Difusão para a Equação de Difusão

Notemos que a equação de convecção-difusão (3.1.1) pode ser transformada na equação de difusão pura (2.1.1) através da substituição

$$u(x, t) = \exp\left(\frac{c}{2v}(x - \frac{c}{2}t)\right)v(x, t). \quad (3.4.49)$$

Ou seja, a equação de convecção-difusão (3.1.1) transforma-se na equação de difusão pura

$$\frac{\partial v}{\partial t} = v \frac{\partial^2 v}{\partial x^2}. \quad (3.4.50)$$

A condição inicial (3.1.2) torna-se agora

$$v(x, 0) = \exp\left(-\frac{c}{2v}x\right)f(x) \quad (3.4.51)$$

e as condições de contorno (3.1.3) mudam para

$$v(0, t) = \exp\left(\frac{c^2}{4v}t\right)a(t), \quad v(l, t) = \exp\left(-\frac{c}{2v}(l - \frac{c}{2}t)\right)b(t). \quad (3.4.52)$$

De fato, calculamos

$$\begin{aligned}\frac{\partial u}{\partial t} &= \exp\left(\frac{c}{2v}\left(x - \frac{c}{2}t\right)\right)\left(-\frac{c}{2}\frac{c}{2v}\right)v(x, t) + \exp\left(\frac{c}{2v}\left(x - \frac{c}{2}t\right)\right) \\ \frac{\partial u}{\partial x} &= \exp\left(\frac{c}{2v}\left(x - \frac{c}{2}t\right)\right)\frac{c}{2v}v(x, t) + \exp\left(\frac{c}{2v}\left(x - \frac{c}{2}t\right)\right)\frac{\partial v}{\partial x} \\ \frac{\partial^2 u}{\partial x^2} &= \frac{c^2}{4v^2}\exp\left(\frac{c}{2v}\left(x - \frac{c}{2}t\right)\right)v(x, t) + \frac{c}{2v}\exp\left(\frac{c}{2v}\left(x - \frac{c}{2}t\right)\right)\frac{\partial v}{\partial x} + \\ &\quad \frac{c}{2v}\exp\left(\frac{c}{2v}\left(x - \frac{c}{2}t\right)\right)\frac{\partial v}{\partial x} + \exp\left(\frac{c}{2v}\left(x - \frac{c}{2}t\right)\right)\frac{\partial^2 v}{\partial x^2}\end{aligned}$$

e substituindo na equação (3.1.1) chegamos em

$$\frac{\partial v}{\partial t} = v\frac{\partial^2 v}{\partial x^2}.$$

Já, trabalhando com as condições de contorno e inicial, obtemos:

$$\begin{aligned}u(x, 0) &= \exp\left(\frac{c}{2v}x\right)v(x, 0) \Rightarrow v(x, 0) = \exp\left(-\frac{c}{2v}x\right)f(x) \\ u(0, t) &= \exp\left(\frac{c}{2v}\left(-\frac{c}{2}t\right)\right)v(0, t) \Rightarrow v(0, t) = \exp\left(\frac{c^2}{4v}t\right)a(t) \\ u(l, t) &= \exp\left(\frac{c}{2v}\left(l - \frac{c}{2}t\right)\right)v(l, t) \Rightarrow v(l, t) = \exp\left(-\frac{c}{2v}\left(l - \frac{c}{2}t\right)\right)b(t).\end{aligned}$$

As equações (3.4.50) com (3.4.51) e (3.4.52) podem ser resolvidas pela família GTF(α)-FDS, como mostrado no capítulo anterior.

Capítulo 4

Resultados Numéricos

Para os experimentos numéricos, selecionamos alguns problemas que melhor traduzem o desempenho de cada método estudado. Restringimos os experimentos e comparações para tamanhos de passos no espaço e no tempo, h e k , fixos, tomando tamanho de passos pequenos apenas por razões de precisão da solução.

4.1 Métodos Avaliados

Avaliaremos os métodos obtidos nos capítulos anteriores, divididos pelas duas equações estudadas:

4.1.1 Equação de Difusão

- Fórmula Trapezoidal Estendida (ETF) com condições de contorno de Dirichlet:

$$\left(I + \frac{2}{3}rJ + \frac{1}{6}r^2J^2\right)\mathbf{u}_{j+1} = \left(I - \frac{1}{3}rJ\right)\mathbf{u}_j + \frac{r}{12}[5\mathbf{c}_j + 2(4I + rJ)\mathbf{c}_{j+1} - \mathbf{c}_{j+2}]$$

e com condições de contorno de Newman:

$$\left(I + \frac{2}{3}rA + \frac{1}{6}r^2A^2\right)\mathbf{u}_{j+1} = \left(I - \frac{1}{3}rA\right)\mathbf{u}_j + \frac{1}{3}rh(6I + rA)\mathbf{b}$$

- Fórmula Trapezoidal Generalizada (GTF(α)) com condições de contorno de Dirichlet:

$$(I + \frac{1}{2}(1 + \alpha)rJ + \frac{1}{2}\alpha r^2 J^2)\mathbf{u}_{j+1} = (I - \frac{1}{2}(1 - \alpha)rJ)\mathbf{u}_j + \frac{r}{2}\mathbf{c}_j + \frac{r}{2}[I + \alpha rJ]\mathbf{c}_{j+1}$$

e com condições de contorno de Newman:

$$(I + \frac{1}{2}(1 + \alpha)rA + \frac{1}{2}\alpha r^2 A^2)\mathbf{u}_{j+1} = (I - \frac{1}{2}(1 - \alpha)rA)\mathbf{u}_j + rh(2I + \alpha rA)\mathbf{b}$$

- Média Aritmética (GTF(0)) ou sistema equivalente ao de Crank-Nicolson (C-N) com condições de contorno de Dirichlet:

$$(I + \frac{1}{2}rJ)\mathbf{u}_{j+1} = (I - \frac{1}{2}rJ)\mathbf{u}_j + \frac{r}{2}(\mathbf{c}_j + \mathbf{c}_{j+1})$$

e com condições de contorno de Newman:

$$(I + \frac{1}{2}rA)\mathbf{u}_{j+1} = (I - \frac{1}{2}rA)\mathbf{u}_j + rh2\mathbf{b}$$

4.1.2 Equação de Convecção-Difusão

- Fórmula Trapezoidal Estendida (ETF(β_0)) com condições de contorno de Dirichlet:

$$(I + \frac{1}{12}(4 + \beta_0)C + \frac{1}{48}(2 + \beta_0)C^2)\mathbf{u}_{j+1} = (I + \frac{1}{12}(\beta_0 - 2)C - \frac{1}{48}\beta_0 C^2)\mathbf{u}_j + \frac{1}{48}[(10I + \beta_0 C)\mathbf{c}_j + \{16I + (2 + \beta_0)C\}\mathbf{c}_{j+1} - 2\mathbf{c}_{j+2}]$$

e com condições de contorno de Newman:

$$(I + \frac{1}{12}(4 + \beta_0)M + \frac{1}{48}(2 + \beta_0)M^2)\mathbf{u}_{j+1} = (I + \frac{1}{12}(\beta_0 - 2)M - \frac{1}{48}\beta_0 M^2)\mathbf{u}_j + \frac{1}{2}[I + \frac{1}{12}(1 + \beta_0)M]\gamma$$

- Regra de Simpson Estendida (ESR(α_0)) com condições de contorno de Dirichlet:

$$(I + \frac{1}{12}(5 - 4\alpha_0)C + \frac{1}{144}(11 - 16\alpha_0)C^2 + \frac{1}{144}(1 - 2\alpha_0)C^3)\mathbf{u}_{j+1} = (I - \frac{1}{12}(1 + 4\alpha_0)C + \frac{1}{144}(8\alpha_0 - 1)C^2)\mathbf{u}_j + \frac{1}{12}[I - \frac{1}{12}(1 + 4\alpha_0)C]\mathbf{c}_j + \frac{1}{3}[I + \frac{1}{6}(1 - 2\alpha_0)C]\mathbf{c}_{j+\frac{1}{2}} + \frac{1}{12}[I + \frac{1}{12}(5 - 4\alpha_0)C + \frac{1}{12}(1 - 2\alpha_0)C^2]\mathbf{c}_{j+1}$$

e com condições de contorno de Newman:

$$\begin{aligned} & \left(I + \frac{1}{12}(5 - 4\alpha_0)M + \frac{1}{144}(11 - 16\alpha_0)M^2 + \frac{1}{144}(1 - 2\alpha_0)M^3\right)\mathbf{u}_{j+1} = \\ & \left(I - \frac{1}{12}(1 + 4\alpha_0)M + \frac{1}{144}(8\alpha_0 - 1)M^2\right)\mathbf{u}_j + \\ & \frac{1}{2}\left[I + \frac{1}{6}(1 - 2\alpha_0)M + \frac{1}{72}(1 - 2\alpha_0)M^2\right]\gamma \end{aligned}$$

• Fórmula Trapezoidal Generalizada (GTF(α)) com condições de contorno de Dirichlet:

$$\left(I + \frac{1}{2}(1 + \alpha)A + \frac{1}{2}\alpha A^2\right)\mathbf{u}_{j+1} = \left(I - \frac{1}{2}(1 - \alpha)A\right)\mathbf{u}_j + \frac{1}{2}[\mathbf{c}_j + (I + \alpha A)\mathbf{c}_{j+1}]$$

• Média aritmética (GTF(0)) ou sistema equivalente ao de Crank-Nicolson (C-N), com condições de contorno de Dirichlet:

$$\left(I + \frac{1}{2}A\right)\mathbf{u}_{j+1} = \left(I - \frac{1}{2}A\right)\mathbf{u}_j + \frac{1}{2}[\mathbf{c}_j + \mathbf{c}_{j+1}]$$

4.2 Implementação dos Métodos

Os métodos foram implementados em Fortran 77 como linguagem de programação, rodando no sistema Linux. Os gráficos foram elaborados no Mathematica 4.0, rodando em Windows 95.

Os programas possuem uma estrutura de fácil compreensão, e escolhemos o método de decomposição LU para resolvermos os sistemas lineares, geralmente pentadiagonais, que surgem nas discretizações dos problemas. No apêndice, listamos o exemplo de um dos programas executados.

4.3 Resultados Numéricos

Dentre todos os problemas para os quais realizamos testes, selecionamos cautelosamente apenas aqueles que refletem de forma explícita, todas as qualidades e falhas dos métodos considerados em cada um deles.

Optamos por cinco problemas lineares, sendo que para a equação de difusão (3.1.1), os problemas são dois: um com condições de contorno de Dirichlet e outro com condições de contorno de Newman; e para a equação de convecção-difusão (4.1.1), os problemas são três: dois com condições de contorno de Dirichlet e um com condições de contorno de Newman.

Para cada problema a seguir, fornecemos a solução exata e o gráfico desta solução. Logo em seguida, mostramos as tabelas contendo os erros absolutos máximos, ou seja, calculados pela norma do máximo, obtidos por cada método nos pontos da malha considerada e os gráficos comparando os métodos aplicados em cada problema. Lembramos que o tamanho dos incrementos espacial e temporal, h e k , são sempre fixos. Procuramos para um certo valor de k , o ponto mais distante da solução exata.

De um modo geral, escolhemos o parâmetro livre que aparece nas equações que definem os métodos, considerando-se a estabilidade, ou seja, diferenciamos-os pela propriedade de A e L-estabilidade, e também o valor que nos desse a ordem mais elevada para cada método.

Problema 4.3.1

Seja

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}, \quad 0 < x < 2, \quad t > 0,$$

com condição inicial

$$u(x, 0) = 1, \quad 0 < x < 2,$$

e com condições de contorno de Dirichlet

$$u(0, t) = 0, \quad u(2, t) = 0, \quad t \geq 0.$$

A solução exata do problema é dada por

$$u(x, t) = \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{2n-1} \operatorname{sen}\left(\frac{2n-1}{2}\pi x\right) \exp\left(-\left(\frac{2n-1}{2}\right)^2 \pi^2 t\right).$$

Resolvemos este problema com os métodos ETF, GTF($\frac{1}{3}$), e o método de Crank-Nicolson (C-N), para valores de $h = 0.05$, e $k = 0.05, 0.1$ e 0.2 . Na tabela 4.3.1, fornecemos o erro absoluto para $x = 1$ e $t = 1$.

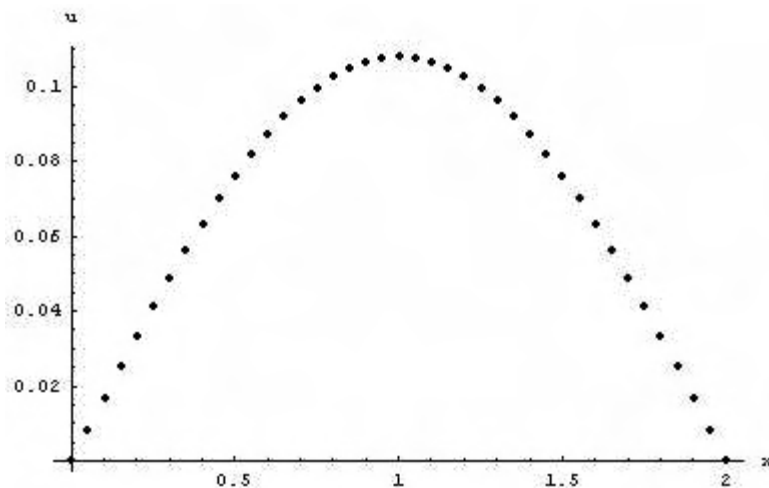
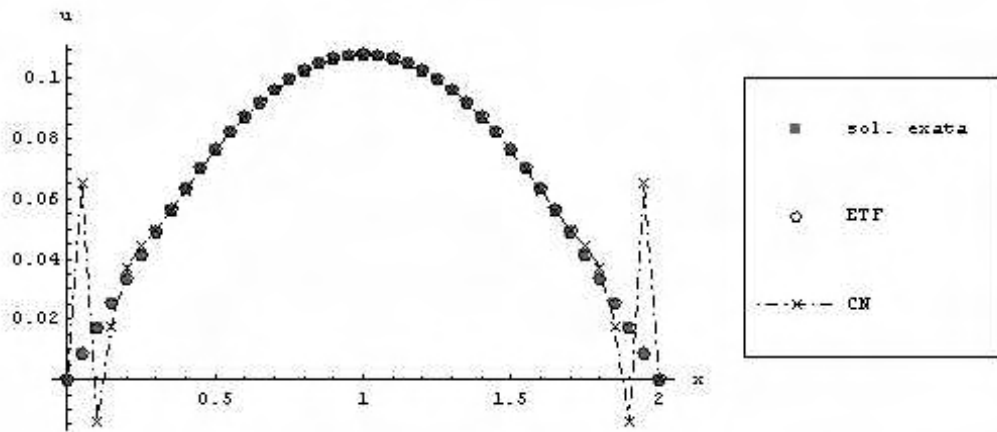


Figura 4.1: Solução Exata

Tabela 4.3.1

k	C-N	ETF(0)=GTF($\frac{1}{3}$)
0.05	2.5×10^{-4}	7.4×10^{-5}
0.1	1.2×10^{-3}	2.9×10^{-5}
0.2	1.5×10^{-2}	3.2×10^{-4}



Observamos que o método de Crank-Nicolson produz oscilações na solução computada, próximo ao contorno. Já os métodos ETF e GTF($\frac{1}{3}$), não produzem oscilações e dão melhores aproximações para a solução exata.

Problema 4.3.2

Seja

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t}, \quad 0 < x < 0.5, \quad t > 0,$$

com condição inicial

$$u(x, 0) = 0, \quad 0 \leq x \leq 0.5,$$

e com condições de contorno de Newman:

$$\frac{\partial u(0, t)}{\partial x} = 0, \text{ e } \frac{\partial u(0.5, t)}{\partial x} = 1.$$

A solução exata do problema é dada por

$$u(x, t) = 2t + \frac{1}{2} \left[\frac{12x^2 - 1}{6} - \frac{2}{\pi^2} \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos(2n\pi x) \exp(-4n^2\pi^2 t) \right].$$

Resolvemos este problema com os métodos ETF, GTF($\frac{1}{3}$), e o método de Crank-Nicolson (C-N), para $h = 0.0125$, e $k = 0.01$, e $t = 0.01, 0.02$ e 0.03 . Mostramos na tabela 4.3.2, o erro absoluto calculado pela norma do máximo.

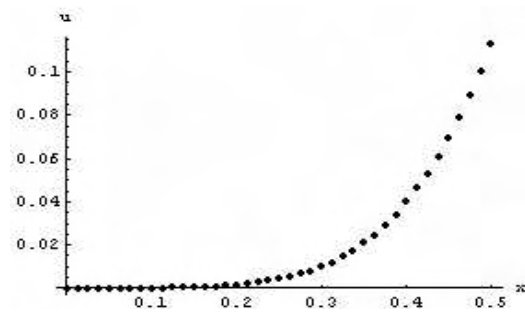
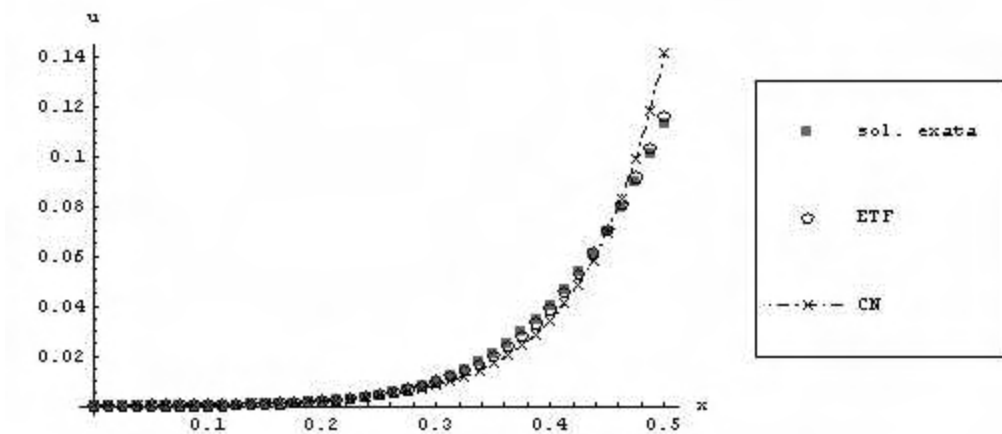


Figura 4.2: Solução Exata

Tabela 4.3.2

t	C-N	ETF(0)	GTF($\frac{1}{3}$)
0.01	2.8×10^{-2}	2.8×10^{-3}	2.8×10^{-3}
0.02	1.8×10^{-2}	5.0×10^{-5}	5.0×10^{-5}
0.03	1.6×10^{-2}	2.1×10^{-5}	2.1×10^{-5}



Observamos que o método de Crank-Nicolson distancia-se da solução exata, próximo ao contorno do lado direito do gráfico. Já os métodos ETF e $GTF(\frac{1}{3})$, são praticamente iguais neste exemplo e dão boas aproximações para a solução exata.

Problema 4.3.3

Seja

$$v \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}, \quad 0 < x < 1, \quad t > 0,$$

com condição inicial e com condições de contorno consistentes com a solução exata do problema

$$u(x, t) = \frac{1}{\sqrt{s}} \exp\left(-50 \frac{(x-t)^2}{s}\right), \quad s = 1 + 200vt.$$

Resolvemos este problema com os métodos $ETF(\beta_0)$, $GTF(\alpha)$, $ESR(\alpha_0)$ e o método de Crank-Nicolson (C-N), para valores de $h = 0.05$, e $k = 0.25$, $v = 1$ e $t = 1$. Mostramos na tabela 4.3.3, o erro absoluto calculado pela norma do máximo.

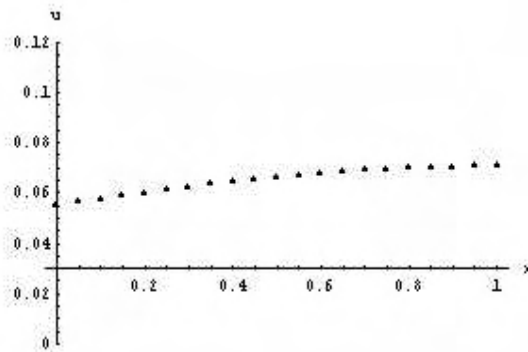
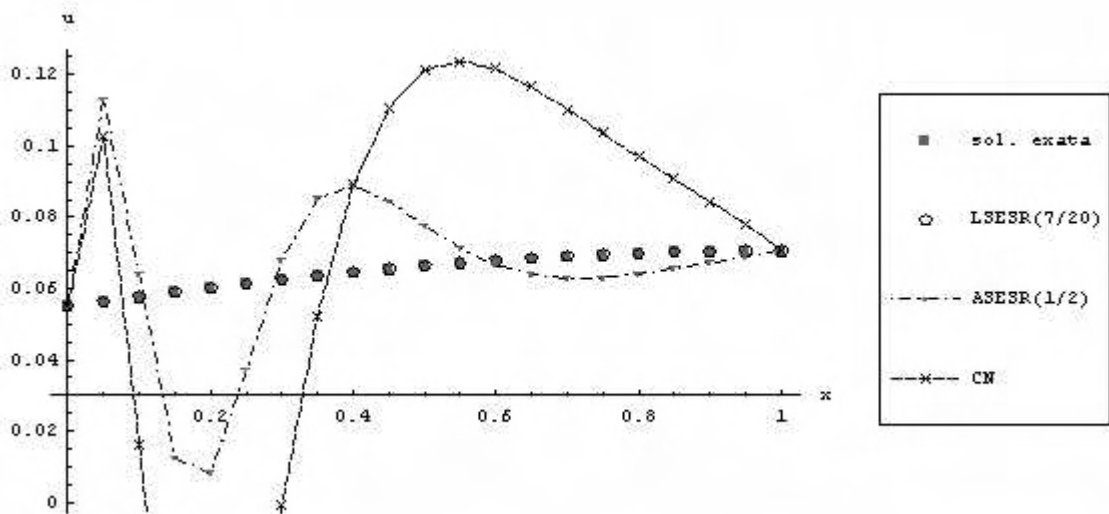


Figura 4.3: Solução Exata

Tabela 4.3.3

C-N	AS-ETF(-1)	LS-ETF(0)	AS-ESR($\frac{1}{2}$)	LS-ESR($\frac{7}{20}$)	GTF($\frac{1}{3}$)
1.6×10^{-1}	6.7×10^{-2}	7.2×10^{-5}	5.7×10^{-2}	1.8×10^{-5}	2.5×10^{-4}



Neste problema, tomamos condições de contorno e inicial consistentes com a solução exata na expectativa de que o método de Crank-Nicolson não produzisse oscilações. Porém, ele ainda produziu oscilações desordenadas na solução aproximada. Estas oscilações se devem ao fato deste método possuir sistema equivalente à Fórmula Trapezoidal Clássica. Esta é conhecida por produzir oscilações indesejáveis

em problemas rígidos, que é o caso deste exemplo. O AS-ESR($\frac{1}{2}$) que tem a Regra de Simpson A-estável associada, também produz oscilações, embora relativamente menores que o método de C-N. Assim, os métodos ETF(β_0), GTF(α) e ESR(α_0) são melhores que o método de Crank-Nicolson. Podemos perceber, observando os gráficos, que o método LS-ESR($\frac{7}{20}$) é o que melhor reproduz a solução exata, já que tem a Regra de Simpson L-estável associada, dando melhor precisão nas aproximações e sem oscilações. Este exemplo é uma indicação de que métodos L-estáveis de ordem mais elevada dão melhores aproximações.

Problema 4.3.4

Seja

$$0.1 \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}, \quad 0 < x < 1, \quad t > 0,$$

com condição inicial

$$u(x, 0) = 3\text{sen}(4\pi x), \quad 0 \leq x \leq 1,$$

e com condições de contorno de Dirichlet

$$u(0, t) = 0, \quad u(1, t) = 0.$$

A solução exata do problema é dada por

$$u(x, t) = \exp\left(\frac{c}{2v}\left(x - \frac{c}{2}t\right)\right) \sum_{n=1}^{\infty} B_n \exp(-vn^2\pi^2 t) \text{sen}(n\pi x),$$

$$B_n = \frac{3c}{2v} [1 + (-1)^{n+1} \exp(-\frac{c}{2v})] \left[\frac{1}{(\frac{c}{2v})^2 + (n-4)^2\pi^2} - \frac{1}{(\frac{c}{2v})^2 + (n+4)^2\pi^2} \right].$$

Resolvemos este problema com os métodos ETF(β_0), GTF(α), ESR(α_0) e o método de Crank-Nicolson (C-N), para valores de $h = 0.05$, e $k = 0.25$, $v = 0.1$ e $t = 1$. Mostramos na tabela 4.3.4, o erro absoluto calculado pela norma do máximo.

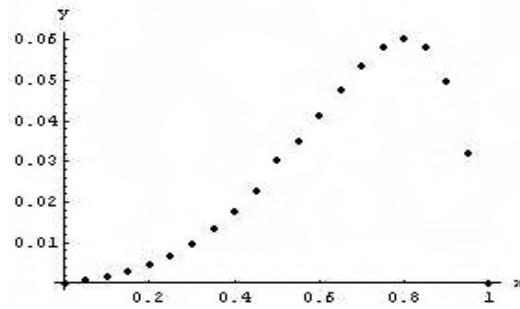
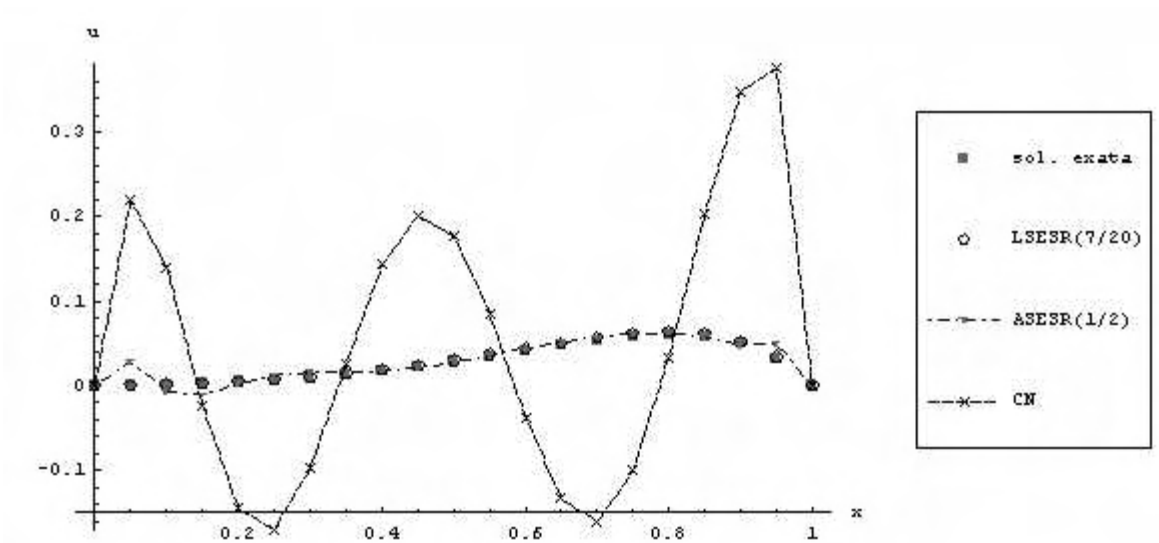


Figura 4.4: Solução Exata

Tabela 4.3.4

C-N	AS-ETF(-1)	LS-ETF(0)	AS-ESR($\frac{1}{2}$)	LS-ESR($\frac{7}{20}$)	GTF(0.45)
3.4×10^{-1}	2.8×10^{-2}	6.3×10^{-3}	2.8×10^{-2}	3.1×10^{-3}	4.8×10^{-3}



Aqui tomamos também condições de contorno e inicial consistentes com a solução exata na expectativa de que o método de Crank-Nicolson não produzisse oscilações. Porém, ele ainda produziu oscilações inaceitáveis. Assim, os métodos ETF(β_0), GTF(α) e ESR(α_0) são melhores que o Crank-Nicolson. Podemos perceber também, analisando os gráficos que os métodos ESR(α_0) são os que melhor reproduzem a solução exata, sendo que o LS-ESR($\frac{7}{20}$) é o que imita sutilmente a solução exata.

Problema 4.3.5

Seja

$$0.1 \frac{\partial^2 u}{\partial x^2} = \frac{\partial u}{\partial t} + \frac{\partial u}{\partial x}, \quad 0 < x < 1, \quad t > 0,$$

com condição inicial

$$u(x, 0) = \exp(5x) \left[\cos\left(\frac{\pi}{2}x\right) + 0.25 \operatorname{sen}\left(\frac{\pi}{2}x\right) \right], \quad 0 \leq x \leq 1,$$

e com condições de contorno de Newman

$$\frac{\partial u(0, t)}{\partial x} = \left(5 + \frac{\pi}{8}\right) u(0, t), \quad e \quad \frac{\partial u(1, t)}{\partial x} = -(2\pi - 5) u(1, t).$$

A solução exata do problema é dada por

$$u(x, t) = \exp\left(5\left(x - \frac{1}{2}t\right)\right) \exp\left(-\frac{\pi^2}{40}t\right) \left[\cos\left(\frac{\pi}{2}x\right) + 0.25 \operatorname{sen}\left(\frac{\pi}{2}x\right) \right].$$

Resolvemos este problema com os métodos ETF(β_0), ESR(α_0), e o método de Crank-Nicolson (C-N), com $h = 0.05$, $k = 0.25$, e $t = 2$. Mostramos na tabela 4.3.5, o erro absoluto calculado pela norma do máximo.

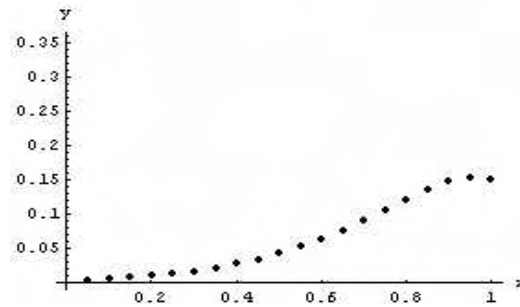
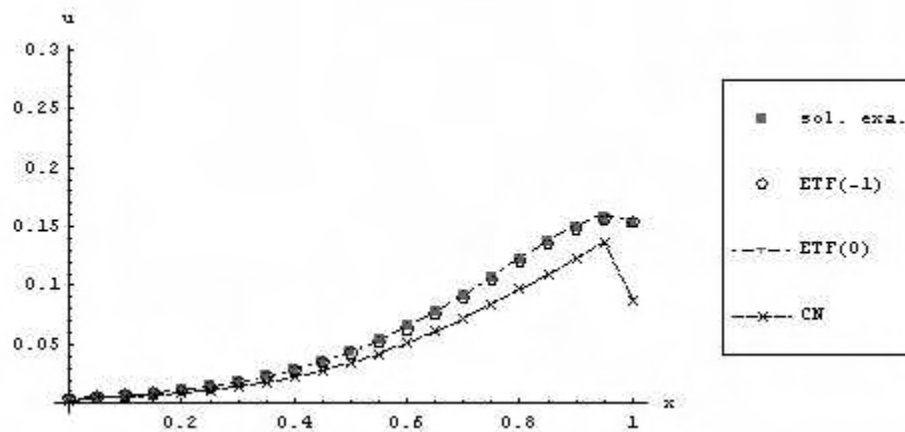


Figura 4.5: Solução Exata

Tabela 4.3.5

$C - N$	$AS - ETF(-1)$	$LS - ETF(0)$	$AS - ESR(\frac{1}{2})$	$LS - ESR(\frac{7}{20})$
6.5×10^{-2}	5.5×10^{-3}	2.4×10^{-3}	5.5×10^{-3}	4.1×10^{-3}



Tomamos novamente condições de contorno e inicial consistentes com a solução exata. Porém, o método de Crank-Nicolson ainda produziu uma solução indesejada do lado direito do gráfico, próximo ao contorno. Assim, os métodos $ETF(\beta_0)$ e $ESR(\alpha_0)$ são melhores que o método de Crank-Nicolson. Neste caso, os métodos que reproduzem melhor a solução exata são o $AS-ETF(-1)$ e o $LS-ETF(0)$.

4.4 Conclusões

Neste trabalho, estudamos a aplicação das Fórmulas Trapezoidais Estendida e Generalizada, e a Regra de Simpson Estendida para equações diferenciais parciais parabólicas em uma dimensão. Estas fórmulas foram propostas numa série de artigos, Chawla et al [2], [3], [4] e [5].

Nossa principal contribuição nesta dissertação, foi reunir as informações contidas em cada um dos artigos, fornecendo uma análise de cada fórmula quando aplicada na resolução de equações de difusão pura e equação de convecção-difusão, e uma análise comparativa entre elas.

Assim, verificamos que a Fórmula Trapezoidal Estendida ($\beta_0 = 0$) (1.4.5), aplicada na equação de difusão (2.1.1), levou à Fórmula Trapezoidal Estendida - Sistema de Diferenças Finitas (ETF-FDS) de ordem 3 na variável temporal e incondicionalmente estável. Esta mesma fórmula (1.4.5) foi aplicada à equação de convecção-difusão (3.1.1) e mostramos que esta ETF(β_0)-FDS será de ordem 3 na variável temporal, mas para $\beta_0 = -1$, terá ordem 4. Para $\beta_0 = 0$, será um sistema incondicionalmente estável com o método aplicado ETF L-estável e para $\beta_0 \geq -1$, será também incondicionalmente estável com o método aplicado ETF A-estável.

Verificamos também que a Fórmula Trapezoidal Generalizada (1.4.6) aplicada à equação (2.1.1), levou ao sistema GTF(α)-FDS que é, em geral, de ordem 2 em ambas variáveis temporal e espacial, e se $\alpha = \frac{1}{3}$, terá ordem 3 na variável temporal e é incondicionalmente estável. Agora, aplicando a mesma fórmula (1.4.6), para a equação (3.1.1), obtemos o método GTF(α)-FDS que é de ordem 2 em ambas variáveis temporal e espacial, e para $\alpha = \frac{1}{3}$, terá ordem 3 na variável temporal e é incondicionalmente estável.

Aplicamos ainda a Regra de Simpson Estendida (1.4.7) para a equação (3.1.1). Mostramos que para $\alpha_0 \leq \frac{1}{2}$, é um sistema incondicionalmente estável com um método aplicado ESR A-estável e para $\alpha_0 < \frac{1}{2}$, é um sistema incondicionalmente estável com um método aplicado ESR L-estável. Concluimos também que ESR(α_0)

é de ordem 4 na variável temporal, mas para $\alpha_0 = \frac{7}{20}$, tem ordem 5.

Agora, analisando os problemas considerados, os resultados obtidos nos permitem então afirmar que o método de Crank-Nicolson é inferior aos outros métodos estudados, na maioria dos casos, principalmente próximo ao contorno da solução exata. Além disso, todos os métodos L-estáveis são os que produzem as melhores aproximações para a solução exata.

Uma importante sequência desta dissertação seria um estudo destes métodos aplicados a problemas não-lineares e ainda para problemas de dimensão 2 e 3, e com aplicações no processamento de imagens, já que tem como modelo uma equação diferencial parcial parabólica.

Apêndice

```
C *****  
C Exemplo do Programa referente ao Problema 4.3.4  
C Método: Regra de Simpson Estendida - L-estável( $\frac{7}{20}$ )  
C LS-ESR( $\frac{7}{20}$ )  
C *****
```

Integer N

Real Nfim

Real*8 L, T, k, h, C1, S, A(100,100), R, V, j, i,

\$Ix, p, Ij, U(500), Idd(100,100), Mat1(100,100), Mat2(100,100),

\$PriM(100,100), SegM(100,100), Number, U2(500), B(100,100),

\$Mat5(100,100), Mat6(100,100), Mat7(100,100), Mat8(100,100),

\$Mat11(100,100), Mat10(100,100), Mat12(100,100), Mat4(100,100),

\$Mat15(100,100), Mat16(100,100), Mat17(100,100), Mat18(100,100),

\$Mat19(100,100), Mat21(100,100), Mat22(100,100), Mat20(100,100),

\$C(100,100), Mat9(100,100)

```

C**** H: tamanho do passo na variável espacial
Write(*,*) 'Digite o valor de H: '
Read(*,*) h
C**** K: tamanho do passo na variável temporal
Write(*,*) 'Digite o valor de k: '
Read(*,*) k
C**** V: constante da equação de convecção-difusão
Write(*,*) 'Digite o valor de V: '
Read(*,*) V
C**** L: tamanho do intervalo na variável espacial
Write(*,*) 'Digite o valor de L: '
Read(*,*) L
C**** T: tamanho do intervalo na variável temporal
Write(*,*) 'Digite o valor de T: '
Read(*,*) T
C**** C: constante da equação de convecção-difusão
Write(*,*) 'Digite o valor de C:'
Read(*,*) C1
Ix=L/h
Nfim=Ix-1
C**** N: dimensão do sistema
N=int(Nfim)
C**** R: constante do método
R=(V*k)/(h*h)
C**** S: constante do método
S=(C1*k)/h
C**** Ij: número de vezes que resolvo o sistema
Ij=(T/k)
Write(*,*) 'N=', N, 'R=', R, 'j=', Ij, 'S=', S

```

```
C**** Monta a matriz J do método
Call MontA(A,N)
C**** Monta o vetor inicial
Call MontU(U,N)
C**** Monta a matriz B do método
Call MontB(B,N)
C**** Monta a matriz identidade
Call MontI(Idd,N)
C**** Zera uma matriz
Call ZerMat(C,100)
Call ZerMat(Mat1,100)
Call ZerMat(Mat2,100)
Call ZerMat(Mat4,100)
Call ZerMat(Mat5,100)
Call ZerMat(Mat6,100)
Call ZerMat(Mat7,100)
Call ZerMat(Mat8,100)
Call ZerMat(Mat9,100)
Call ZerMat(Mat10,100)
Call ZerMat(Mat11,100)
Call ZerMat(Mat12,100)
Call ZerMat(Mat15,100)
Call ZerMat(Mat16,100)
Call ZerMat(Mat17,100)
Call ZerMat(Mat18,100)
Call ZerMat(Mat19,100)
Call ZerMat(Mat20,100)
Call ZerMat(Mat21,100)
Call ZerMat(Mat22,100)
```

```

Call ZerMat(PriM,100)
Call ZerMat(SegM,100)

C***** Cálculo do Primeiro Membro do sistema
Number=2*R
C**** Multiplica um número por uma matriz
Call MultMatNum(A,Number,N)
Call MultMatNum(B,S,N)
C**** Soma duas matrizes
C**** C matriz do método
Call SomaMat(A,B,C,N)
C**** Move uma matriz para outra
Call MoveMat(C,Mat1,N)
Number=8.246664888
Number=Number/60
Call MultMatNum(Mat1,Number,N)
Call SomaMat(Mat1,Idd,Mat21,N)
C*****
Call MoveMat(C,Mat22,N)
Number=9.753335112
Number=Number/60
Call MultMatNum(Mat22,Number,N)
C**** Multiplica duas matrizes
Call MultMat(C,C,Mat2,N,N)
Number=3.637834253
Number=Number/240
Call MultMatNum(Mat2,Number,N)
Call SomaMat(Mat22,Mat2,Mat20,N)
Call SomaMat(Mat20,Idd,PriM,N)

```

```

C*****Calculo do Segundo Membro do sistema
Number=-1
Number=Number/5
Call MoveMat(C,Mat4,N)
Call MultMatNum(Mat4,Number,N)
Call MultMat(C,C,Mat8,N,N)
Number=1
Number=Number/80
Call MultMatNum(Mat8,Number,N)
Call SomaMat(Idd,Mat4,Mat9,N)
Call SomaMat(Mat9,Mat8,SegM,N)
C**** Saída12.dat: arquivo com saída dos resultados
OPEN(1,FILE='Saida12.dat')
k=0.25
j=0.0
OPEN(1,FILE='Saida12.dat')
Do i=0, Ij-1
j=j+k
Write(1,*) 'U(',i,')',',-',',',',j,')'
C**** Multiplica matriz por vetor
Call MultVet(SegM,U,U2,N,N)
C**** Resolve o sistema no máximo pentadiagonal por decomposição L.U.
Call Penta(Mat21,U2,N)
Call Penta(PriM,U2,N)
Do p=1,N
Write(1.100)U2(p)
EndDo
C**** Repassa um vetor em outro
Call RepVet(U2,U,N)

```

```

100 Format(E40.20)
EndDo
C**** Fim do Algoritmo
Close(1)
stop
end
C**** Fim do Programa
C-----
C Subrotinas e funções auxiliares
C -----
C***** Move a matriz M1 para M2
SubRoutine MoveMat(M1,M2,N)
Real*8 M1(100,100), M2(100,100)
Integer N, i, j
Do i=1, N
Do j=1, N
M2(i,j)=M1(i,j)
EndDo
EndDo
Return
End

C***** Monta a matriz J do método
SubRoutine MontA(A,N)
Integer j, i, N
Real*8 A(100,100)
Do i=1,N
Do j=1, N
If (Abs(i-j).GT.1) Then

```

```

A(i,j)=0
else
If (i.EQ.j) Then
A(i,j)=2
else
A(i,j)=-1
EndIf
EndIf
EndDo
EndDo
Return
End

```

```

C***** Monta a matriz B do método
Subroutine MontB(B,N)
Integer j,i,N
Real*8 B(100,100)
Do i=1,N
Do j=1,N
If ((Abs(i-j).GT.1).or.(i.EQ.j)) Then
B(i,j)=0
else
If (i.EQ.(j-1)) Then
B(i,j)=1
else
IF (i.EQ.(j+1)) Then
B(i,j)=-1
Endif
Endif
Endif

```

```

Endif
EndDo
EndDo
Return
End

```

C***** Monta o vetor inicial

```
SubRoutine MontU(U,N)
```

```
Integer N, p
```

```
Real*8 U(500),x,i,j,Pi
```

```
Pi=3.14159265359
```

```
x=0.05
```

```
i=x*N
```

```
p=1
```

```
Do j=x,i,x
```

```
U(p)= 3*sin(4*Pi*j)
```

```
p=p+1
```

```
EndDo
```

```
Return
```

```
End
```

C***** Monta a matriz identidade

```
SubRoutine MontI(Idd,N)
```

```
Integer N, i, j
```

```
Real*8 Idd(100,100)
```

```
Do i=1, N
```

```
Do j=1, N
```

```
If (i.EQ.j) Then
```

```
Idd(i,j)=1
```

```
else  
Idd(i,j)=0  
EndIf  
EndDo  
EndDo  
Return  
End
```

```
C***** Zera uma matriz
```

```
SubRoutine ZerMat(T,N)
```

```
Integer N, i, j
```

```
Real*8 T(100,100)
```

```
Do i=1, N
```

```
Do j=1, N
```

```
T(i,j)=0
```

```
EndDo
```

```
EndDo
```

```
Return
```

```
End
```

```
C***** Repassa o vetor U2 em U
```

```
SubRoutine RepVet(U2,U,N)
```

```
Integer i
```

```
Real*8 U2(500), U(500)
```

```
Do i=1,N
```

```
U(i)=0
```

```
U(i)=U2(i)
```

```
EndDo
```

```
Return
```

```
End
```

C***** Multiplica a matriz M1 por M2 e guarda em P

SubRoutine MultMat(M1,M2,P,m,n)

Integer m, n, i, j, k, l

Real*8 M1(100,100), M2(100,100), P(100,100)

Do i=1,m

Do j=1,n

P(i,j)=0

endDo

enddo

Do i=1,n

Do j=1,m

Do k=1,n

P(i,j)=P(i,j)+(M1(i,k)*M2(k,j))

EndDo

EndDo

EndDo

Return

End

C***** Multiplica matriz M1 pelo vetor M2 e guarda em P

SubRoutine MultVet(M1,M2,P,m,n)

Integer m, n, i, j, k, l

Real*8 M1(100,100), M2(500), P(500)

Do i=1,m

P(i)=0

enddo

Do i=1,n

Do j=1,m

```

P(i)=P(i)+(M1(i,j)*M2(j))
EndDo
EndDo
Return
End

```

```

C***** Multiplica número por matriz
SubRoutine MultMatNum(M,Num,N)
Integer N, i, j
Real*8 M(100,100), Num
Do i=1,N
Do j=1,N
M(i,j)=M(i,j)*Num
endDo
endDo
Return
End

```

```

C***** Soma duas matrizes
SubRoutine SomaMat(M1,M2,M,N)
Integer N, i, j
Real*8 M1(100,100), M2(100,100), M(100,100)
Do i=1,N
Do j=1,N
M(i,j)=M1(i,j)+M2(i,j)
EndDo
EndDo
Return
End

```

```

C***** Soma dois vetores
Subroutine SomaVet(d8,d9,d10,N)
integer N,i
Real*8 d8(500),d9(500),d10(500)
Do i=1,N
d10(i)=0
d10(i)=d8(i)+d9(i)
EndDo
Return
End

```

```

C***** Calcula um sistema no máximo pentadiagonal por Decomposição L.U.
Subroutine penta(Mat,b,m)
Real*8 Mat(100,100),sup(m),sub(m),diag(m),b(m),supp(m),subb(m)
Do i=1,m
diag(i)=Mat(i,i)
EndDo
Do i=1,m-1
sup(i)=Mat(i,i+1)
EndDo
Do i=1,m-2
supp(i)=Mat(i,i+2)
EndDo
Do i=1,m-1
sub(i)=Mat(i+1,i)
EndDo
Do i=1,m-2
subb(i)=Mat(i+2,i)

```

```

EndDo
n=m
nn=n-1
do 2 i=2,nn
xmult=sub(i-1)/diag(i-1)
diag(i)=diag(i)-xmult*sup(i-1)
sup(i)=sup(i)-xmult*supp(i-1)
b(i)=b(i)-xmult*b(i-1)
xmult=subb(i-1)/diag(i-1)
sub(i)=sub(i)-xmult*sup(i-1)
diag(i+1)=diag(i+1)-xmult*supp(i-1)
b(i+1)=b(i+1)-xmult*b(i-1)
2 continue
xmult=sub(nn)/diag(nn)
diag(n)=diag(n)-xmult*sup(nn)
b(n)=(b(n)-xmult*b(nn))/diag(n)
b(nn)=(b(nn)-sup(nn)*b(n))/diag(nn)
do 3 i=n-2,1,-1
b(i)=(b(i)-supp(i)*b(i+2)-sup(i)*b(i+1))/diag(i)
3 continue
return
end

```

Referências Bibliográficas

- [1] Ames, W.F., *Numerical Methods for Partial Differential Equations*, Academic Press-Inc., Third Edition, 1992.
- [2] Chawla, M.M. & Al-Zanaidi, M.A., An Extended Trapezoidal Formula for the Diffusion Equation, *Intern. J. Computer Math.*, v38, p.51-59,1998
- [3] Chawla, M.M. & Al-Zanaidi, M.A. & Evans, D.J., Generalized Trapezoidal Formula for Parabolic Equations, *Intern. J. Computer Math.*, v70, p.429-443,1998
- [4] Chawla, M.M. & Al-Zanaidi, M.A. & Al-Aslab, M.G., Extended One-Step Time-Integration Schemes for Convection-Diffusion Equations, *Intern. J. Computer Math.*, v39, p.71-84,1999
- [5] Chawla, M.M. & Al-Zanaidi, M.A. & Evans, D.J., Generalized Trapezoidal Formulas for Convection-Diffusion Equations, *Intern. J. Computer Math.*, v72, p.141-154,1998
- [6] Chawla, M.M. & Al-Zanaidi, M.A. & Evans, D.J., A Class of Generalized Trapezoidal Formulas for the Numerical Integration of $y' = f(x, y)$, *Intern. J. Computer Math.*, v62, p.131-142,1996
- [7] Chawla, M.M., Karaballi, A.A.& Al-Sahar, M.S., Extended double-stride L-estable methods for the numerical solution of ODEs, *Computers Math. Applic.*, v31, p.1-6,1996

- [8] Crank, J. and Nicolson, P., A practical method for numerical evaluation of solutions of partial differential equations of the heat-conduction type, Proc. Camb. Phil Soc., v43, p.50-67,1947
- [9] Dallquist, G. & Bjorck, A., *Numerical Methods*, Prentice Hall, Englewood Cliffs, 1974
- [10] Faires, B., *Numerical Analysis*, PWS-Kent Publishing Company, Fifth Edition, 1993.
- [11] Jacques, I.B., Extended one-step methods for the numerical solution of ordinary differential equations, Intern. J. Computer Math., v29,p247-255,1989.
- [12] Noble, B. and Daniel, J.W., *Álgebra Linear Aplicada*, Prentice-Hall, 1986.
- [13] Smith, G.D., *Numerical Solution of Partial Differential Equations - Finite Difference Methods*, Clarendon Press-Oxford, Third Edition, 1985.
- [14] Usmani, R.A. and Agarwal, R.P., An A-stable extended trapezoidal rule for the numerical integration of ordinary differential equations, Computers Math. Applic., v11,p1183-1191,1985.
- [15] Wheatley, P.O. & Gerald, C.F., *Applied Numerical Analysis*, Addison-Wesley Publishing Company, Fourth Edition, 1992.