

# XII ERMAC

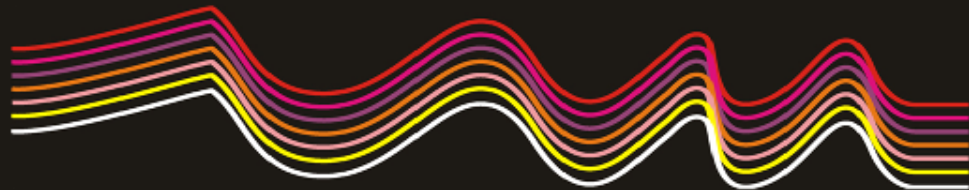
XII ENCONTRO REGIONAL DE MATEMÁTICA APLICADA E COMPUTACIONAL

---

11, 12 E 13 DE AGOSTO DE 2008 - UNIOESTE - FOZ DO IGUAÇU



# XII ERMAC



## Apresentação

APRESENTAÇÃO



Iniciar



# XII ERMAC



*Home*

Pôster

Comunicação

Cronograma



# XII ERMAC

## *Pôster*

- CÓDIGOS PERFEITOS EM ESPAÇOS DE HAMMING
- COMPARAÇÃO COMPUTACIONAL ENTRE OS MÉTODOS NUMÉRICOS DE EULER, HEUN E RUNGE-KUTTA PARA EQUAÇÕES DIFERENCIAIS ORDINÁRIAS DE 1ª ORDEM.
- CONDIÇÕES DE OTIMALIDADE E DE QUALIFICAÇÃO PARA PROBLEMAS DE PROGRAMAÇÃO NÃO LINEAR
- CONJUNTOS PARCIALMENTE ORDENADOS, LEMA DE ZORN E AXIOMA DA ESCOLHA
- ESTIMATIVA DE ERROS DE ITERAÇÃO EM MÉTODOS DE RELAXAÇÃO







## *Pôster*

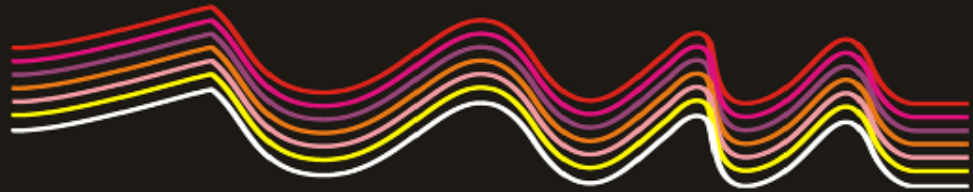
- MATEMÁTICA: RESULTADO DE UMA CONSTRUÇÃO SOCIAL
- MODELAGEM MATEMÁTICA DAS PROPRIEDADES DOS PRODUTOS DE COMBUSTÃO DE COMBUSTÍVEIS VEGETAIS, CONSIDERANDO O EQUILÍBRIO QUÍMICO
- MODELAGEM PROBABILÍSTICA NO DIMENSIONAMENTO DA PROGRAMAÇÃO DE UM SEMÁFORO
- PESQUISA E EXTENSÃO: CONSTRUINDO ALTERNATIVAS METODOLÓGICAS PARA AUXILIAR O PROCESSO DE ENSINO E APRENDIZAGEM DO CÁLCULO DIFERENCIAL E INTEGRAL NOS CURSOS DO CECE
- (RE) CONSTRUINDO O TANGRAM



## *Pôster*

- REPETÊNCIA E EVASÃO NO CÁLCULO: PROPONDO ALTERNATIVAS
- RESOLUÇÃO DO PROBLEMA DE STEINER
- TÉCNICAS DE ESTATÍSTICA MULTIVARIADA APLICADAS NA CLASSIFICAÇÃO DOS CANDIDATOS AO VESTIBULAR DA FECILCAM
- UMA APLICAÇÃO DA TEORIA FUZZY NA BIOMATEMÁTICA: DIAGNÓSTICO DE DOENÇAS INFANTIS
- UMA MODELAGEM DO TRANSPORTE E REAÇÕES DOS POLUENTES NO LAGO IGAPÓ I - LONDRINA, PARANÁ





## Comunicação

- AGRUPAMENTO BASEADO EM COLÔNIA DE FORMIGAS: ESTUDO COMPARATIVO DE ALGORITMOS PARA RECUPERAÇÃO DOS GRUPOS
- ANÁLISE DA INFLUÊNCIA DO TEOR DE UMIDADE INICIAL DO GRÃO, VELOCIDADE E TEMPERATURA DO AR NA SECAGEM DE SOJA
- ANÁLISE MORFOLÓGICA DOS ERITRÓCITOS NAS ANEMIAS ATRAVÉS DA APLICAÇÃO DE REDES NEURAIAS ARTIFICIAIS NO PROCESSAMENTO DE IMAGENS
- APLICAÇÃO DE TRANSFORMADAS DISCRETAS WAVELETS AO SOM DE UM CRUSTÁCEO
- AVALIAÇÃO DA QUALIDADE DE UM SERVIÇO DE TRANSMISSÃO DE VOZ SOBRE IP USANDO UM SISTEMA DE REGRAS NEBULOSAS



# XII ERMAC

## Comunicação

- CODES SATISFYING THE CHAIN CONDITION OVER ROSENBLOOM-TSFASMAN SPACES
- CÓDIGOS CORRETORES DE ERROS EM ESPAÇOS DE BLOCOS DE ROSENBLOOM-TSFASMAN
- CÓDIGOS LINEARES: CODIFICAÇÃO E DECODIFICAÇÃO DE CÓDIGOS COM UM ERRO
- CONSTRUÇÃO DOS NÚMEROS REAIS ATRAVÉS DOS CORTES DE DEDEKIND
- DESIGUALDADE ISOPERIMÉTRICA



- DETERMINAÇÃO COMPUTACIONAL DE ÁREAS E VOLUMES DE TUMORES MEDIANTE IMAGENS
- ESTABILIDADE DO SISTEMA DE ENERGIA ELÉTRICA COM FACTS
- ESTUDO DA CURVATURA DOS AEROFÓLIOS GERADOS PELA TRANSFORMAÇÃO GENERALIZADA DE JOUKOWSKI
- ESTUDO E CONSTRUÇÕES DE FRACTAIS
- EXISTENCE OF SOLUTIONS IN WEIGHTED SOBOLEV SPACES FOR SOME DEGENERATE QUASILINEAR ELLIPTIC EQUATIONS







## Comunicação

- EXISTÊNCIA DA SIGMA-ÁLGEBRA DE LEBESGUE
- FUNÇÕES ELÍPTICAS
- HOMOGENEIZAÇÃO DA EQUAÇÃO DA ONDA COM DOMÍNIOS E OPERADORES VARIANDO SIMULTANEAMENTE
- MODELAGEM E SIMULAÇÃO COMPUTACIONAL DE UM ATUADOR HIDRÁULICO CONSIDERANDO O EFEITO DO ATRITO DINÂMICO
- MODELAGEM MATEMÁTICA DOS PROCESSOS DE EVAPORAÇÃO E TRANSPIRAÇÃO NO SISTEMA SOLO - PLANTA



# XII ERMAC

## Comunicação

- MODELOS MATEMÁTICOS PARA DESCREVER O AUMENTO DE PESO EM GADO DE CORTE
- MULTICOLINEARIDADE EM ANÁLISE DE REGRESSÃO
- O CÓDIGO BINÁRIO DE HAMMING
- O MÉTODO SIMPLEX E O MODELO DE SISTEMA DE TRANSPORTES
- O PROBLEMA DA COBERTURA SOBRE ANÉIS E CONJUNTOS LIVRE DE PRODUTO



- ON THE NORM OF THE FOURIER-LAPLACE PROJECTION
- PARAMETRIZAÇÃO DE PARÁBOLAS EM COORDENADAS POLARES
- REGISTROS DE REPRESENTAÇÃO SEMIÓTICA E TECNOLOGIA NO ENSINO DE FUNÇÕES
- SIMULAÇÃO DA AERAÇÃO DE GRÃOS ARMazenADOS VARIANDO AS CONDIÇÕES DO AR AMBIENTE
- SIMULAÇÃO NUMÉRICA DE ESCOAMENTOS VISCOELÁSTICOS ATRAVÉS DE UMA CONTRAÇÃO BRUSCA



# XII ERMAC

## Comunicação

- SIMULAÇÃO NUMÉRICA DO PROCESSO DE VAZAMENTO E SOLIDIFICAÇÃO DA LIGA Al-22%Ni EM MOLDE METÁLICO
- SLOTTING DE ARMAZÉNS UTILIZANDO ALGORITMO GENÉTICO
- TÉCNICA GEOMÉTRICA DO PREDITOR SECANTE APLICADA AO FLUXO DE CARGA CONTINUADO NO PLANO MAGNITUDE DE TENSÃO NODAL E FATOR DE CARREGAMENTO
- UM ESTUDO NUMÉRICO DO PROBLEMA DE QUADRADOS MÍNIMOS NÃO LINEARES COM RESTRIÇÕES
- USO DOS ÂNGULOS E DAS TENSÕES NODAIS COMO PARÂMETRO NA SOLUÇÃO DE SISTEMAS DE EQUAÇÕES NÃO LINEARES DOS SISTEMAS DE POTÊNCIA



# XII ERMAC



---

## *Cronograma*

---

Geral

Comunicações





# XII ERMAC



---

## *Cronograma*

---



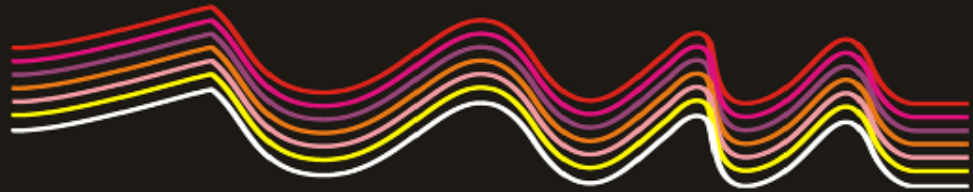
# XII ERMAC



## Cronograma



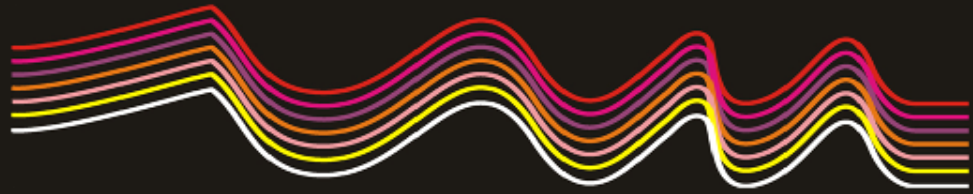
# XII ERMAC



## Cronograma



# XII ERMAC



## Cronograma



# XII ERMAC

---

## *Cronograma*

---

### Segunda-Feira -- Manhã

8:30 – 10:00	Entrega de material
10:00 – 10:40	Solenidade de Abertura
10:40 – 12:00	Palestra de Abertura: João Frederico da Costa Azevedo Meyer (UNICAMP)





# XII ERMAC

## *Cronograma*

### Segunda-Feira -- Tarde

- |               |   |
|---------------|---|
| 13:30 – 14:30 | Palestra 1: Mariza Stefanello Simsen (UNIFEI) – A Equação de Schrödinger        |
| 14:30 – 15:30 | Palestra 2: Irene Naomi Nakaoka (UEM) – O Princípio da Casa dos Ponbos          |
| 15:30 – 16:00 | Intervalo   |
| 16:00 – 17:00 | Palestra 3: Jacson Simsen (USP) – Teoria de Semi-Grupos e Equações Diferenciais |
| 17:00 – 18:00 | Palestra 4: Emerson Luiz do Monte Carmelo (UEM) – Euler e os 36 Oficiais        |



# XII ERMAC

## Cronograma

### Terça-Feira -- Manhã

- |               |   |
|---------------|---|
| 9:00 – 10:15  | Minicurso 1: Luiz Carlos Matioli (UFPR) – Formação de Carteiras de Investimentos pelo Modelo da Média Variância |
| 9:00 – 10:15  | Minicurso 2: Claiton Massarolo (UNIOESTE) – Introdução aos Problemas do Cálculo Variacional                     |
| 10:15 – 10:45 | Intervalo   |
| 10:45 – 12:00 | Minicurso 3: Emerson Luiz do Monte Carmelo (UEM) – Matemática na Loteria Esportiva                              |
| 10:45 – 12:00 | Minicurso 4: Tomas Edson Barros (UFSCAR) – Espaços Métricos Completos e o Método das Aproximações Sucessivas    |



# XII ERMAC

## *Cronograma*

### Terça-Feira -- Tarde

- |               |  |
|---------------|--|
| 13:30 – 14:30 | Palestra 5: Yuan Jin Yan (UFPR) – Metodologia Matemática: Ler os Livros e Artigos              |
| 14:30 – 15:30 | Palestra 6: Marcelo Escudeiro Hernandez (UEM) – Classificação Analítica de Curvas Irreduzíveis |
| 15:30 – 16:00 | Intervalo/Pôster   |
| 16:00 – 18:00 | Comunicações Orais   |



# XII ERMAC

## Cronograma

### Quarta-Feira -- Manhã

- |               |   |
|---------------|---|
| 9:00 – 10:15  | Minicurso 1: Luiz Carlos Matioli (UFPR) – Formação de Carteiras de Investimentos pelo Modelo da Média Variância |
| 9:00 – 10:15  | Minicurso 2: Claiton Massarolo (UNIOESTE) – Introdução aos Problemas do Cálculo Variacional                     |
| 10:15 – 10:45 | Intervalo   |
| 10:45 – 12:00 | Minicurso 3: Emerson Luiz do Monte Carmelo (UEM) – Matemática na Loteria Esportiva                              |
| 10:45 – 12:00 | Minicurso 4: Tomas Edson Barros (UFSCAR) – Espaços Métricos Completos e o Método das Aproximações Sucessivas    |



## *Cronograma*

### Quarta-Feira -- Tarde

13:30 – 15:30	Comunicações Orais
15:30 – 16:00	Intervalo/Pôster
16:00 – 17:00	Palestra 7: Ryuichi Fukuoka (UEM) – Introdução à Análise Técnica dos Mercados Financeiros
17:00 – 18:00	Palestra de Encerramento: José Plínio de Oliveira Santos (UNICAMP) – Alguns Problemas em Combinatória
18:00 – 18:10	Encerramento





# Uma modelagem do transporte e reações dos poluentes no Lago Igapó I - Londrina, Paraná

Suellen Ribeiro Pardo, Paulo Laerte Natti, Neyva Maria Lopes Romeiro, Eliandro Rodrigues Cirilo.

Universidade Estadual de Londrina - UEL, suellenrpardo@yahoo.com.br; Universidade Estadual de Londrina - UEL, plnatti@uel.br; Universidade Estadual de Londrina - UEL, nromeiro@uel.br; Universidade Estadual de Londrina - UEL, ercirilo@uel.br

## Resumo:

A crescente expansão demográfica e industrial observada nas últimas décadas trouxe como consequência o comprometimento das águas dos rios, lagos e reservatórios. Neste contexto, o presente projeto, através da modelagem matemática e simulação numérica, é uma contribuição para a compreensão dos processos químicos, físicos e biológicos que geram a dispersão de poluentes em meios aquáticos. Este estudo é realizado no Lago Igapó I, situado na cidade de Londrina, Paraná. Inicialmente modelou-se a geometria do Lago I, através de discretização estruturada em coordenadas generalizadas, sendo a fronteira discreta obtida a partir do método de interpolação polinomial spline cúbico parametrizado. Em seguida modelou-se o escoamento, admitindo-o como incompressível, utilizando equações de conservação de quantidade de movimento e da continuidade (modelo hidrodinâmico). Admitindo condições iniciais e de contorno, obteve-se os campos de velocidade e pressão em todo domínio do espelho d'água do lago. Enfim, descreveu-se numericamente o transporte laminar dos poluentes, por meio do campo de velocidades obtido, e suas reações, a partir da versão linearizada do modelo WASP5, ou seja, a evolução das concentrações locais da amônia ( $S_{nh}$ ), do nitrito + nitrato ( $S_{n03}$ ), da demanda bioquímica de oxigênio ( $X_s$ ) e do oxigênio dissolvido ( $S_0$ ).

**Palavras-chave:** modelo de transporte de poluentes, ciclo do nitrogênio, ciclo do carbono, lago Igapó I.

## 1. Introdução

A crescente urbanização nas cidades, sem um planejamento adequado, tem contribuído para a degradação dos corpos d'água. Ainda são poucas as pesquisas sobre a qualidade da água nas cidades, apesar de já existirem muitos levantamentos relativos à poluição de águas. Com o objetivo de frear esse tipo de degradação ambiental, muitas alternativas são discutidas atualmente e, neste contexto, este trabalho é uma contribuição para um melhor entendimento do efeito que uma descarga de poluentes pode causar no lago Igapó I. O lago localiza-se na cidade de Londrina, Paraná, na microbacia do Ribeirão Cambé, que tem suas nascentes na cidade de Cambé, aproximadamente 10 Km a oeste de Londrina. O lago é artificial e surgiu, em 1959, como solução para o problema de drenagem, dificultada por uma barragem natural de pedra. Foi esvaziado e limpo algumas vezes, sendo a última em 1996, quando suas margens foram revitalizadas. É um local de lazer da cidade, que além da represa, propicia a prática de esportes náuticos. Porém, o lago recebe despejos de poluentes e despejos clandestinos, não tratados, dos lagos Igapó IV, Igapó III, Igapó II e afluentes, que poluem o lago Igapó I. Com o objetivo de estudar o impacto que esses despejos podem causar no lago Igapó I, apresenta-se o texto na forma a seguir. Na seção 2, descreve-se a geometria do lago Igapó I através de uma malha computacional. Na seção 3, modela-se matematicamente o escoamento, admitindo-o como incompressível, gerando os campos de velocidade e de pressão em todo o seu domínio. Na seção 4, modela-se o transporte laminar advectivo-difusivo-reativo de poluentes utilizando o modelo bidimensional WASP5, em sua forma linearizada (ROMEIRO, 2003). Na seção 5, simula-se numericamente, através do método de diferenças finitas, o ciclo do carbono e do nitrogênio (AMBROSE *et al*, 1993).

## 2. Geometria computacional do lago Igapó I

A água do lago Igapó II escoava para o Igapó I, na passagem da Avenida Higienópolis,

como pode ser observado na figura (??), caracterizando a entrada. A saída trata-se de uma barragem física com passagem controlada de água por adutoras e rampas. A margem esquerda é uma região de vegetação rasteira e no alto existe um canal por onde o lago recebe água. A margem direita possui propriedades particulares e também tem um canal de recebimento de água em sua parte-média inferior.



Figura 1: Domínio físico do lago Igapó I

Ao longo do tempo, teorias matemáticas foram desenvolvidas com o objetivo de fornecer uma representação fiel da geometria física de um problema. Para discretizar o interior da malha será utilizado um sistema de EDP's do tipo elíptico (CIRILO; BORTOLI, 2006) dado por

$$\alpha x_{\xi\xi} + \gamma x_{\eta\eta} - 2\beta x_{\xi\eta} + \frac{1}{J^2} (Px_{\xi} + Qx_{\eta}) = 0 \quad (1)$$

$$\alpha y_{\xi\xi} + \gamma y_{\eta\eta} - 2\beta y_{\xi\eta} + \frac{1}{J^2} (Py_{\xi} + Qy_{\eta}) = 0 \quad , \quad (2)$$

onde  $x$  e  $y$  são as coordenadas cartesianas do domínio físico,  $\xi$  e  $\eta$  são as coordenadas generalizadas do domínio computacional,  $P$  e  $Q$  são os termos fonte,  $J$  é o jacobiano dado por

$$J = x_\xi y_\eta - x_\eta y_\xi \quad (3)$$

e

$$\alpha = x_\eta^2 + y_\eta^2, \quad \beta = x_\eta x_\xi + y_\eta y_\xi, \quad \gamma = x_\xi^2 + y_\xi^2, \quad (4)$$

com  $x_\xi$ ,  $x_\eta$ ,  $y_\xi$  e  $y_\eta$  denotando derivadas parciais. Enfim, as margens foram obtidas com interpolação polinomial spline cúbico parametrizado, cujas equações são

$$x(h) = \sum_{k=0}^3 a_{k_t} (h-t)^k \quad (5)$$

$$y(h) = \sum_{k=0}^3 b_{k_t} (h-t)^k, \quad (6)$$

onde a variável  $h$  é ajustada na medida que se deseja efetuar, ou não, a concentração de pontos à esquerda ou à direita de um ponto base parametrizado em  $t$ . Os coeficientes  $a_{k_t}$  e  $b_{k_t}$  são determinados via resolução de sistemas lineares tridiagonais (CIRILO; BORTOLI, 2006). Tais procedimentos foram adotados devido ao seu desempenho computacional e ainda pela rápida similaridade obtida com a geometria física, a partir de poucos pontos conhecidos do domínio.

Considerando 839 pontos espaçados ao longo das margens esquerda e direita e 35 pontos espaçados nos contornos de entrada e saída, a malha que descreve a geometria do lago Igapó I, está representada pela figura (??). A similaridade entre as figuras (??) e (??), referente ao domínio físico do lago Igapó I, fará com que os resultados gerados pelos modelos hidrodinâmico e de transporte estejam próximos da realidade.



Figura 2: Geometria computacional do lago Igapó I

### 3. Modelo hidrodinâmico

Para modelar o escoamento no Lago Igapó I, devido a sua característica tipo laminar (profundidade baixa e densidade constante), considerou-se um modelo bidimensional na forma adimensional, constituído pelas leis de conservação. Admitiu-se o escoamento do lago como incompressível e em regime permanente. Note que o escoamento é incompressível quando as variações de densidade são desprezíveis. Por outro lado, dizer que o escoamento ocorre em regime permanente, significa dizer que as propriedades do escoamento no lago, em um determinado ponto, não variam no tempo. Neste tipo de escoamento, as propriedades podem variar de ponto para ponto, mas devem permanecer constante em relação ao tempo num ponto. Nesta situação, a ação do vento e as trocas de calor do lago com o meio externo não são expressivas, por isso são irrelevantes. Enfim, não foram admitidas variações no contorno do lago, na massa do fluido e na área do espelho d'água ao longo do tempo. Assumindo tais hipóteses, as equações do modelo hidrodinâmico (ROSMAN, 1997) são

$$\frac{\partial u}{\partial t} + u \frac{\partial u}{\partial x} + v \frac{\partial u}{\partial y} = -\frac{\partial p}{\partial x} + \frac{1}{Re} \left( \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \right) \quad (7)$$

$$\frac{\partial v}{\partial t} + u \frac{\partial v}{\partial x} + v \frac{\partial v}{\partial y} = -\frac{\partial p}{\partial y} + \frac{1}{Re} \left( \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \right) \quad (8)$$

$$\nabla^2 p = -\frac{\partial^2 uu}{\partial x^2} - 2\frac{\partial^2 uv}{\partial xy} - \frac{\partial^2 vv}{\partial y^2} - \frac{\partial d}{\partial t} + \frac{1}{Re} \left( \frac{\partial^2 d}{\partial x^2} + \frac{\partial^2 d}{\partial y^2} \right) , \quad (9)$$

onde  $t$  é o tempo,  $u$  e  $v$  são as componentes do vetor velocidade nas direções  $x$  e  $y$ , respectivamente,  $Re$  é o número de Reynolds,  $d$  o divergente e  $p$  a pressão. Utilizando coordenadas generalizadas e aproximando as derivadas parciais por diferenças finitas, um sistema algébrico de equações é obtido. A equação algébrica proveniente da equação (??), foi resolvida iterativamente pelo método de relaxações sucessivas (SMITH, 1990; DE BORTOLI, 2000).

O campo de pressão obtido em todo domínio computacional foi inserido nas equações algébricas advindas das equações (??) e (??). Tais equações foram resolvidas iterativamente pelo método de Runge-Kutta de terceira ordem (KROLL; ROSSOW, 1989; DE BORTOLI, 2000) fornecendo as componentes  $u$  e  $v$  do vetor velocidade em toda a geometria.

#### 4. Modelo de transporte e reações dos ciclos do carbono e nitrogênio

O modelo das reações a ser incluído no modelo de transporte é a versão linearizada do modelo WASP5, limitado aos ciclos do carbono (C) e do nitrogênio (N), (ROMEIRO, 2001). Neste modelo o esquema de reações dos ciclos do carbono (C) e do nitrogênio (N) são descritos na figura (??), nas cores vermelha e azul, respectivamente.

O sistema acoplado de equações das reações cinéticas do modelo linearizado WASP5, que descrevem o comportamento das quatro espécies reativas presentes nos ciclos descritos na figura (??), ou seja, amônia ( $S_{nh}$ ), o nitrito + nitrato ( $S_{n03}$ ), a demanda bioquímica de oxigênio ( $X_s$ ) e o oxigênio dissolvido ( $S_0$ ), após um período de tempo, são dadas abaixo

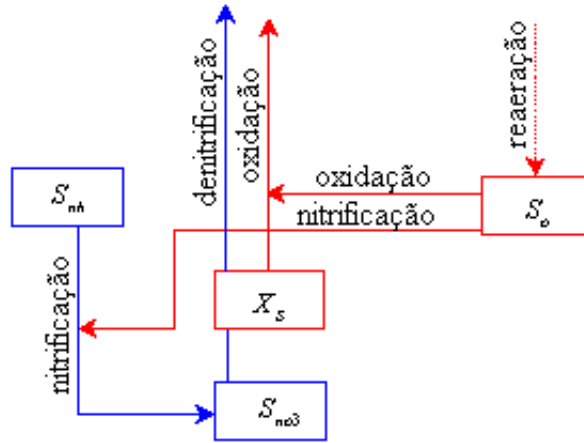


Figura 3: Esquemas de reações e ciclos utilizados no modelo WASP5

$$\frac{dS_{nh}}{dt} = -K_1 S_{nh} - K_7 S_0 + \tau_{S_{nh}}$$

$$\frac{dS_{n03}}{dt} = K_1 S_{nh} - K_2 S_{n03} + K_8 S_0 - \tau_{S_{n03}}$$

$$\frac{dX_S}{dt} = -\frac{20}{7} - K_2 S_{n03} - K_5 X_S - K_4 S_0 + \tau_{X_S}$$

$$\frac{dS_0}{dt} = -\frac{32}{7} - K_1 S_{nh} - K_5 X_S - K_6 S_0 + \tau_{S_0} ,$$

(10)

onde define-se as constantes

$$\begin{aligned}
K_1 &= k_{12}\Theta_{12}^{T-20} \left( \frac{\bar{S}_0}{k_{nit} + \bar{S}_0} \right) \\
K_2 &= k_{2D}\Theta_{2D}^{T-20} \left( \frac{K_{n03}}{k_{n03} + \bar{S}_0} \right) \\
K_3 &= k_D\Theta_D^{T-20} \left( \frac{\bar{X}_S k_{DBO}}{(k_{DBO} + \bar{S}_0)^2} \right) + \frac{32}{7}k_{12}\Theta_{12}^{T-20} \left( \frac{\bar{S}_{nh}k_{nit}}{(k_{nit} + \bar{S}_0)^2} \right) \\
K_4 &= k_D\Theta_D^{T-20} \left( \frac{\bar{X}_S k_{DBO}}{(k_{DBO} + \bar{S}_0)^2} \right) - \frac{32}{7}k_{2D}\Theta_{2D}^{T-20} \left( \frac{k_{n03}\bar{S}_{n03}}{(k_{n03} + \bar{S}_0)^2} \right) \\
K_5 &= k_D\Theta_D^{T-20} \left( \frac{\bar{S}_0}{k_{DBO} + \bar{S}_0} \right) \tag{11} \\
K_6 &= k_2\Theta_D^{T-20} + K_D\Theta_D^{T-20} \left( \frac{\bar{X}_S k_{DBO}}{(k_{DBO} + \bar{S}_0)^2} \right) + \frac{32}{7}k_{12}\Theta_{12}^{T-20} \left( \frac{\bar{S}_{nh}k_{nit}}{(k_{nit} + \bar{S}_0)^2} \right) \\
K_7 &= k_{12}\Theta_{12}^{T-20} \left( \frac{\bar{S}_{nh}k_{nit}}{(k_{nit} + \bar{S}_0)^2} \right) \\
K_8 &= k_{12}\Theta_{12}^{T-20} \left( \frac{\bar{S}_{nh}k_{nit}}{(k_{nit} + \bar{S}_0)^2} \right) + k_{2D}\Theta_{2D}^{T-20} \left( \frac{k_{n03}\bar{S}_{n03}}{(k_{n03} + \bar{S}_0)^2} \right) \\
\tau_{S_{nh}} &= k_7\bar{S}_0 & \tau_{S_{n03}} &= k_8\bar{S}_0 \\
\tau_{X_S} &= k_4\bar{S}_0 & \tau_{S_0} &= k_2\Theta_D^{T-20}C_{sat} + K_3\bar{S}_0 .
\end{aligned}$$

Os símbolos, valores e unidades dos parâmetros utilizados no modelo de reações são dados na tabela (1).



Tabela 1: Valores das constantes utilizadas no modelo WASP5

Símbolo	Valor	Unidade	Descrição do parâmetro
$T$	—	$^{\circ}C$	Temperatura
$\Theta_{2D}$	1,045		Coefficiente da temperatura para denitrificação
$\Theta_{12}$	1,08		Coefficiente da temperatura para nitrificação
$\Theta_D$	1,047		Coefficiente da temperatura para oxidação do carbono
$\Theta_2$	1,028		Coefficiente da temperatura para reaeração
$k_{2D}$	0,09	$h^{-1}$	Coefficiente da temperatura para denitrificação
$k_{12}$	0,22	$h^{-1}$	Taxa de nitrificação
$k_D$	0,38	$h^{-1}$	Taxa de oxidação do OD em $20^{\circ}C$
$k_2$	1,2517	$h^{-1}$	Taxa de reaeração em $20^{\circ}C$
$k_{DBO}$	0,001	$mgL^{-1}$	Constante de meia saturação da DBO carbonada
$k_{nit}$	0,2	$mgL^{-1}$	Constante de meia saturação para OD na nitrificação
$k_{n03}$	0,1	$mgL^{-1}$	Constante de meia saturação para OD na denitrificação
$S_{sat}$	8,3	$mgL^{-1}$	Concentração de saturação do OD

Assim, o modelo 2DH de transporte de poluentes, com reações linearizadas, é descrito pelo sistema

$$\begin{aligned}
 \frac{\partial S_{nh}}{\partial t} + U_i \frac{\partial S_{nh}}{\partial x_i} - \frac{1}{Re} D \frac{\partial^2 S_{nh}}{\partial^2 x_i} &= -K_1 S_{nh} - K_7 S_0 + \tau_{S_{nh}} \\
 \frac{\partial S_{n03}}{\partial t} + U_i \frac{\partial S_{n03}}{\partial x_i} - \frac{1}{Re} D \frac{\partial^2 S_{n03}}{\partial^2 x_i} &= K_1 S_{nh} - K_2 S_{n03} + K_8 S_0 - \tau_{S_{n03}} \\
 \frac{\partial X_S}{\partial t} + U_i \frac{\partial X_S}{\partial x_i} - \frac{1}{Re} D \frac{\partial^2 X_S}{\partial^2 x_i} &= -\frac{20}{7} - K_2 S_{n03} - K_5 X_S - K_4 S_0 + \tau_{X_S} \\
 \frac{\partial S_0}{\partial t} + U_i \frac{\partial S_0}{\partial x_i} - \frac{1}{Re} D \frac{\partial^2 S_0}{\partial^2 x_i} &= -\frac{32}{7} - K_1 S_{nh} - K_5 X_S - K_6 S_0 + \tau_{S_0} ,
 \end{aligned} \tag{12}$$

onde os índices  $i = 1, 2$ , representam as direções  $x$  e  $y$ , respectivamente,  $D$  a dispersão da espécie reativa no fluido,  $Re$  o número de Reynolds e  $U_i$  são as componentes do vetor velocidade fornecidas pelo modelo hidrodinâmico, equações (7-8).

## 5. Simulação numérica do modelo 2DH

A partir de uma simulação numérica para o transporte laminar de poluentes, descrito pelo modelo 2DH (modelo horizontal 2-dimensional), pode-se representar o impacto que uma descarga de amônia, na entrada do lago Igapó I, produz em toda sua extensão, caracterizada pela região compreendida entre a Avenida Higienópolis até a barragem (veja figura (??)). Para isso, admitiu-se em (12), o mesmo número de Reynolds  $Re$  e as componentes do campo de velocidades  $U_i$ , em todo domínio do lago Igapó I, fornecidas pelo modelo hidrodinâmico (7-8). Calculou-se os valores de  $K_1, K_2, K_4, K_5, K_6, K_7, K_8, \tau_{X_s}, \tau_{S_0}$  e  $\tau_{S_{n03}}$ , dados em (??), utilizando os valores da tabela (1). Obteve-se os valores dados na tabela (2).

Tabela 2: Valores das constantes utilizadas no modelo 2DH

$K_1 = 0,2320094118$	$K_2 = 0,1119642857x10^{-2}$	$K_4 = 0,2915820793x10^{-4}$
$K_5 = 0,3978120708$	$K_6 = 0,1,256960821$	$K_7 = 0,1144426298x10^{-2}$
$K_8 = 0,1144426298x10_{-2}$	$\tau_{S_{nh}} = 0,9498738273x10^{-2}$	$\tau_{S_{n03}} = -0,9498738273x10^{-2}$
$\tau_{X_s} = 0,2420131258x10^{-3}$	$\tau_{S_0} = 10,43277482$	

Também, estabeleceu-se-se como condições iniciais para as concentrações de amônia ( $S_{nh}$ ), de nitrito + nitrato ( $S_{n03}$ ), da demanda bioquímica de oxigênio ( $X_s$ ) e do oxigênio dissolvido ( $S_0$ ), na entrada do lago Igapó I, sob a passagem da Avenida Higienópolis (ver

figura (1)), os valores

$$\begin{aligned}
 S_{nh}(x_i, 0) &= 0,0 \text{ mgL}^{-1} \\
 S_{n03}(x_i, 0) &= 0,0 \text{ mgL}^{-1} \\
 X_S(x_i, 0) &= 0,0 \text{ mgL}^{-1} \\
 S_0(x_i, 0) &= 0,0 \text{ mgL}^{-1} \text{ ,}
 \end{aligned}
 \tag{13}$$

onde  $i = 1, 2$  e para condições nesta fronteira as concentrações abaixo

$$\begin{aligned}
 S_{nh}(0, t) &= 1,74 \text{ mgL}^{-1} \\
 S_{n03}(0, t) &= 0,0 \text{ mgL}^{-1} \\
 X_S(0, t) &= 5,05 \text{ mgL}^{-1} \\
 S_0(0, t) &= 8,3 \text{ mgL}^{-1} \text{ .}
 \end{aligned}
 \tag{14}$$

Enfim, tomando para o coeficiente de dispersão o valor de  $D = 10 \text{ m}^2\text{s}^{-1}$ , considerando um lançamento contínuo de  $1,74 \text{ mgL}^{-1}$  de amônia na entrada do lago Igapó I, consistente com a condição na fronteira (14), simula-se a variação das concentrações de amônia, nitrito+nitrato, demanda bioquímica de oxigênio e oxigênio dissolvido.

Tabela 3: Variação das concentrações das espécies

Horas	$S_{nh}$	$S_{n03}$	$X_S$	$S_0$
1 h	1,74 – 0,00	0,00 – 0,015	5,05 – 0,02	8,30 – 0,03
5 h	1,74 – 0,00	0,00 – 0,046	5,05 – 0,00	8,30 – 1,48
25 h	1,74 – 0,00	0,00 – 0,20	5,05 – 0,00	8,30 – 5,85
50 h	1,74 – 0,03	0,00 – 0,38	5,05 – 0,06	8,30 – 6,61
75 h	1,74 – 0,12	0,00 – 0,52	5,05 – 0,22	8,30 – 6,57
100 h	1,74 – 0,34	0,00 – 0,68	5,05 – 0,37	8,30 – 6,54
150 h	1,74 – 0,21	0,00 – 0,82	5,05 – 0,52	8,30 – 6,52

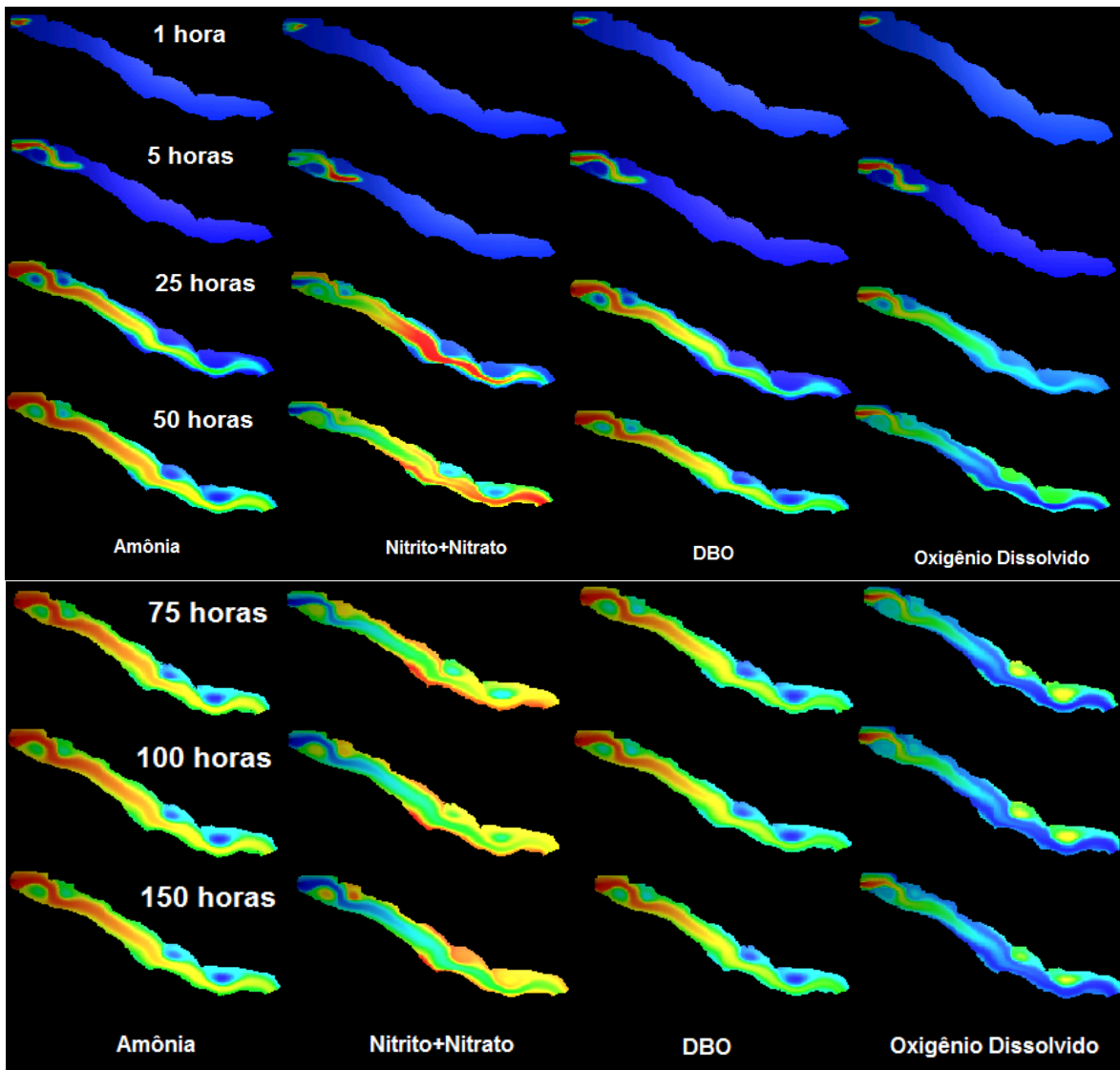


Figura 4: Evolução temporal e espacial considerando as quatro espécies reativas.

A figura (??) apresenta os resultados das simulações destas quatro espécies reativas, num intervalo de tempo de 150 horas de lançamento contínuo, enquanto a tabela (3) apresenta valores quantitativos das variações das concentrações das espécies reativas com seus valores máximos e mínimos no tempo indicado. Notamos que da tabela (3) verifica-se que, após 150 horas de lançamento contínuo, a concentração de oxigênio dissolvido atinge uma situação quase estacionária com uma taxa de  $6,52 \text{ mgL}^{-1}$ , o que de acordo com a Resolução do Conselho Nacional do Meio Ambiente (CONAMA, 1995), esta água pode ser destinada ao abastecimento para consumo humano, após tratamento, e à recreação de contato primário. Da figura (5), para  $t = 150$  vê-se que a concentração de amônia é alta na entrada (cor vermelha) e decai ao longo do lago, devido à decomposição através da oxidação biológica (processo em as bactérias degradam amônia em nitrato). Para uma concentração da ordem de  $0,2 \text{ mgL}^{-1}$  a  $0,3 \text{ mgL}^{-1}$  na região final do lago, como mostra a tabela (3). Por outro lado, a taxa de concentração de nitrito +nitrato aumentam nas regiões de vórtice, atingindo a concentração da ordem de  $0,8 \text{ mgL}^{-1}$ .

#### **Referências Bibliográficas:**

AMBROSE, R.B., WOOL, T.A., MARTIN, J.L.: **The Water Quality Analysis Simulation Program, WASP5**, Environmental Research Laboratory, US Environmental Protection Agency, Athens, G.A, (1993).

CIRILO, E.R., BORTOLI, A.L. DE.: **Geração do grid da traquéia e dos tubos bronquiais por splines cúbicos**, Semina: Ciências Exatas e Tecnológicas, vol. 27, 2, (2006), 147-155.

CONAMA 357/2005: **Resolução N° 357 de 17 de Março de 2005**. Disponível em : <http://www.mna.gov.br/ort/conama/res/res05/res35705.pdf>. Acesso em: junho/2008.

DE BORTOLI, A. L. **Introdução à dinâmica de fluidos computacional**. Porto Alegre: UFRGS, 2000.

ROMEIRO, N.M.L, CASTRO, R.G.S. E LANDAU, L.: **Simulação numérica do transporte de poluentes envolvendo quatro espécies reativas nos rios Negro e Solimões**, I Simpósio de Recurso Hídricos da Amazônia, Caderno de resumos e em CD-ROM, (2003), pp. 50.

ROMEIRO, N.M.L.: **Simulação numérica de modelos de qualidade de água usando o método de elementos finitos estabilizados**, Tese de doutorado, COPPE/UFRJ/RJ, (2003).

ROSMAN, P.C.C.: **Subsídios para Modelagem de Sistemas Estuarinos**. In: Rui Vieira. (Org.). Métodos Numéricos em Recursos Hídricos. : ABRH, 1997, v. 3, p. 231-343.

SMITH, G. D. **Numerical solution of partial differential equations: finite difference methods**. New York: Oxford University Press, 1990.

# ANÁLISE MORFOLÓGICA DOS ERITRÓCITOS NAS ANEMIAS ATRAVÉS DA APLICAÇÃO DE REDES NEURAIIS ARTIFICIAIS NO PROCESSAMENTO DE IMAGENS

Fernando Pinto, Luiz Carlos Radtke, Oleg Khatchatourian

Universidade Regional do Noroeste do Estado do Rio Grande do Sul – UNIJUÍ,  
fernexp@hotmail.com; Universidade Regional do Noroeste do Estado do Rio Grande do Sul –  
UNIJUÍ, lcradtke@hotmail.com; Universidade Regional do Noroeste do Estado do Rio Grande  
do Sul – UNIJUÍ, olegkha@unijui.edu.br

## **Resumo:**

Com o avanço tecnológico ocorrido nos últimos anos, foram criadas muitas tecnologias computacionais, usadas na medicina, sendo uma delas, utilizada na análise de imagens laboratoriais, na qual, observa-se uma padronização em suas características. A grande quantidade de imagens adquiridas nos procedimentos técnicos realizados em exames, exige uma forma de automatizá-las, de modo a agilizar o processamento. Este estudo propôs a utilização de uma rede neural artificial, a fim de auxiliar o processamento de imagens na identificação de padrões morfológicos presentes nas anemias. Foram analisados quatro tipos de eritrócitos: normais, microcíticos, megaloblásticos e falciformes. Com a intenção de automatizar a identificação de alguns tipos de anemias e validar a utilização de redes neurais artificiais no reconhecimento de padrões, foi desenvolvida uma ferramenta utilizando a plataforma Matlab. A rede configurada foi composta por 2500 neurônios de entrada na primeira camada, 100 neurônios na camada oculta e 4 neurônios na camada de saída, utilizando a função sigmóide como função de transferência e o algoritmo *backproagation* para realizar o treinamento. Os resultados obtidos indicam que a rede teve um bom desempenho, identificando cerca de 95% das células apresentadas, mostrando-se uma ferramenta eficaz no reconhecimento de padrões.

**Palavras-chaves:** redes neurais artificiais, processamento de imagens, análise morfológica dos eritrócitos.

## 1. Introdução

Devido à grande quantidade de imagens adquiridas nos procedimentos técnicos realizados em exames, faz-se necessário que haja uma forma de automatizá-las, a fim de agilizar a identificação.

Com o avanço tecnológico ocorrido nos últimos anos, foram criadas muitas tecnologias ligadas à computação, que, atualmente estão sendo usadas na solução de diversos problemas, em diversas áreas da medicina, sendo uma delas, utilizada na análise de imagens laboratoriais, na qual, observa-se uma padronização das características da mesma.

De acordo com o exposto, tomou-se como estudo de caso, as alterações morfológicas dos eritrócitos (ou hemácias) em determinadas doenças hematológicas, como nas anemias. As anemias são caracterizadas pela redução do número de eritrócitos (hemácias, glóbulos vermelhos) no sangue ou pela redução da taxa de hemoglobina (proteínas que transportam o oxigênio).

Há diferentes tipos de anemias, das quais, três foram abordadas para a análise: anemia ferropriva, megaloblástica e falciforme.

A anemia ferropriva é ocasionada por níveis baixos de ferro no organismo, onde ao tentar compensar à falta de ferro a medula óssea acelera a divisão celular, isso faz com sejam produzidos eritrócitos imaturos e pequenos (microcíticas), sua incidência é tão elevada que constitui um problema mundial de saúde pública.

A anemia megaloblástica é caracterizada pelo aumento no tamanho das células (macrocíticas), normalmente causadas pela insuficiência de vitamina B12 ou de ácido fólico.

A anemia falciforme é uma anemia hemolítica grave, conseqüência de um distúrbio genético onde sua característica fundamental é uma alteração na hemoglobina dos eritrócitos, que modifica seu formato normal, assumindo o formato de uma foice.

A característica de um eritrócito normal (sadio) é redonda e uniforme no tamanho, qualquer mudança significativa na morfologia, representa alguma anormalidade, onde cada grupo de anemias causa uma anormalidade diferente. Estas alterações são observadas microscopicamente, através de lâminas de microscopia preparadas com esfregaço do sangue, e tem por finalidade o estudo visual da morfologia celular (VALLADA, 1997). Tal procedimento busca concluir o exame laboratorial chamado de hemograma.



No hemograma completo existe uma série de elementos figurados que devem ser avaliados quantitativamente, para os quais, já existem meios eletrônicos para esta realização. Ao mesmo tempo, apesar do progresso da tecnologia eletrônica, a observação do esfregaço de sangue com o microscópio ainda é indispensável, para a avaliação qualitativa das mesmas. O estudo visual permite identificar as alterações da morfologia celular que não foram descobertas pelos caminhos eletrônicos, e avaliar a coerência dos resultados do hemograma.

Se houver índices hematimétricos anormais ou sinais de anemia, o técnico faz a confirmação visualmente, observando as alterações numéricas e pesquisando dados morfológicos (FAILACE, 1995).

Com o propósito de automatizar esse processo, utilizou-se uma ferramenta criada em cima da plataforma Matlab, a qual, apropriou-se de interações algébricas, a fim de desenvolver o processamento das imagens.

Segundo Gonzáles, 2000, o processamento de imagens pode ser dividido em cinco etapas: aquisição da imagem, pré-processamento, segmentação, representação/descrição e reconhecimento/interpretação

Um grande problema encontrado no processamento de imagens, se dá na hora de identificar (reconhecer) os padrões desejados, porém, nos últimos anos, têm se desenvolvido estudos utilizando técnicas de inteligência artificial, apresentando resultados positivos na solução de problemas complexos.

Redes neurais artificiais, são estruturas computacionais que tentam simular os cálculos sinápticos realizados pelo sistema nervoso biológico, tentando assim, reproduzir a inteligência humana artificialmente (TAFNER, 1996).

Baseando-se neste pressuposto, este estudo propôs a utilização de uma rede neural artificial, a fim de auxiliar no processamento de imagens na identificação dos padrões morfológicos presentes nas anemias.

As figuras abaixo representam as quatro morfologias consideradas. A imagem (a) representa eritrócitos normais; a imagem (b) representa eritrócitos com anemia ferropriva; imagem (c) a anemia megaloblástica; e a imagem (d) representa a anemia falciforme.

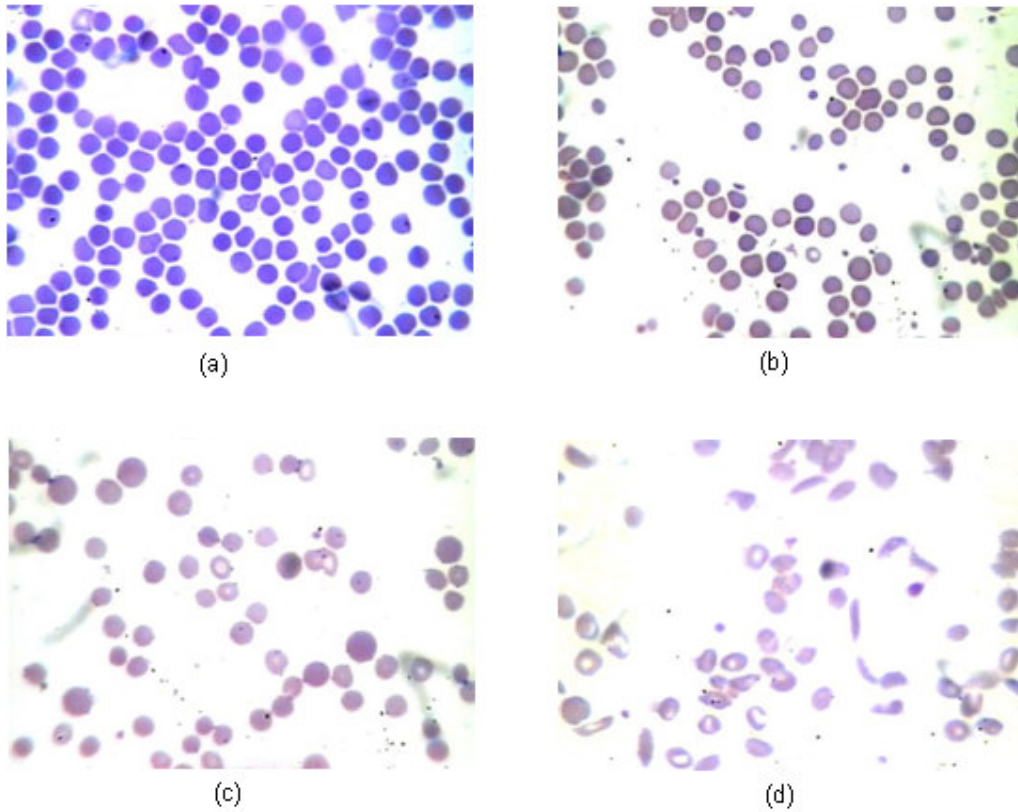


Figura 1: Morfologia dos eritrócitos: (a) Eritrócitos Normais; (b) Eritrócitos microcíticos; (c) Eritrócitos megaloblásticos ; (d) Eritrócitos falciformes

## 2. O processamento das imagens

O processamento das imagens foi dividido em cinco etapas: aquisição da imagem, pré-processamento, segmentação, representação, e reconhecimento.

A aquisição das imagens das lâminas com esfregaço de sangue foi realizada através de uma câmera de CCD acoplada a um microscópio, com resolução de 640 x 480 pixels, e salvas em formato Jpeg.

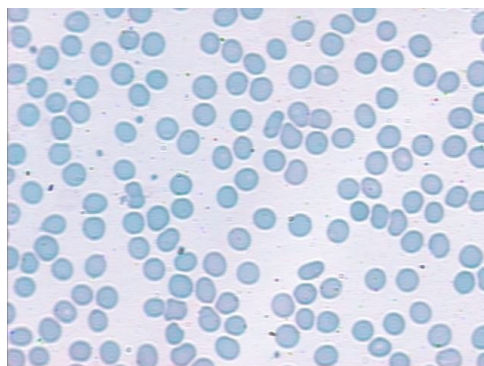


Figura 2: Imagem original

As demais etapas do processamento, após a aquisição da imagem, foram desenvolvidas através de uma ferramenta criada em cima do software matlab.

Na etapa de pré-processamento, após a imagem ser inserida no programa é feito a suavização da imagem para tons de cinza através da função *rgb2gray*, como mostra a figura 3. A imagem é transformada em uma matriz, em que, cada posição da matriz representa um pixel e seu valor representa o tom cinza do referido pixel.

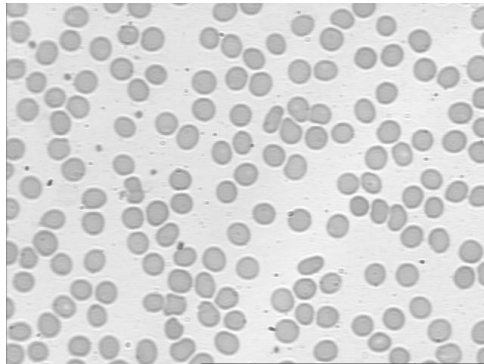


Figura 3: Imagem em tons de cinza

Na etapa seguinte, como pode ser observado na figura 4, é realizada a segmentação, utilizou-se o método de Canny para efetuar a detecção dos contornos das células.

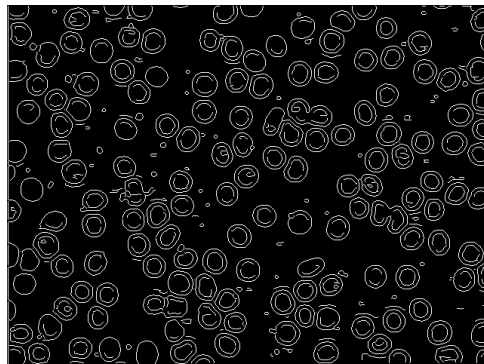


Figura 4: Contornos detectados através do método de Canny

Contudo, o resultado não foi satisfatório, apresentando contornos fracos e descontínuos, devido a isso, foi necessário realizar a dilatação dos contornos, o qual, foi realizado através da função *imdilate* do matlab, como pode ser visto na figura 5 .

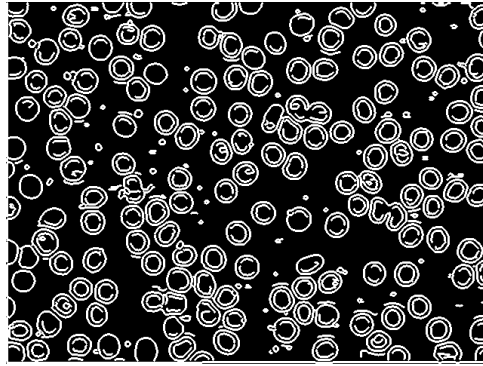


Figura 5: Contornos dilatados

Porém, ainda restaram contornos descontínuos, sendo necessário aplicar a função *imfill*, a qual, preenche todas as células com contornos contínuos, figura 6.

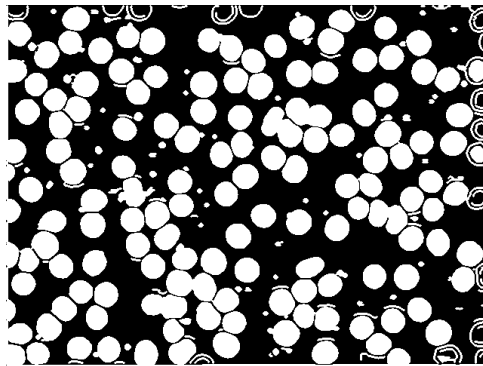


Figura 6: Contornos preenchidas

As células localizadas nas bordas da imagem por serem difíceis de identificar foram removidas através da função *imclearborder*, figura 7.

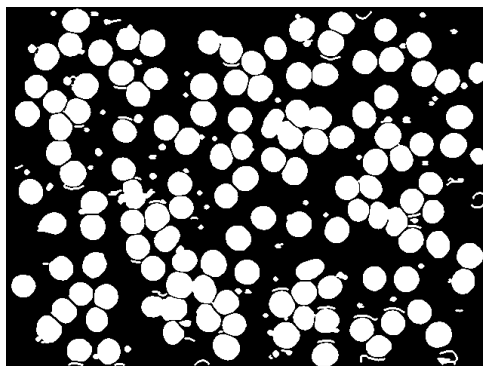


Figura 7: Remoção das células localizadas sobre a borda

Devido à dilatação dos contornos, realizada anteriormente, algumas células podem perder as suas características impedindo assim a identificação das mesmas, com a intenção de restaurar essa dilatação, foi aplicado a função *imerode*, reduzindo a dilatação realizada anteriormente, de acordo com a figura 8.

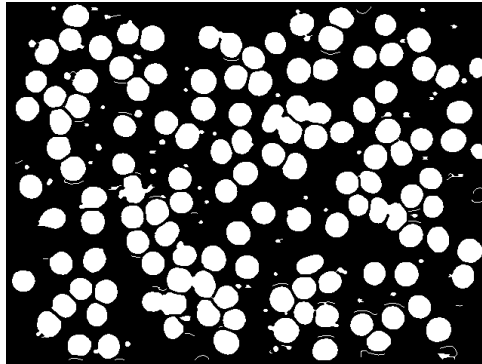


Figura 8: Restauração da dilatação

Para concluir a etapa de segmentação, foram removidos através da função *bwareaopen* aglomerados com mais de 700 pixels, por serem considerados agrupamentos com mais de uma célula e ruídos com menos de 300 pixels, por estes não representarem células, mas sim sujeiras, fungos, ou até mesmo interferências na câmera, figura 9.

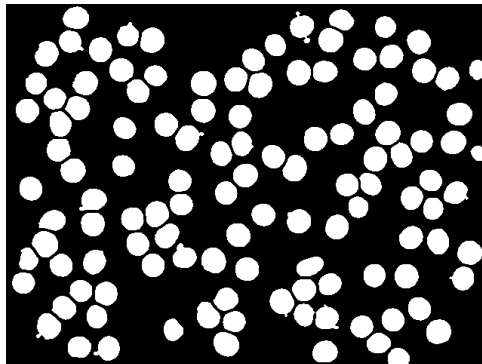


Figura 9: Remoção de aglomerados menores que 300 e maiores que 700 pixels

Com o encerramento da etapa de segmentação iniciou-se a etapa de representação, onde através da função *bwboundaries*, foram obtidas características morfológicas de cada célula extraídas da imagem resultante da etapa de segmentação, retornando as coordenadas de cada pixel que formam uma célula, e numerando-as seqüencialmente como mostra a figura 10.



Porém, antes de utilizar uma rede neural para reconhecer algum padrão é necessário determinar a configuração da arquitetura e realizar o treinamento da mesma.

### 3. Configuração da arquitetura da Rede Neural

A arquitetura da rede neural foi dada através de consulta bibliográfica, segundo Haykin(2001), a arquitetura de rede neural mais indicada para este trabalho é a rede *feedforward* (propagação) multicamadas, com uma camada oculta, e seu aprendizado dado através do algoritmo de aprendizado *backpropagation* (retropropagação).

A criação da rede no matlab foi desenvolvida através da função *newff* , a qual, foi configurada com 2500 neurônios de entrada na primeira camada, 100 neurônios na camada oculta, 4 neurônios na camada de saída, utilizando-se de uma função sigmóide como função de transferência entre as camadas .

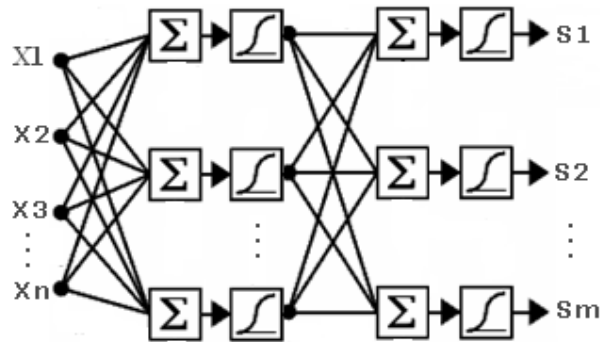


Figura 12: Arquitetura da rede utilizada

### 4. Treinamento de Rede

Depois da configuração da arquitetura da rede executou-se o treinamento da mesma com intuito de habilitá-la no reconhecimento morfológico das células. Para isso, foi definido um conjunto de treinamento e um conjunto de saídas desejadas. O conjunto de células selecionadas para realizar o treinamento foi submetido a etapa de pré-processamento que resultou em uma matriz de dimensão 2500 por 261, a qual, cada coluna representa uma célula reconhecida no conjunto de treinamento e será utilizada como vetor de entrada da rede.

Após, definir as dimensões do conjunto de treinamento, foi determinado o conjunto de saídas desejadas, como sendo uma matriz 4 por 261, onde cada coluna

representa a saída desejada para cada célula e as linhas representam os padrões a serem treinados, ou seja, cada coluna foi preenchida com zeros, exceto, na posição que representa esse padrão, onde foi inserido o valor 1, formando uma matriz de rótulos para as células do conjunto de treinamento, figura 13.

	Normal	Ferropirva	Megaloblástica	Falciforme
1	0	0	0	0
0	1	0	0	0
0	0	0	1	0
0	0	0	0	1

Figura 13: Vetores objetivos

## 5. Algoritmo de Treinamento

O procedimento para ajustar os pesos em uma rede neural (algoritmo de treinamento), os quais, inicialmente possuem valores randômicos consiste de duas etapas: a feedforward (propagação) e a backpropagation (retropropagação).

O treinamento de uma rede, com aprendizado backpropagation, em geral requer muitas épocas de treinamento (RICH,1993).

O algoritmo backpropagation, foi desenvolvido com base na regra de Widrow e Hoff (regra delta). O ajuste dos pesos ( $W_s$ ) entre a camada oculta e a camada de saída é dado por:

$$W_{sn} = W_s + \alpha \cdot S_o \cdot \Delta_s$$

$$\Delta_s = Er \cdot g'(Ft)$$

$$Er = S_d - S_e$$

Onde:

$W_s$  são os pesos da camada de saída

$W_{sn}$  são os novos pesos da camada de saída

$\alpha$  é a constante de aprendizado

$S_d$  é a saída desejada

$S_e$  é a saída da camada de saída

$S_o$  é a saída da camada oculta



$g'(Ft)$  é a derivada da função de transferência

$Er$  é o erro entre a saída desejada e a saída esperada

O ajuste dos pesos ( $W_o$ ) entre a camada de entrada e a camada oculta, foi definido por uma quantidade análoga de erro em relação ao erro da camada de saída, pois os neurônios da camada oculta são responsáveis por uma parte do erro da camada de saída.

$$W_{on} = W_o + \alpha \cdot X \cdot \Delta_o$$

$$\Delta_o = g'(Ft) \cdot \sum W_s \cdot \Delta_s$$

Onde:

$W_o$  são os pesos da camada oculta

$W_{on}$  são os novos pesos da camada oculta

$X$  são os valores de entradas da rede

A rede treinada para reconhecer os quatro tipos de morfologias alcançou o erro estipulado de 1% com 376 épocas. Uma época é configurada cada vez que todos os vetores de entrada são apresentados ao treinamento da rede.

O erro estipulado é o percentual de células utilizadas no treinamento não reconhecidas pela rede.

## 6. Testes e resultados

Com o término do treinamento, a rede foi submetida a testes para verificar a sua eficácia. Os quais estão descritos na tabela 1.

Tabela de resultados			
Tipo de eritrócito	Número de células analisadas	Células analisadas erradamente	Percentual de erro
Normal	118	3	2,54%
Ferropriva	107	4	3,73%
Megaloblástica	73	4	5,47%
Falciforme	51	6	11,76%
Total	349	17	4,87%

Tabela 1: Resultado das imagens analisadas

Como pode-se observar a morfologia com reconhecimento mais eficaz foi a dos eritrócitos normais com apenas 2,54% de erro. Doutro norte a anemia falciforme obteve o menor número de reconhecimentos com 11,76% de erro.

Fazendo a análise do conjunto total de reconhecimento para os quatro padrões, concluí-se que a rede obteve um bom desempenho, identificando cerca de 95,13% das células apresentadas, mostrando-se uma ferramenta eficaz no reconhecimento de padrões.

### **Referências bibliográficas**

ANDRADE, M. C.: **Um Algoritmo Interativo para Suavização e Segmentação de Imagens Digitais**. Centro de Desenvolvimento da Tecnologia Nuclear – CDTN. UFMG, Belo Horizonte, MG.

BULLOCK, D.: **Neural Networks for you toolbox**. Agricultural Engineering, July, 1962.

FAILACE, R.: **Hemograma: Manual de Interpretação** –3. ed. Ver. Aum. - Editora Artes Médicas Sul Ltda, Porto Alegre – RS – Brasil, 1995.

FAUSETT, L. V.: **Fundamentals of Neural Networks: Architectures, Algorithms, and Applications**. Prentice-Hall, Inc. New Jersey, 1994.

GONZALES, R. C., Woods, R. E.: **Processamento de Imagens Digitais**. Ed. Edgard Blucher – LTDA, São Paulo – SP, 2000.

HAYKIN, S. **Redes Neurais: Princípios e práticas**. 2. Ed., traduzido por Paulo Martins Engel, Porto Alegre: Bookmam, 2001.

MATLAB Versão 7.0. **Neural Network Toolbox 4.0.1 Release Notes**. The MathWorks, 2004.

RAPAPORT, S. I.: **Introdução a hematologia**. 2. Ed., traduzido por Ruth Moreira Leite, São Paulo: Roca, 1990.

SCHALKOFF, R. J.: **Artificial Neural Networks**. New York: McGraw-Hill, 1997.

TAFNER, M. A., XEREZ, M., FILHO, I. W. R.: **Redes Neurais Artificiais: Introdução e Princípios de Neurocomputação**. Ed da FURB, Blumenau, 1995.

VALLADA, E. P.: **Manual de Técnicas Hematológicas**. Ed. Atheneu, SP-Brasil, 1997.

VERRASTRO, T.: **Hematologia e Hemoterapia: Fundamentos de Morfologia, Fisiologia, Patologia e Clínica**. 2. Ed. Rio de Janeiro, 2005.

# DESIGUALDADE ISOPERIMÉTRICA

Anderson Bruneira Lopes, Fernando Luiz dos Reis, Jonhey Nazario Lucizani, Kelly  
Roberta Mazzutti Lübeck.

Universidade Estadual do Oeste do Paraná - UNIOESTE, ander.tk@gmail.com; Universidade Estadual  
do Oeste do Paraná - UNIOESTE, fer\_nando\_reis@hotmail.com; Universidade Estadual do Oeste do  
Paraná - UNIOESTE, jonheynazario@hotmail.com ; Universidade Estadual do Oeste do Paraná -  
UNIOESTE, kellyrobertaml@gmail.com

## Resumo:

A Desigualdade Isoperimétrica está associada a problemas que relacionam os perímetros de curvas (simples, fechadas e planas) com a área limitada pelo “interior” destas curvas. A relação estabelecida por esta desigualdade permite investigar questões do tipo: fixado um comprimento e uma área, é possível determinar uma curva suave por partes, fechada e simples com tais especificações? A Desigualdade Isoperimétrica estabelece um parâmetro para resolução destas. Uma primeira prova deste problema foi dada por K. Weierstrass, em 1870. A prova que apresentaremos aqui é devido a E. Schmidt (1939).

**Palavras-chave:** Geometria Diferencial, Geometria Plana, Curvas, Áreas.

## Apresentação:

Dentre todas as curvas simples fechadas no plano, existe alguma que tenha comprimento igual a 6 metros e que limita uma área de 3 metros quadrados? Em busca desta resposta, estaremos apresentando a seguir a Desigualdade Isoperimétrica. Para tanto, começamos com a apresentação de um lema, que também pode ser obtido utilizando-se o Teorema de Green, para áreas que são limitadas por uma curva simples, fechada, com orientação positiva.

**Lema 1.** *Seja  $\alpha : [a, b] \rightarrow \mathbb{R}^2$  uma curva regular por partes, fechada, simples, orientada positivamente e definida por  $\alpha(t) = (x(t), y(t))$ . Então*

$$A = - \int_a^b y(t)x'(t)dt = \int_a^b x(t)y'(t)dt = \frac{1}{2} \int_a^b (x(t)y'(t) - y(t)x'(t))dt,$$

onde  $A$  é a região limitada pela curva  $\alpha$ .

*Demonstração.* Inicialmente mostraremos que a primeira e a segunda integral são equivalentes. De fato,

$$\int_a^b x(t)y'(t)dt = \int_a^b (x(t)y(t))'dt - \int_a^b x'(t)y(t)dt,$$

pois,

$$\int_a^b (x(t)y(t))'dt = \int_a^b x'(t)y(t)dt + \int_a^b x(t)y'(t)dt,$$

mas,

$$\int_a^b (x(t)y(t))'dt = [xy(b) - xy(a)] = 0,$$

uma vez que a curva  $\alpha$  é fechada e, assim,  $\alpha(a) = \alpha(b)$ , ou seja,  $x(a) = x(b)$  e  $y(a) = y(b)$ .

Então

$$\int_a^b x(t)y'(t)dt = - \int_a^b y(t)x'(t)dt.$$

Das relações acima é imediato que a terceira integral, é resultado das duas primeiras integrais. Agora iremos provar que a área da região limitada pela curva  $\alpha$  é dada pela primeira integral. Consideremos inicialmente a Figura 1, onde a curva é formada de dois segmentos de reta paralelos ao eixo  $Oy$  e por dois arcos que podem ser descritos pelas funções

$$y = f_1(x) \text{ e } y = f_2(x), \quad x \in [x_0, x_1], \quad f_1 > f_2.$$

A área da região limitada pelas curvas é dada pelo limite das somas de Riemann, isto é :

$$A = \int_{x_0}^{x_1} f_1(x)dx - \int_{x_0}^{x_1} f_2(x)dx.$$

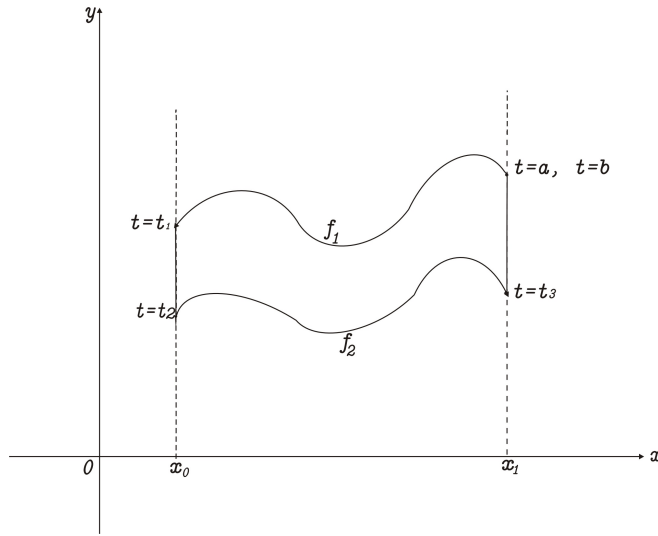


Figura 1: respectivamente

Pelo fato da curva ter orientação positiva, obtemos,  $\alpha(t) = (x(t), y(t))$  com

$$\alpha(t) = \begin{cases} (t, f_1(t)) & \text{para } a < t < t_1, \\ (x_0, t) & \text{para } t_1 < t < t_2, \\ (t, f_2(t)) & \text{para } t_2 < t < t_3, \\ (x_1, t) & \text{para } t_3 < t < b. \end{cases}$$

Assim,

$$\begin{aligned} A &= \int_{x_0}^{x_1} f_1 dx - \int_{x_0}^{x_1} f_2 dx \\ &= - \int_a^{t_1} y(t)x'(t)dt - \int_{t_2}^{t_3} y(t)x'(t)dt \\ &= - \int_a^{t_1} y(t)x'(t)dt - \int_{t_1}^{t_2} y(t)x'(t)dt - \int_{t_2}^{t_3} y(t)x'(t)dt - \int_{t_3}^b y(t)x'(t)dt \\ &= - \int_a^b y(t)x'(t)dt. \end{aligned} \tag{1}$$

Para o caso geral, basta observar que é possível dividir uma região limitada por uma curva suave por partes, simples e fechada em finitas regiões do tipo acima, e assim poderemos aplicar as fórmulas em cada uma delas. Para isto basta existir uma reta  $E$

no plano tal que a distância  $\rho(t)$  de  $\alpha(t)$  a esta reta é uma função com um número finito de pontos críticos, como ilustramos na Figura 2.

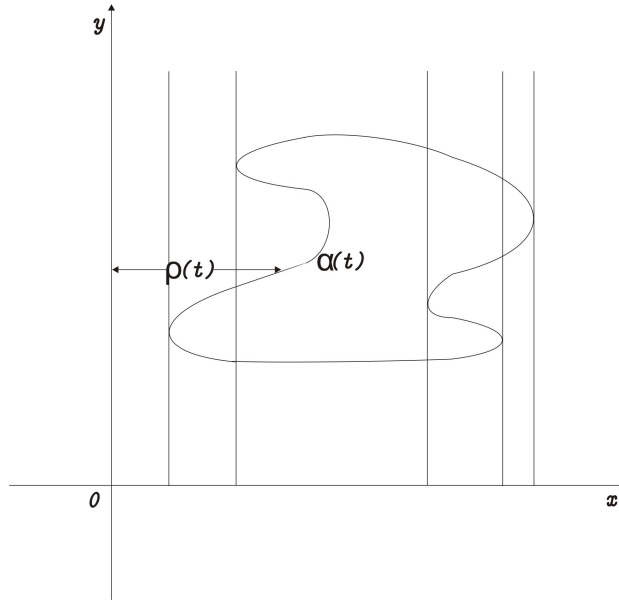


Figura 2: Função  $\rho$ : distância de  $\alpha(t)$  à reta  $E$ .

**Teorema 1 (Desigualdade Isoperimétrica).** *Seja  $C$  uma curva plana regular por partes, simples e fechada com comprimento  $l$ , e seja  $A$  a área da região limitada por  $C$ . Então*

$$l^2 - 4\pi A \geq 0, \quad (2)$$

*e verifica-se a igualdade se, e somente se,  $C$  é um círculo.*

*Demonstração.* Sejam  $E$  e  $E'$  duas retas paralelas que intersectam a curva fechada  $C$ , e considere o movimento de translação destas retas até que toquem a curva  $C$  pela primeira vez. Obtemos, assim, duas retas paralelas  $L$  e  $L'$ , tangentes a curva  $C$ , de forma que  $C$  está totalmente contida na faixa limitada por  $L$  e  $L'$ . Considere agora um círculo  $S^1$  que seja tangente a  $L$  e  $L'$  em  $\bar{Q}$  e  $\bar{P}$  respectivamente e não intercepta  $C$ . Seja  $O$  o centro de  $S^1$  e o eixo  $Ox$  perpendicular a  $L$  e  $L'$  (conforme Figura 3).

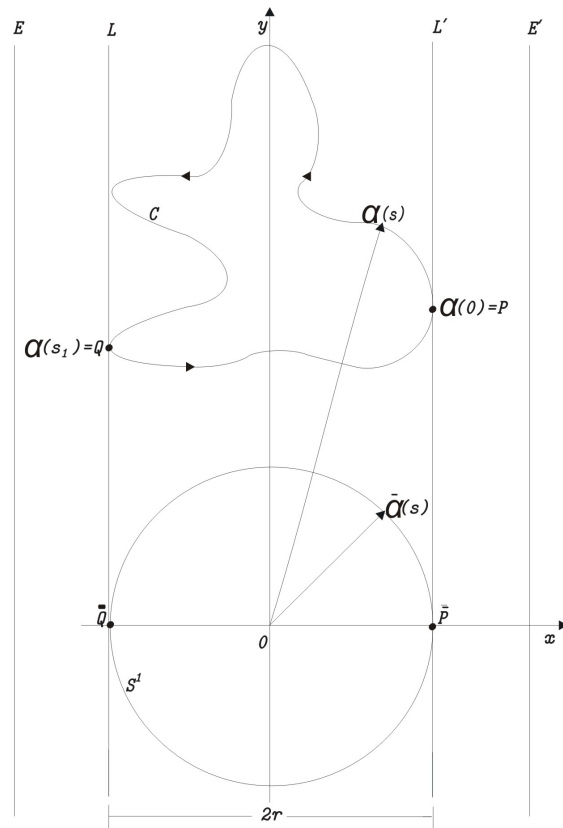


Figura 3: Representação das curvas  $C$  e  $S^1$ .

Seja  $\bar{\alpha}$  uma curva que traço descreve o traço do círculo  $S^1$  tangente em  $\bar{P}$  e  $\bar{Q}$ . Denotemos por  $r$  o raio de  $S^1$  e tomemos seu centro sendo a origem do sistemas de coordenadas. Podemos supor, sem perda de generalidade, que a curva  $C$  esteja parametrizada por comprimento de arco, tal que  $\alpha : [0, l] \rightarrow \mathbb{R}^2$ , sendo  $\alpha(0) = P$ , onde  $\alpha$  está orientada positivamente, e  $\alpha(s_1) = Q$  e, em relação a esse sistema de coordenadas,  $\alpha$  seja dada por

$$\alpha(s) = (x(s), y(s)) \quad x \in [0, l], \text{ com } \alpha(0) = \alpha(l).$$

Agora parametrizemos  $\bar{\alpha} : [0, l] \rightarrow \mathbb{R}^2$ , por

$$\bar{\alpha}(s) = (\bar{x}(s), \bar{y}(s)),$$

tal que

$$\bar{x}(s) = x(s)$$

e

$$\bar{y}(s) = \begin{cases} \sqrt{r^2 - (x(s))^2}, & 0 \leq s \leq s_1, \\ -\sqrt{r^2 - (x(s))^2}, & s_1 \leq s \leq l, \end{cases}$$

Note que  $\bar{\alpha}(0) = \bar{\alpha}(l) = \bar{P}$ .

Agora utilizando o Lema 1, e denotando por  $A$  a região limitada por  $C$  e por  $\bar{A}$  a região limitada por  $S^1$ , temos:

$$A = \int_0^l x(s)y'(s)ds,$$

e

$$\bar{A} = - \int_0^l \bar{y}(s)x'(s)ds = \pi r^2.$$

Então,

$$\begin{aligned} A + \pi r^2 &= \int_0^l x(s)y'(s) - \bar{y}(s)x'(s)ds \\ &\leq \int_0^l \sqrt{(x(s)y'(s) - \bar{y}(s)x'(s))^2}ds \\ &= \int_0^l \sqrt{(x(s))^2(y'(s))^2 - 2x(s)y'(s)\bar{y}(s)x'(s) + (\bar{y}(s))^2(x'(s))^2}ds. \end{aligned}$$

Observe que,

$$-2cd \leq c^2 + d^2, \text{ onde } c, d \in \mathbb{R},$$

tomemos  $c = x'(s)x(s)$  e  $d = \bar{y}(s)y'(s)$ , logo

$$-2x'(s)x(s)\bar{y}(s)y'(s) \leq (x'(s)x(s))^2 + (\bar{y}(s)y'(s))^2 \quad (3)$$

deste modo segue que,



$$\begin{aligned}
A + \pi r^2 &\leq \int_0^l \sqrt{(x(s)y'(s))^2 + (x(s)x'(s))^2 + (\bar{y}(s)y'(s))^2 + (\bar{y}(s)x'(s))^2} ds \\
&= \int_0^l \sqrt{((x(s))^2 + (\bar{y}(s))^2)((x'(s))^2 + (y'(s))^2)} ds \\
&= \int_0^l \sqrt{(x(s))^2 + (\bar{y}(s))^2} ds \\
&= \int_0^l \sqrt{(\bar{x}(s))^2 + (\bar{y}(s))^2} ds \\
&= \int_0^l r ds = lr.
\end{aligned} \tag{4}$$

Usando o fato de que a média geométrica de dois números positivos é menor ou igual a média aritmética, ou seja,

$$cd \leq \frac{1}{2}(c^2 + d^2),$$

então,

$$\sqrt{A}\sqrt{\pi r^2} \leq \frac{1}{2}(A + \pi r^2) \leq \frac{1}{2}lr, \tag{5}$$

logo,

$$4\pi Ar^2 \leq l^2 r^2.$$

Portanto,

$$l^2 - 4\pi A \geq 0.$$

Consideremos agora, a igualdade em (2), ou seja,  $l^2 = 4\pi A$ . Usando (5) temos que  $A\pi r^2 = \frac{1}{4}(A + \pi r^2)^2$ , o que implica diretamente que  $A = \pi r^2$ , portanto  $l = 2\pi r$ . A distância entre  $E'$  e  $F'$  é  $2r$ , não dependendo da escolha da direção destas retas, e além disso, a igualdade em (4) implica que vale a igualdade em (3), ou seja,

$$x'(s)x(s) = c = -d = -\bar{y}(s)y'(s).$$

Daí segue que,

$$(x'(s)x(s))^2 + (y'(s)x(s))^2 = (y'(s)x(s))^2 + (\bar{y}(s)y'(s))^2.$$

Então,

$$(x(s))^2[(x'(s))^2 + y'(s)^2] = [(x(s))^2 + (\bar{y}(s))^2](y'(s))^2 = r^2(y'(s))^2,$$

logo

$$(x(s))^2 = r^2(y'(s))^2. \tag{6}$$

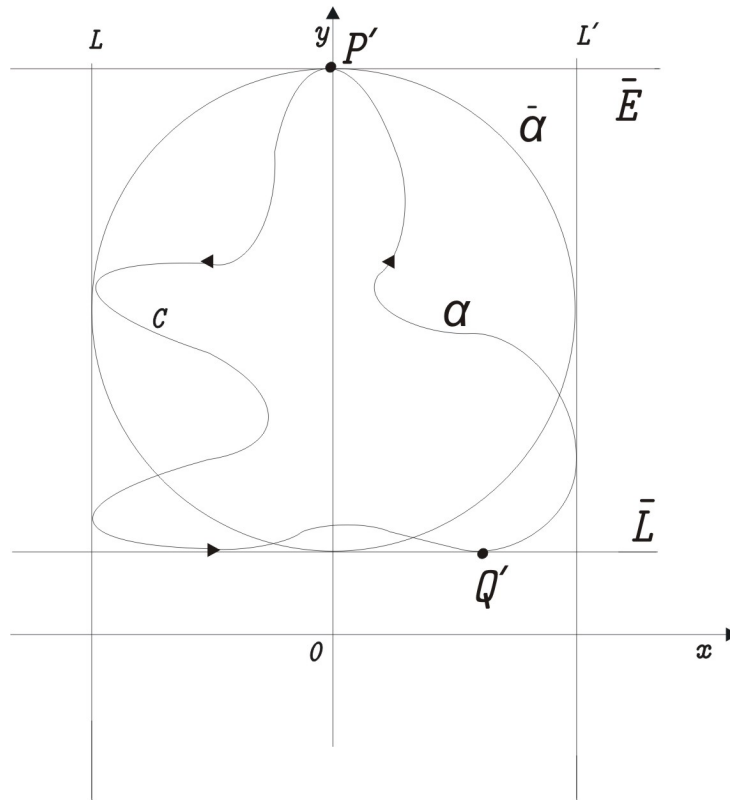


Figura 4: respectivamente

Sejam  $P'$  e  $Q'$  pontos sobre o traço da curva  $\alpha$ , tais que suas ordenadas sejam valores de mínimo e máximo globais respectivamente. Do fato de  $\alpha$  ser regular em todos os pontos, implica que as retas tangentes a  $P'$  e  $Q'$  são paralelas ao eixo  $Ox$ . Tomando  $\bar{E}$  e  $\bar{L}$ , as respectivas tangentes, a distância entre estas retas è  $2r$ . Por um movimento rígido sobre o círculo  $\bar{\alpha}$ , dado pelo vetor  $(0, a)$ , de modo que o círculo fique tangente as

retas  $\bar{E}$  e  $\bar{L}$  (conforme a Figura 4), temos

$$(y(s) + a)^2 = r^2(x'(s))^2. \quad (7)$$

Portanto, somando as equações (6) e (7), obtemos

$$(x(s))^2 + (y(s) + a)^2 = r^2[(x'(s))^2 + (y'(s))^2] = r^2,$$

donde segue-se o resultado.

Agora, para responder ao nosso problema inicial, ou seja, verificar se existe dentre todas as curvas simples fechadas no plano alguma que tenha comprimento igual a 6 metros e que limita uma área de 3 metros quadrados, vamos aplicar a fórmula da desigualdade isoperimétrica, ou seja,

$$l^2 - 4\pi A \geq 0.$$

Tomando  $l = 6$  e  $A = 3$ , temos que

$$6^2 - 4\pi 3 \geq 0,$$

logo,

$$36 \geq 12\pi,$$

o que é um absurdo. Portanto não existe uma curva com tais características.

### **Referências Bibliográficas:**

CARMO, M.P.: **Differential Geometry of Curves and Surfaces**, Prentice-Hall, Englewood Cliffs, New Jersey: 1976.

ARAÚJO, P.V.: **Geometria Diferencial**, IMPA, Rio de Janeiro: 2004.

CARMO, M.P.: **Elementos de Geometria Diferencial**, Ao livro técnico S. A., Rio de Janeiro: 1971.

LIMA, E.L.: **Álgebra Linear**, Ciência Moderna, Rio de Janeiro: 2003.

ALENCAR, H., SANTOS, W.: **Geometria Diferencial das Curvas Planas**, Universidade Federal de Goiás: 2002.

# MULTICOLINEARIDADE EM ANÁLISE DE REGRESSÃO

Luís Fernando Moreira  
Doutorando em Administração [Finanças] – PPGA – USP  
Mestre em Administração [Finanças] – PPGA – UFRGS  
moreiralfm@yahoo.com.br

*Felix qui potuit rerum cognoscere causas.*  
Virgílio

## **Resumo:**

A regressão múltipla visa entre outras coisas estabelecer a importância relativa e a magnitude do efeito das variáveis preditoras sobre a variável dependente, identificar preditores que deveriam ser eliminados do modelo e apontar a necessidade de inclusão de variáveis preditoras ainda não participantes do modelo. Na ausência de multicolinearidade esses objetivos podem ser facilmente alcançados. Infelizmente, nas ciências sociais, em função dos experimentos não serem controlados, a existência de multicolinearidade é a regra. Esse trabalho apresenta um exemplo de como identificar e tratar a multicolinearidade

**Palavras chave:** Regressão múltipla, multicolinearidade, econometria.

## **Abstract:**

The multiple regression model aims among other things to establish the relative importance and magnitude of the effect of a given predictor variable on the response variable, identify predictor variables that should be dropped from the model, and point out the need to include predictor variables not considered to inclusion yet. In the absence of multicollinearity these goals can be easily attainable. Unfortunately, in social sciences as the experiments are not controlled the presence of multicollinearity is the rule. This paper shows an example of how to identify and deal with multicollinearity.

**Key words:** Multiple regression, multicollinearity, econometrics.

---

## 1. Introdução

---

Nos modelos de regressão múltipla de primeira ordem, quando as variáveis preditoras não estão correlacionadas, o efeito atribuído a cada uma delas é o mesmo, não importa quais variáveis estejam inclusas no modelo. Ainda nesse contexto, a contribuição marginal para a redução da soma dos quadrados da regressão de uma das preditoras quando a outra está no modelo é a mesma de quando ela aparece sozinha no modelo (Kutner et al., 2004). Mais formalmente:

$$SSR(x_1) = SSR(x_1 / x_2) \quad (I)$$

Em modelos de regressão com duas ou mais variáveis explicativas é usual que tais variáveis apresentem algum tipo de interdependência. A essa relação de interdependência chamamos multicolinearidade (Fonseca, 2003). A presença de multicolinearidade independe da existência de relação de dependência entre a variável dependente e os termos independentes (Ferrar & Glauber, 1967).

De acordo com Kutner et al. (2004) a regressão múltipla visa entre outras coisas estabelecer a importância relativa e a magnitude do efeito das variáveis preditoras sobre a variável dependente, identificar preditores que deveriam ser eliminados do modelo e apontar a necessidade de inclusão de variáveis preditoras ainda não participantes do modelo. Na ausência de multicolinearidade esses objetivos podem ser facilmente alcançados. Infelizmente, nas ciências sociais, em função dos experimentos não serem controlados, a existência de multicolinearidade é a regra.

A relação de interdependência quase sempre presente pode ser facilmente identificada; medi-la, entretanto, não é tarefa tão simples. De modo bastante pragmático identificar a presença de multicolinearidade em pouco contribui no sentido de saber se isso representa um problema ou não para o modelo. É preciso aquilatar a magnitude da multicolinearidade para saber se ela é severa – caso em que exige tratamento – ou não; surge então o problema de saber qual a medida adequada a ser utilizada (Gujarati, 2000).

---

## 2. O Que a Multicolinearidade Causa

---

Mas afinal o que ocorre se houver multicolinearidade severa? Primeiramente conforme observado por Ferrar e Glauber (1967) em um artigo seminal sobre o assunto, se o conjunto de variáveis independentes for totalmente interdependente os coeficientes da regressão não poderão ser estimados uma vez que a matriz resultante da multiplicação

da matriz transposta das variáveis independentes pela matriz das variáveis independentes será singular e não será possível a inversão dessa matriz necessária para cálculo dos coeficientes da regressão, ou seja, não existe  $(X^T X)^{-1}$ . Nesse caso a multicolinearidade é, obviamente, severa e o modelo deve ser revisto.

O problema é mais sutil quando a matriz  $X^T X$  não é singular. Nesse caso os coeficientes serão estimados normalmente. Uma vez que a multicolinearidade implica que as variáveis preditoras estão linearmente relacionadas, o problema usualmente implica em alta correlação entre as variáveis preditoras, mas nem sempre isso ocorre (Tamhane & Dunlop, 2000).

A presença de multicolinearidade severa faz com que os coeficientes não reflitam nenhum efeito particular da variável explicativa na regressora, pois o coeficiente estará contaminado pelo efeito de outras variáveis e apenas um efeito parcial lhe poderá ser atribuído. Em outros termos isso equivale a dizer que os coeficientes terão valores que vão depender da existência ou não de outras variáveis explicativas no modelo. Além disso, a variabilidade amostral dos coeficientes é grande gerando grandes mudanças na magnitude de  $\hat{\beta}_j$  por conta de pequenas alterações nos dados,  $s^2\{\hat{\beta}_j\}$  terá valores elevados e crescerá na medida em que novos preditores forem incluídos (Kutner et al., 2004; Tamhane & Dunlop, 2000).

Quanto à soma dos quadrados da regressão (SSR), como a informação de um coeficiente já está contida em outro, a redução por conta da adição da variável explicativa é modesto e, como no caso dos coeficientes, seu valor vai depender de quais outras variáveis estão no modelo não podendo ser atribuído a nenhuma delas em particular. Em casos especiais, em que as variáveis preditoras estão correlacionadas entre si, mas uma delas não tem correlação com a variável dependente; pode até ocorrer aumento na SSR (Kutner et al., 2004). Os coeficientes de determinação parcial também serão afetados uma vez que a SSR é utilizada em seu cálculo.

Pode ocorrer também de haver uma relação de regressão – estatística F da regressão apontando que pelo menos um  $\hat{\beta}_j$  é diferente de zero – e ainda assim os coeficientes não serem estatisticamente significantes (Kutner et al., 2004; Tamhane & Dunlop, 2000).

Por outro lado, se o objetivo for tão somente prever a variável dependente o problema da multicolinearidade é menos relevante haja vista sua presença, em geral, não impedir que um ajuste seja obtido e tampouco afetar as inferências sobre novas observações ou sobre a resposta média da variável dependente. Mais tecnicamente, o MSE é reduzido na medida em que novas variáveis explicativas são adicionadas e a  $s^2\{\hat{Y}_h\}$  é reduzida (Kutner et al., 2004).

---

### 3. Um Comentário Adicional a Respeito dos Coeficientes ( $\hat{\beta}_j$ )

---

É muito comum em se tratando de regressão múltipla o desejo de interpretar o coeficiente  $\hat{\beta}_j$  observando diretamente seu valor e atribuindo maior magnitude de seu efeito sobre a variável dependente em função disso. Cabe frisar em primeiro lugar que as magnitudes das estimativas de  $\beta_j$  não podem ser comparadas diretamente, uma vez que guardam relação com as unidades das variáveis preditoras. É importante ter em mente que a unidade de  $\hat{\beta}_j$  é dada pela divisão da unidade de  $y$  pela unidade de  $x_j$ , como as variáveis preditoras não costumam ter a mesma unidade de medida a comparação direta é incoerente. A solução consiste em padronizar as variáveis antes da regressão, o que permite que os coeficientes obtidos sejam comparados entre si (Tamhane & Dunlop, 2000).

Outro deslize usual diz respeito ao exame da estatística a seguir:

$$t^* = \frac{\hat{\beta}_j}{s\{\hat{\beta}_j\}} \quad (\text{II})$$

para decidir se  $\hat{\beta}_j$  é igual a zero para um  $k$  qualquer. Nesse caso, conforme observam Kutner et al. (2004), o nível de significância deve ser controlado no nível global da regressão, isto é levando em conta o número de variáveis preditoras. Porém, mesmo com um procedimento de inferência simultânea, há problemas quando as variáveis dependentes estão fortemente correlacionadas, uma vez que cada teste  $t^*$  é um teste de significância marginal (Kutner et al., 2004).

---

### 4. Diagnosticando a Multicolinearidade

---

Há alguns indicadores da presença de multicolinearidade que são calculados pela maioria dos softwares que fazem regressão múltipla. Um indicador da provável presença de multicolinearidade ocorre quando como um conjunto as variáveis preditoras



são significantes, ao passo que individualmente não (Newbold, 1994). Isso reflete o caso em que a estatística F da regressão é altamente significativa, mas o valor da estatística  $t^*$  não aponta algum  $\hat{\beta}_j$  significativo. Isso ocorre porque F mede a relação de regressão do conjunto de preditores contra a variável dependente, ao passo que  $t^*$  mede a contribuição de uma única variável explicativa (McClave & Sincich, 2000). Some-se a isso um  $R^2$  alto e a sintomatologia está completa (Gujarati, 2000).

Outro indicador é a matriz de correlação de ordem zero das variáveis explicativas, embora deva ser enfatizado que alta correlação simples entre as variáveis é condição suficiente, mas não necessária para que exista multicolinearidade, portanto sua presença não é indício infalível de multicolinearidade salvo no caso de apenas duas variáveis preditoras (Fonseca, 2003; Gujarati, 2000). O exame das correlações parciais pode também ser útil (Gujarati, 2000).

Ainda de acordo com Gujarati (2000) como a multicolinearidade se manifesta em função dos vetores das variáveis explicativas serem linearmente dependentes, um meio de descobrir qual variável  $x$  se relaciona com outras é regressar  $x_k$  sobre as demais e calcular o  $R^2$  para cada regressão. Calcular a seguinte estatística F, onde  $n$  é o tamanho da amostra e  $k$  o número de variáveis explicativas incluindo o intercepto:

$$F = \frac{\frac{R^2_{x_1, x_2, \dots, x_k}}{(k-2)}}{\frac{(1 - R^2_{x_1, x_2, \dots, x_k})}{(n-k+1)}} \quad (\text{III})$$

Se o F calculado é significativo há multicolinearidade. Porém se a multicolinearidade envolver associações lineares complexas esse processo de identificação pode ser infrutífero. Há ainda uma regra prática que estabelece que o  $R^2$  obtido na regressão auxiliar só é preocupante se for maior que o  $R^2$  global, isso evita o teste de cada regressão auxiliar.

Outra abordagem para diagnosticar e medir a multicolinearidade é através da análise dos autovalores (eigenvalues) e de seu derivado índice de condição (IC). Autovalores muito próximos de zero podem indicar multicolinearidade. O índice de condição é dado por:

$$IC = \sqrt{\frac{A_{Max}}{A_{Min}}} \quad (\text{IV})$$

Onde  $A_{Max}$  é o valor do maior autovalor e  $A_{Min}$  é o valor do menor autovalor. Entre 10 e 30 há multicolinearidade moderada, valores acima de 30 implicam multicolinearidade severa (Gujarati, 2000).

O fator de inflação da variância (VIF) também é utilizado para medir a multicolinearidade, tal fator é calculado conforme demonstrado no quadro 1.

**Quadro 1 – Como é calculado o VIF**

<p>Seja:</p> $y_i^* = \frac{y_i - \bar{y}}{s_y} \quad (i = 1, \dots, n \text{ observações})$ $x_{ij}^* = \frac{x_{i,j} - \bar{x}_j}{s_{x_j}} \quad (i = 1, \dots, n \text{ observações}; j = 1, \dots, k \text{ variáveis})$ <p>Seja <math>X^*</math> a matriz <math>k \times k</math> sem a coluna de 1 correspondente a constante</p> $\frac{1}{n-1} X^{*T} X^* = R = \begin{bmatrix} 1 & r_{x_1 x_2} & \dots & r_{x_1 x_k} \\ r_{x_2 x_1} & 1 & \dots & r_{x_2 x_k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{x_k x_1} & r_{x_k x_2} & \dots & 1 \end{bmatrix}$ $R^{-1} = \begin{bmatrix} VIF_{11} & \dots & \dots & \dots \\ \vdots & VIF_{11} & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \dots & \dots & VIF_{n,k} \end{bmatrix}$
--

Pode ser demonstrado que o VIF é igual a:

$$VIF_j = \frac{1}{1 - r_j^2} \quad (j = 1, 2, \dots, k) \quad (V)$$

Quando as variáveis estão linearmente relacionadas o coeficiente de determinação múltipla  $r_j^2$  será próximo de 1,0 e o VIF será alto; esse coeficiente de determinação é obtido quando se faz a regressão de  $x_j$  sobre as variáveis k-1 remanescentes. Quanto maior o VIF maior a variância do  $\hat{\beta}_j$ , valores maiores que 10 correspondem a um coeficiente de determinação múltipla  $r_j^2 > 0,90$  e são considerados inaceitáveis. (Kutner et al., 2004; Tamhane & Dunlop, 2000).

Finalmente Ferrar e Glauber (1967) indicam um índice, aqui chamado de FG, que, segundo os autores, gera uma medida objetiva para a multicolinearidade. A fórmula de Ferrar e Glauber é a seguinte:

$$FG = - \left[ N - 1 - \frac{1}{6}(2n + 5) \right] \ln |R| \quad (\text{VI})$$

Nessa fórmula  $N$  é o tamanho da amostra,  $n$  é o número de variáveis e  $\ln |R|$  é o logaritmo natural do determinante da matriz  $R$  que aparece no quadro 1. Essa estatística tem distribuição aproximadamente qui-quadrado com  $\nu = 1/2n(n-1)$  graus de liberdade. Como a distribuição  $\chi^2$  segue uma escala conhecida essa seria uma estatística das mais úteis para detectar e quantificar de modo realmente objetivo a multicolinearidade; estranhamente até onde foi possível constatar nenhum pacote estatístico trás em seu bojo essa estatística. No exemplo apresentado na próxima seção ela será calculada.

---

### 5. Um Exemplo Ilustrativo

---

O problema a seguir visa ilustrar os aspectos tratados nas seções anteriores, além de alguns outros ainda não contemplados e não necessariamente relacionados a multicolinearidade, mas usuais na análise de regressão. Os dados para análise foram obtidos de McClave & Sincich (2000), a amostra original continha 129 dados, foram eliminados 16 deles por falta de alguma variável explicativa ou por possuírem algum valor padronizado superior a 2,5, portanto, a análise inicial parte de uma base com 113 informações a respeito de carros de marcas e modelos variados. As variáveis que entrarão inicialmente no modelo são as seguintes:  $y$  = Preço;  $x_1$  = Aceleração em segundos de 0-60 milhas por hora;  $x_2$  = Aceleração em segundos para percorrer ¼ de milha;  $x_3$  = Velocidade máxima em milhas por hora;  $x_4$  = Distância de frenagem em pés a 70 milhas por hora;  $x_5$  = Consumo em milhas por galão;  $x_6$  = Aderência em g's durante a curva. A idéia inicial é optar por um modelo de primeira ordem linear nos parâmetros no seguinte formato:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \beta_6 x_6 \quad (\text{VII})$$

Os dados que serviram de base para a análise estão disponíveis no anexo 1.

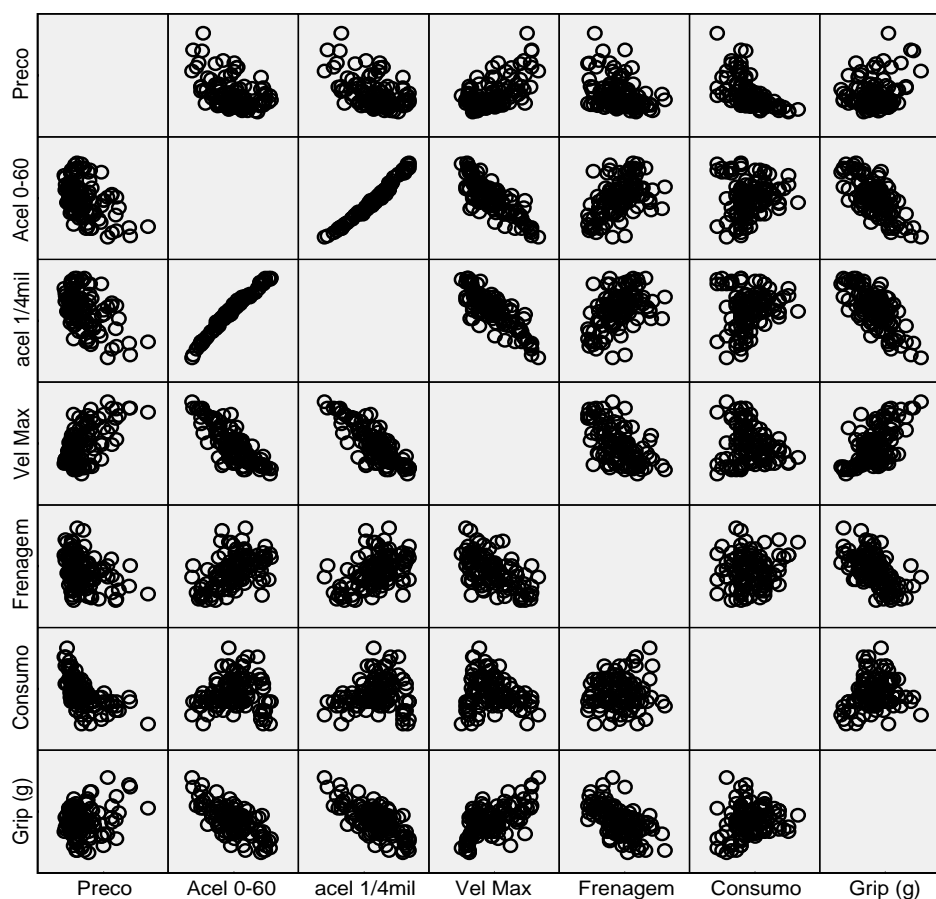
Sempre que se inicia a análise de um conjunto de dados é recomendável explorá-los através de algumas estatísticas descritivas a tabela 1 apresenta algumas dessas estatísticas.

**Tabela 1 – Estatísticas Descritivas**

		y	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	x <sub>4</sub>	x <sub>5</sub>	x <sub>6</sub>
N	Valid	113	113	113	113	113	113	113
	Missing	362	362	362	362	362	362	362
Mean		33695,86	8,02	16,148	126,78	184,65	18,89	,7951
Std. Deviation		18021,771	1,679	1,2636	17,263	12,258	3,363	,06239
Skewness		1,606	-,050	-,451	,517	,222	,474	,200
Std. Error of Skewness		,227	,227	,227	,227	,227	,227	,227
Kurtosis		2,752	-,443	-,143	-,697	-,266	,254	,231
Std. Error of Kurtosis		,451	,451	,451	,451	,451	,451	,451

Um Scatter Plot desses dados já padronizados está na figura 1, é clara a forte associação entre x1 e x2.

**Figura 1 – Box Plot das variáveis explicativas.**



Em seguida, com os dados padronizados, uma primeira regressão incluindo todas as variáveis e dados disponíveis foi realizada e os resultados são apresentados abaixo:

**Tabela 2 – Regressão com MINITAB**

The regression equation is  
 Preço = - 0,0000 + 0,658 Acel 0-60 - 0,676 acel 1/4mil + 0,538 Vel Max  
 - 0,0072 Frenagem - 0,429 Consumo - 0,062 Grip (g)

Predictor	Coef	SE Coef	T	P	VIF
Constant	-0,00000	0,06149	-0,00	1,000	
Acel 0-60	0,6584	0,4880	1,35	0,180	62,4
acel 1/4mil	-0,6761	0,5197	-1,30	0,196	70,8
Vel Max	0,5382	0,1243	4,33	0,000	4,1
Frenagem	-0,00721	0,08396	-0,09	0,932	1,8
Consumo	-0,42940	0,07688	-5,59	0,000	1,5
Grip (g)	-0,0625	0,1106	-0,56	0,573	3,2

S = 0,653659 R-Sq = 59,6% R-Sq(adj) = 57,3%  
 PRESS = 52,3401 R-Sq(pred) = 53,27%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	6	66,709	11,118	26,02	0,000
Residual Error	106	45,291	0,427		
Total	112	112,000			

Aqui a atenção deve se voltar primeiramente para o fato de haver uma relação de regressão entre as variáveis, a estatística F na análise da variância é significativa, porém nem todas as variáveis no modelo são relevantes conforme denota o valor p do teste t para os  $\hat{\beta}_j$ . O fator de inflação da variância VIF confirma claramente que as variáveis x1 e x2 apresentam colinearidade o valor da tolerância calculado pelo SPSS aponta no mesmo sentido com valores de 0,016 e 0,014 respectivamente para x1 e x2. A estatística de Ferrar e Glauber (FG) foi calculada conforme demonstrado no quadro 2.

**Quadro 2 – Estatística de Ferrar e Glauber**

Seja:

$$y_i^* = \frac{y_i - \bar{y}}{s_y} \quad (i = 1, \dots, n)$$

$$x_i^* = \frac{x_{i,j} - \bar{x}_j}{s_{x_j}} \quad (i = 1, \dots, n; j = 1, \dots, k)$$

Seja  $X^*$  a matriz  $k \times k$  sem a coluna de 1 correspondente a constante

$$\frac{1}{n-1} X^{*T} X^* = R = \begin{bmatrix} 1 & r_{x_1 x_2} & \dots & r_{x_1 x_k} \\ r_{x_2 x_1} & 1 & \dots & r_{x_2 x_k} \\ \vdots & \vdots & \ddots & \vdots \\ r_{x_k x_1} & r_{x_k x_2} & \dots & 1 \end{bmatrix}$$

$$FG = - \left[ N - 1 - \frac{1}{6} (2n + 5) \right] \ln |R|$$

Essa estatística produziu um valor de 758 contra um valor crítico de 12,6 também indicando forte presença de multicolinearidade.

Por outro lado, as variáveis velocidade máxima e consumo são as únicas que produziram estimativas significantes. O passo seguinte foi realizar uma regressão utilizando o algoritmo best subset do MINITAB que produziu a saída da tabela 3. O resultado indica que 58% da variância da variável dependente na amostra é explicada pelas variáveis  $x_3 =$  Velocidade máxima em milhas por hora e  $x_5 =$  Consumo em milhas por galão. A tabela apresenta também o CP de Mallows igual a 1,0 o que implica um erro quadrático médio total pequeno.

**Tabela 3 – Regressão usando o algoritmo “best subsets” do MINITAB**

Response is Preco						Mallows					
Vars	R-Sq	R-Sq(adj)	C-p	S	x						
					1	2	3	4	5	6	
1	36,3	35,7	58,0	0,80177			X				
1	32,9	32,3	66,9	0,82290						X	
1	24,9	24,2	87,9	0,87052		X					
1	20,9	20,2	98,2	0,89315	X						
1	11,7	10,9	122,6	0,94415					X		
2	58,8	58,0	1,0	0,64774			X			X	
2	50,4	49,5	23,1	0,71093		X				X	
2	49,2	48,3	26,2	0,71925	X					X	
2	43,2	42,2	41,8	0,76022						X	X
2	41,2	40,1	47,2	0,77393					X	X	
3	58,9	57,8	2,7	0,64988	X		X			X	
3	58,9	57,7	2,8	0,65012			X		X	X	X
3	58,8	57,7	3,0	0,65053		X	X			X	
3	58,8	57,7	3,0	0,65066			X	X	X		
3	51,4	50,1	22,4	0,70669		X		X	X		
4	59,4	57,9	3,3	0,64864	X	X	X			X	

Com base nessas novas informações uma nova regressão agora incluindo apenas a velocidade máxima e o consumo foi conduzida e as estatísticas recalculadas. Os resultados da regressão aparecem na tabela 4.

A estatística de Ferrar e Glauber para esses dados resultou em um valor calculado de 3,52 contra um valor crítico de 5,99; o valor da estatística de tolerância calculada no SPSS foi de 0,969. Com essas duas variáveis no modelo os números da regressão melhoraram substancialmente e foi eliminado o problema da multicolinearidade.

**Tabela 4 – Regressão com  $x_3$  e  $x_5$  resultado do MINITAB.**

The regression equation is  
 Preço = - 0,0000 + 0,517 Vel Max - 0,482 Consumo

Predictor	Coef	SE Coef	T	P	VIF
Constant	-0,00000	0,06093	-0,00	1,000	
Vel Max	0,51713	0,06219	8,32	0,000	1,0
Consumo	-0,48197	0,06219	-7,75	0,000	1,0

S = 0,647738 R-Sq = 58,8% R-Sq(adj) = 58,0%  
 PRESS = 49,2308 R-Sq(pred) = 56,04%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	2	65,848	32,924	78,47	0,000
Residual Error	110	46,152	0,420		
Lack of Fit	101	43,680	0,432	1,57	0,235
Pure Error	9	2,472	0,275		
Total	112	112,000			

Lack of fit test

Possible interaction in variable Vel Max (P-Value = 0,042 )

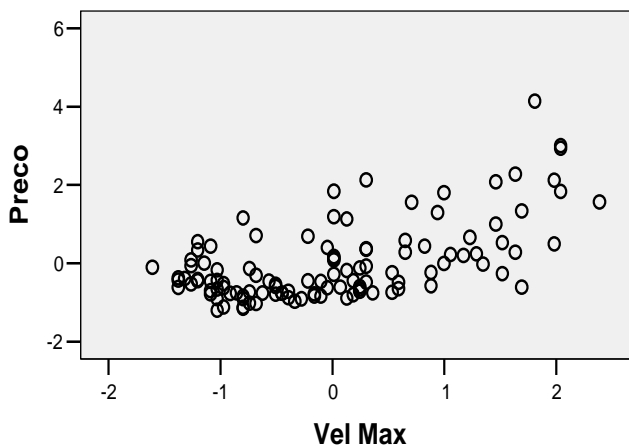
Possible curvature in variable Consumo (P-Value = 0,047 )

Possible lack of fit at outer X-values (P-Value = 0,018)

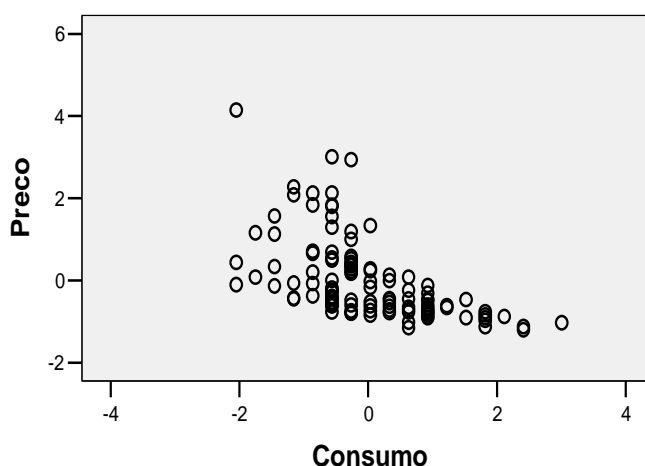
Overall lack of fit test is significant at P = 0,018

Porém a natureza das coisas é tal que basta resolver um problema para nos depararmos com outro, os resultados parecem indicar uma relação não linear e interação em certas variáveis. As figuras 2 e 3 a seguir mostram scatter plot do preço contra a velocidade máxima e consumo respectivamente.

**Figura 2 – Scatter plot preço versus velocidade máxima**



**Figura 3 – Scatter plot preço versus consumo**



Com base nas figuras 2 e 3 e na informação da regressão anterior foram criadas as seguintes variáveis transformadas assim definidas  $y' = e^y$  = Exponencial do preço;  $x_7 = e^{x_3}$  = Velocidade máxima em milhas por hora;  $x_8 = e^{x_5}$  = Exponencial do consumo em milhas por galão;  $x_9 = e^{x_3} e^{x_5}$  = Produto das variáveis  $e^{x_3}$  e  $e^{x_5}$ ;  $x_{10} = x_3 x_5$  = Produto da velocidade pelo consumo;  $x_{11} = x_3^2$  = Quadrado da velocidade;  $x_{12} = x_5^2$  = Quadrado do consumo;  $x_{13} = x_3^2 x_5^2$  = Produto do quadrado da velocidade pelo quadrado do consumo. Com essas variáveis explicativas e todas as anteriores foi conduzida uma regressão usando o algoritmo best subset do MINITAB alguns resultados parciais aparecem na tabela 5.

**Tabela 5 – Resultado parcial da regressão usando o algoritmo “best subsets” do MINITAB**

Vars	Response is Preço		Mallows		x												
	R-Sq	R-Sq(adj)	C-p	S	1	2	3	4	5	6	7	8	9	10	11	12	13
4	65,6	64,3	1,6	0,59761							X	X		X			X
4	64,9	63,6	3,7	0,60362			X		X			X		X			
4	64,8	63,5	4,0	0,60445			X		X					X			X
4	64,3	63,0	5,4	0,60836			X					X	X				X
4	64,2	62,9	5,8	0,60946			X					X	X				X
5	66,3	64,8	1,2	0,59364							X	X	X	X			X
5	65,9	64,3	2,6	0,59752			X				X	X	X				X
5	65,8	64,2	2,9	0,59830							X	X	X	X			X
5	65,8	64,2	3,0	0,59857			X					X	X		X		X
5	65,8	64,2	3,0	0,59864							X	X	X	X	X		X
6	66,8	64,9	1,8	0,59229			X				X	X	X	X			X
6	66,7	64,9	2,0	0,59275							X	X	X	X	X		X
6	66,6	64,8	2,3	0,59367							X	X	X	X	X		X
6	66,5	64,6	2,9	0,59533			X				X	X	X	X	X		X
6	66,4	64,5	2,9	0,59552							X	X	X	X		X	X

Os demais resultados apresentavam resultados superiores a 2,0 para o CP de Mallows e não foram reportados por essa razão. Foram testados os três modelos destacados na tabela 5 e o melhor ajuste foi obtido com a regressão com CP de Mallows de 1,6 cuja



saída do MINITAB é apresentada na tabela 6. Esse modelo parece ser o melhor por combinar as seguintes características: (1) Estatística F com valor 51,40 a maior dos três modelos, (2) todos os valores p das estimativas de  $\hat{\beta}_k$  altamente significantes; (3) Valor do  $R^2_{ajustado}$  0,643 apenas 0,006 inferior ao valor máximo obtido em outras regressões, (4) não há indicação de *lack of fit* ( $p > 0,10$ ), (5) com intervalo de confiança de 95%, nenhum dos  $\hat{\beta}_j$  tem valor nulo, (6) o Cp de Mallows é de apenas 1,6 implicando um erro quadrático médio pequeno e um modelo parcimonioso, (7) índice de condição com valor máximo de 5,814, (8) índice de inflação da variância com valor máximo de 3,2, (9) tolerância com valor mínimo de 0,31, e (10) correlações *zero order* e parciais aceitáveis entre os regressores.

O único indicador que parece apresentar valores preocupantes é a estatística de Ferrar e Glauber que retornou um valor FG de 172,12 contra um valor crítico de 9,49 apontando que ainda existe multicolinearidade. Tal resultado para essa estatística parece contra intuitivo em função de todas as características positivas e mais especificamente dos valores aceitáveis dos demais indicadores da presença e intensidade da multicolinearidade.

**Tabela 6 – Resultado da regressão do modelo final.**

The regression equation is  
 Preço = - 1,06 + 0,239 Vel e + 0,507 Cons e - 0,323 VxC - 0,187 sqVxC

Predictor	Coef	SE Coef	T	P	VIF
Constant	-1,05523	0,09543	-11,06	0,000	
Vel e	0,23863	0,03498	6,82	0,000	1,6
Cons e	0,50733	0,06658	7,62	0,000	3,2
VxC	-0,32284	0,09081	-3,56	0,001	2,6
sqVxC	-0,18729	0,04255	-4,40	0,000	2,8

S = 0,597613 R-Sq = 65,6% R-Sq(adj) = 64,3%

PRESS = 45,7517 R-Sq(pred) = 59,15%

Analysis of Variance

Source	DF	SS	MS	F	P
Regression	4	73,429	18,357	51,40	0,000
Residual Error	108	38,571	0,357		
Lack of Fit	99	36,099	0,365	1,33	0,341
Pure Error	9	2,472	0,275		
Total	112	112,000			

Source	DF	Seq SS
Vel e	1	49,170
Cons e	1	16,881
VxC	1	0,459
sqVxC	1	6,919

Durbin-Watson statistic = 1,24117  
 No evidence of lack of fit (P >= 0,1).

Tais resultados levam a duas possibilidades (1) o teste de Ferrar e Glauber não é eficaz. Gujarati (2000) chama a atenção para essa possibilidade ao mencionar três estudos<sup>1</sup> que lançam dúvidas sobre a capacidade do teste; (2) o teste de Ferrar e Glauber é eficiente e todos os demais indicadores obtidos estão viesados e as respectivas regras de bolso são falaciosas. Tal resultado seria pelo menos surpreendente e colocaria em cheque um bom número de contribuições científicas em diversas áreas do conhecimento que tenham utilizado esses indicadores.

O modelo final considerando as variáveis iniciais na forma normalizada é o da equação II que aparece a seguir:

$$y = \beta_0 + \beta_1 \exp x_3 + \beta_2 \exp x_5 + \beta_3 x_3 x_5 + \beta_4 (x_3 x_5)^2 \quad (\text{VIII})$$

O modelo obtido é de primeira ordem linear nos parâmetros. O modelo poderia ainda ser melhorado uma vez que há leve heteroscedasticidade remanescente.

---

## 6. Referências Bibliográficas

---

FERRAR D. E. & GLAUBER, R. R. *Multicollinearity in regression Analysis – the problem revisited*. The Review of Economics and Statistics, v. 49, n. 1, Feb. 1967.

FONSECA, M. A. R. *Álgebra linear aplicada a finanças, economia e econometria*. Barueri: Manole, 2003.

GUJARATI, D. N. *Econometria básica*. 3 ed. São Paulo: Makron Books, 2000.

KUTNER, M. H. et al. *Applied linear models*. 5th ed. New York: McGraw-Hill Irwin, 2005.

McCLAVE, J. T. & SINCICH, T. *Statistics*. Upper Saddle River: Prentice-Hall, 2000.

NEWBOLD, P. *Statistics for business and economics*. 4 ed. Upper Saddle River: Prentice-Hall, 1994.

TAMHANE, A. C. & Dunlop D. D. *Statistics and Data Analysis – from elementary to intermediate*. Upper Saddle River: Prentice-Hall, 2000.

---

<sup>1</sup> WICHERS, C. R. *The detection of multicollinearity: a comment*. Review of Economics and Statistics, v. 57, 1975.

KUMAR, T. K. *Multicollinearity in regression analysis*. Review of Economics and Statistics, v. 57, 1975.

O'HOGAN, J. & McCABE, *Tests for the severity of multicollinearity in regression Analysis: a comment*. Review of Economics and Statistics, v. 57, 1975.

Dados utilizados no exemplo.

n	Fabricante/Modelo	y	X1	X2	X3	X4	X5	X6	n	Fabricante/Modelo	y	X1	X2	X3	X4	X5	X6
1	Acura Integra GS-R	20015	7.0	15.6	133	185	25	0.84	64	Lexus LS400	57828	7.8	15.8	156	182	19	0.74
2	Acura Integra LS	18560	7.5	15.9	125	189	25	0.82	65	Lexus SC300	44225	6.8	15.4	138	178	18	0.84
3	Acura NSX-T	86642	5.2	13.8	162	173	18	0.95	66	Lincoln Continental	46152	7.6	15.7	123	178	17	0.78
4	AM General Hummer	71760	18.1	21.2	83	253	10	0.62	67	Lincoln Town Car Signature Serie	43627	9.0	16.9	106	186	17	0.79
5	Audi A6 Quattro	36802	8.3	16.5	127	191	18	0.79	68	Lingenfelter Firebird 383	34958	4.1	12.6	183	150	15	0.95
6	Audi Cabriolet	40960	9.7	17.2	126	189	18	0.73	69	Lotus Esprit S4S	87904	4.4	13.0	162	189	17	0.94
7	BMW M3	38782	5.3	14.0	138	165	19	0.88	70	Mazda Miata R Edition	19905	8.2	16.4	119	169	22	0.86
8	BMW 325i	36015	7.7	16.0	127	182	20	0.82	71	Mazda Millenia S	33595	7.7	16.0	144	189	20	0.81
9	BMW 540i Six-speed	54044	5.7	14.3	129	186	14	0.78	72	Mazda Protege ES	17870	8.9	17.0	120	195	26	0.76
10	BMW 740i	66837	8.4	16.6	127	181	16	0.79	73	Mazda RX-7	35252	5.3	14.0	157	156	17	0.99
11	BMW 840Ci	74790	7.1	15.5	155	170	15	0.83	74	Mazda 626ES	24015	8.4	16.8	118	189	20	0.80
12	BMW 850CSi	1E+05	5.3	13.9	158	167	12	0.85	75	Mercedes-Benz C36	51750	6.0	14.6	152	163	18	0.83
13	Buick Park Avenue Ultra	33797	7.0	15.5	107	216	17	0.74	76	Mercedes-Benz C220	31500	8.8	16.9	131	178	22	0.80
14	Buick Regal Gran Sport	24305	8.3	16.5	110	202	19	0.79	77	Mercedes-Benz E420	55130	7.3	15.7	127	185	18	0.73
15	Buick Riviera	30643	8.2	16.3	109	199	19	0.75	78	Mercedes-Benz S320	72037	8.1	16.4	132	178	17	0.72
16	Cadillac Eldorado Touring Coupe	45626	6.4	14.8	148	184	16	0.80	79	Mitsubishi Eclipse GS	18668	8.6	16.7	124	189	22	0.82
17	Cadillac SLS	46501	6.7	15.0	115	195	16	0.77	80	Mitsubishi Galant GS	21838	7.8	16.1	131	207	22	0.77
18	Chevrolet Blazer LT	26969	9.1	17.0	103	218	16	0.67	81	Mitsubishi Montero SR	39796	10.5	17.9	106	191	14	0.66
19	Chevrolet Camaro	20509	7.7	15.9	114	192	19	0.83	82	Mitsubishi 3000GT VR4	43190	5.1	13.9	153	164	18	0.92
20	Chevrolet Camaro Convertible	22484	9.3	17.0	110	195	19	0.84	83	Nissan Maxima SE	25586	7.3	15.9	123	187	21	0.75
21	Chevrolet Camaro Z28	22686	5.5	14.1	156	162	17	0.86	84	Nissan Quest GXE	24219	11.2	18.3	105	178	17	0.71
22	Chevrolet Cavalier	13417	8.8	16.7	113	209	25	0.76	85	Nissan Sentra GXE	13914	8.5	16.7	109	208	30	0.78
23	Chevrolet Corvette	42614	5.1	13.7	161	166	17	0.85	86	Nissan 200SX SE-R	17400	8.0	16.1	109	182	24	0.85
24	Chevrolet Impala SS	23355	6.5	15.0	142	179	17	0.86	87	Nissan 240SX SE	23000	7.7	16.0	118	170	22	0.89
25	Chevy Suburban Turbodiesel	35242	13.6	19.2	96	218	15	0.69	88	Nissan 300ZX Turbo	38806	5.6	14.2	155	175	18	0.89
26	Chevrolet Lumina LS	19900	8.0	16.3	119	205	17	0.79	89	Nissan 300ZX 2+2	37688	6.8	15.3	145	177	18	0.87
27	Chevrolet Monte Carlo LS	18530	8.7	16.7	113	197	19	0.80	90	Oldsmobile Aurora	32390	8.2	16.4	132	185	16	0.78
28	Chevrolet Tahoe LT	31877	11.3	18.2	99	198	12	0.68	91	Oldsmobile Eighty Eight LSS	25817	7.2	15.5	109	198	17	0.80
29	Chrysler Cirrus LXi	19600	9.0	17.0	124	202	20	0.76	92	Pontiac Bonneville SE	28662	7.0	15.5	127	186	17	0.81
30	Chrysler Sebring LX	17660	8.2	16.4	129	191	22	0.76	93	Pontiac Grand Am GT Coupe	17414	9.5	17.0	113	185	22	0.79
31	Chrysler Town & Country	26160	10.9	18.1	106	196	17	0.71	94	Pontiac Sunfire GT	15400	8.9	16.9	114	183	21	0.79
32	Dodge Avenger ES	20800	9.1	17.0	120	190	20	0.80	95	Pontiac Trans Am 25th Anniversary	28965	6.1	14.6	153	166	17	0.84
33	Dodge Intrepid	19837	9.5	17.2	111	195	20	0.74	96	Porsche Carrera 4	71973	4.9	13.6	161	163	16	0.90
34	Dodge Neon Sport Coupe	15155	7.6	16.0	115	207	29	0.82	97	Porsche 911 Carrera	66775	4.7	13.4	162	168	17	0.86
35	Dodge Stratus ES	16195	8.9	16.8	121	197	25	0.78	98	Porsche 911 Turbo	1E+05	3.7	12.3	176	162	13	0.95
36	Dodge Viper RT/10	61975	4.3	12.8	168	180	14	0.98	99	Range Rover 4.0 SE	54625	10.5	17.9	113	185	13	0.73
37	Eagle Talon TSi AWD	22000	6.6	15.3	137	169	21	0.85	100	Rolls-Royce Silver Spur III	2E+05	9.3	17.0	130	191	10	0.70
38	Ferrari F355	1E+05	4.5	13.0	179	165	10	1.02	101	Saab 900S	22430	7.6	16.1	126	188	20	0.75
39	Ford Contour SE	20140	8.8	16.9	112	183	22	0.82	102	Saab 900SE Convertible	40377	8.6	16.7	132	184	18	0.77
40	Ford Crown Victoria LX	25900	8.8	16.8	106	190	17	0.81	103	Saab 900SE Turbo	29460	6.5	15.3	142	171	21	0.82
41	Ford Explorer XLT	25625	10.7	17.9	106	199	15	0.67	104	Saab 900CDE	40157	8.3	16.3	132	178	18	0.75
42	Ford Probe GT	20815	7.4	15.9	131	164	21	0.86	105	Saleen Mustange S351	37319	5.1	13.9	147	171	16	0.87
43	Ford Windstar LX	26700	11.4	18.3	104	195	17	0.72	106	Saturn SC2	17340	7.6	16.0	122	174	25	0.84
44	Guldstrand GS90	1E+05	4.6	12.9	176	172	17	0.95	107	Subaru Legacy L	19584	10.4	17.9	108	176	22	0.75
45	Honda Accord LX V-6	22680	8.8	17.0	128	182	23	0.75	108	Subaru Legacy Outback Wagon	21123	10.6	17.9	108	166	22	0.82
46	Honda Civic DX	13210	8.7	16.7	110	198	35	0.72	109	Subaru SVX LS	29420	7.6	15.9	136	174	17	0.83
47	Honda Odyssey EX	25390	10.7	18.1	108	192	20	0.73	110	Toyota Avalon XL	25689	7.5	15.7	130	175	20	0.81
48	Honda Passport EX	25975	10.4	18.0	103	189	15	0.70	111	Toyota Camry DX	19978	10.7	18.1	116	196	21	0.75
49	Honda Prelude VTEC	24942	6.5	15.1	137	181	22	0.86	112	Toyota Camry SE	24984	7.9	16.2	132	177	18	0.78
50	Hyundai Elantra GLS	12965	10.0	17.3	113	193	21	0.81	113	Toyota Celica GT	25459	8.0	16.5	117	173	22	0.87
51	Hyundai Sonata GLS	19516	9.3	17.1	118	189	18	0.77	114	Toyota Celica GT Convertible	28120	9.0	16.9	115	175	22	0.85
52	Infiniti G20t	25325	8.3	16.5	125	176	24	0.83	115	Toyota Corolla DX	13498	9.3	17.1	110	197	27	0.77
53	Infiniti I30t	35200	7.7	16.1	127	189	21	0.78	116	Toyota Land Cruiser	41631	10.7	17.9	108	178	12	0.72
54	Infiniti Q45	57025	7.8	16.1	143	187	17	0.69	117	Toyota Supra	41642	6.5	15.0	141	173	18	0.92
55	Isuzu Trooper LS	32580	10.9	18.1	105	193	15	0.69	118	Toyota Tercel	12595	9.2	17.0	110	195	33	0.75
56	Jaguar XJR	71193	6.4	14.8	152	162	15	0.80	119	Toyota T100	22634	8.9	16.8	103	201	17	0.69
57	Jaguar XJS 4.0 Convertible	66210	8.3	16.4	144	194	17	0.81	120	Volkswagen Cabrio	21800	10.3	17.6	109	182	23	0.81
58	Jaguar XJ6	61788	8.1	16.3	139	189	17	0.74	121	Volkswagen GTI VR6	19265	6.6	15.2	130	179	18	0.78
59	Jeep Grand Cherokee Limited	31332	8.0	16.3	114	180	14	0.75	122	Volkswagen Jetta III GLX	20365	6.9	15.5	136	208	18	0.78
60	Kia Sephia GS	12065	10.2	17.6	109	197	27	0.75	123	Volkswagen Passat GLX	22930	8.6	16.5	131	178	18	0.81
61	Lamborghini Diablo VT	3E+05	5.9	14.0	185	164	9	0.89	124	Volvo 850 Turbo	33266	6.1	14.6	150	172	19	0.81
62	Land Rover Discovery	35192	10.9	18.3	105	204	13	0.67	125	Volvo 850 T-5R	38056	6.7	15.2	149	181	19	0.79
63	Lexus LS400	57828	7.8	15.8	156	182	19	0.74	126	Volvo 960	30360	8.6	16.6	129	190	17	0.80

y = Preço; x1 = Aceleração em segundos de 0-60 milhas/hora; x2 = Aceleração em segundos para percorrer ¼ de milha; x3 = velocidade máxima em milhas por hora; x4 = Distância de frenagem em pés a 70 milhas/hora; x5 = Consumo em milhas por galão; x6 = Aderência em g's durante a curva.

Adaptado de McCLAVE, J. T. & SINCICH, T. *Statistics*. Upper Saddle River: Prentice-Hall, 2000.

# DETERMINAÇÃO COMPUTACIONAL DE ÁREAS E VOLUMES DE TUMORES MEDIANTE IMAGENS

Camila Becker<sup>1</sup>, Dr. Ruben Edgardo Panta Pazos<sup>2</sup>.

<sup>1</sup> Universidade de Santa Cruz do Sul - UNISC, camilabecker@mx2.unisc.br;

<sup>2</sup> Universidade de Santa Cruz do Sul - UNISC, rpazos@unisc.br.

## Resumo:

No presente trabalho se estuda o problema de calcular a área de um tumor presente num pulmão humano, a partir de uma tomografia. Para isso emprega-se o processamento por programas de edição gráfica da imagem. Visando gerar uma região de tipo poligonal, escolhem-se pontos na fronteira da região a ser considerada e, sobre esses pontos, aplica-se interpolação utilizando posteriormente a fórmula de Green e o método de Monte Carlo para o cálculo de áreas.

**Palavras-chave:** Cálculo de áreas, Método de Monte Carlo, Teorema de Green, Método dos Trapézios.

## Apresentação:

### 1. Introdução

Na determinação de áreas e volumes são usados diversos tipos de imagens (como tomografias, radiografias, imagens de ressonância magnética, entre outros) e geralmente são obtidas mediante métodos de aproximação numérica, particularmente mediante triangulações de regiões poligonais inscritas que aproximam a região.

Neste trabalho, aplicam-se alguns métodos para determinar áreas derivadas de observações experimentais de uma tomografia de um pulmão com tumor.



Figura 1: Imagem utilizada para o cálculo da área de um tumor.

A metodologia a ser utilizada consiste na seguinte seqüência:

- Captar informação gráfica, mediante fotografias, mapas ou outros objetos de igual nível, que podem ser processados por programas de edição gráfica.
- Num programa de edição gráfica, escolher pontos na fronteira da região a ser considerada, visando gerar uma região de tipo poligonal.
- Aplicar uma interpolação por *splines* cúbicos sobre os pontos da região poligonal utilizando posteriormente a fórmula de Green de cálculo de áreas.
- Calcular a área utilizando, também, o método de Monte Carlo.

Tudo isto permite com suficiente resolução e velocidade colocar resultados que antes eram de grande complexidade.

Este trabalho está organizado da seguinte maneira: Inicialmente imprime-se uma visão dos métodos que são aplicados na resolução do problema proposto no trabalho. Posteriormente, são apresentados os resultados obtidos com um sistema de computação algébrica, Maple. Finalmente, se expressa as conclusões e possíveis extensões da pesquisa.

## 2. Métodos Utilizados

### 2.1 *Splines* Cúbicos

As funções *Splines* Cúbicas são bastante utilizadas por serem funções suaves para o ajuste de dados e também por não produzirem comportamentos com oscilações na interpolação.

*Spline* Cúbica é uma função ( $S$ ) definida num intervalo  $[a, b]$ , e possui as seguintes propriedades (Sperandio (2003)):

1.  $S(x_i) = y_i, \quad 0 \leq i \leq n.$
2.  $S(x), S''(x), S'''(x)$  são funções contínuas em  $[a, b].$
3. Em cada subintervalo  $[x_{i-1}, x_i], \quad i = 1, 2, \dots, n, S(x)$  é um polinômio cúbico, isto é,  

$$S(x) = a_i + b_i x + c_i x^2 + d_i x^3, \quad x_{i-1} \leq x \leq x_i.$$
4.  $S^{(j)}(x_i^+) = S^{(j)}(x_i^-), \quad i = 1, 2, 3, \dots, n-1, j = 0, 1, 2,$  tendo em vista a questão de continuidade.

Por exemplo: Dados os pontos  $(1, 1), (4, 3), (6, -1), (8, 1)$ , as funções de interpolação mediante *splines* cúbicos serão:

$$\begin{cases} -\frac{107}{171} + \frac{79}{57}x + \frac{41}{114}x^2 - \frac{41}{342}x^3 & x < 4 \\ -\frac{659}{19} + \frac{1535}{57}x - \frac{229}{38}x^2 + \frac{47}{114}x^3 & x < 6 \\ \frac{1987}{19} - \frac{2434}{57}x + \frac{106}{19}x^2 - \frac{53}{228}x^3 & \text{no caso contrário} \end{cases}$$

Abaixo, representam-se graficamente os pontos e a função de interpolação mediante *splines* cúbicos:

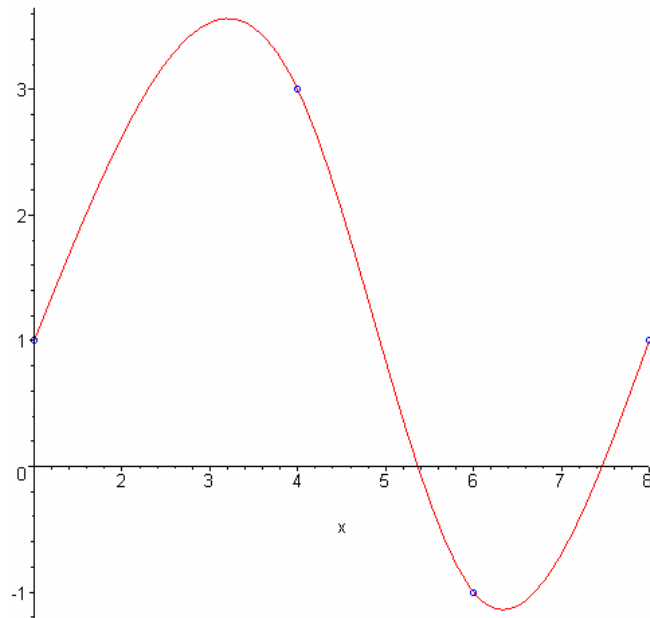


Figura 2: Representação gráfica da interpolação mediante *splines* cúbicos.

## 2.2 Teorema de Green

O Teorema de Green foi desenvolvido pelo matemático e físico inglês, George Green. O Teorema fornece a relação entre uma integral de linha ao redor de uma curva fechada orientada positivamente e uma integral dupla sobre a região do plano cercada pela curva.

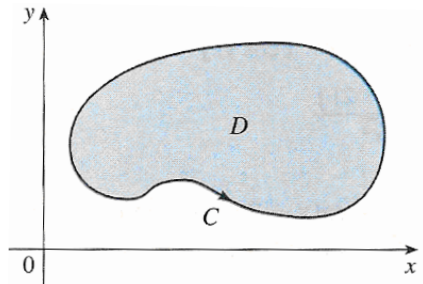


Figura 3: Representação de uma curva fechada.

Seja  $C$  uma curva plana simples, fechada, contínua por trechos, orientada positivamente, e seja  $D$  a região delimitada por  $C$ . Se  $P$  e  $Q$  têm derivadas parciais de primeira ordem contínuas sobre uma região aberta que contenha  $D$ , então:

$$\oint_C (P dx + Q dy) = \iint_D \left( \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) dA$$

Assim, a integral de linha em uma curva de um campo  $P$  é igual a integral dupla, na região cercada pela curva, da diferença das derivadas parciais das componentes do campo,  $P$  e  $Q$ , em relação a  $y$  e  $x$  respectivamente.

Desta forma se a diferença de derivadas parciais do membro da direita for igual a uma constante, podemos retirá-la da integral e teremos a seguinte situação:

$$\oint_C (P dx + Q dy) = K * \iint_D dx dy = K * \text{área cercada por } C$$

Podemos observar que o teorema de Green pode ser usado para encontrarmos a área de uma região plana limitada por uma curva  $C$ .

## 2.3 Método Monte Carlo

O Método de Monte Carlo é uma técnica que vem sendo utilizada a séculos. Entretanto, sua importância aumentou nas últimas décadas. Durante a Segunda Guerra Mundial foi utilizada em pesquisas relacionadas com a difusão aleatória de nêutrons. O seu

nome é inspirado em jogos de azar, uma vez que a simulação estatística possui similaridade com jogos de azar, e Monte Carlo é uma cidade que possui muitos cassinos.

Esse Método é amplamente utilizado em diversas áreas, como:

- Simulação de sistemas físicos complexos e de jogos eletrônicos;
- Estudo de reações nucleares;
- Radioterapia;
- Fluxo de tráfego;
- Economia.

De maneira ampla, podemos definir o Método de Monte Carlo com sendo a simulação para obter aproximações de variáveis de interesse, objetivando reproduzir o que ocorre no mundo real.

Neste método, quanto maior for o número de pontos gerados aleatoriamente mais o valor obtido se aproximará do valor real.

Um exemplo simples de aplicação do Método de Monte Carlo: Dada a função  $f(x) = \text{sen}(x^2)$ , definindo os valores para  $x$  entre 0 e  $\sqrt{\pi}$ , teremos o seguinte gráfico:

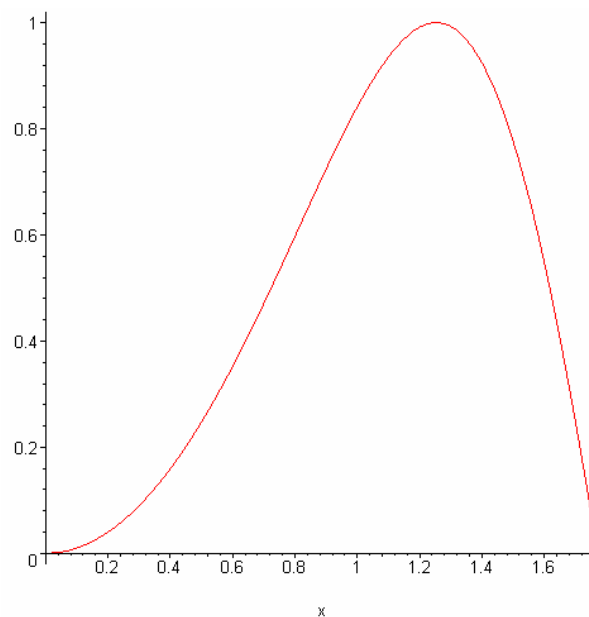


Figura 4: Gráfico para a função  $f(x) = \text{sen}(x^2)$ .

Calculando a área da região limitada superiormente pelo gráfico da função e inferiormente pelo eixo das abscissas mediante o método de Monte Carlo, através da



geração de pontos aleatoriamente para a função, termos o seguinte resultado aproximadamente para a área 0.8950891948 unidades de área ao quadrado.

Um outro exemplo de aplicação do método de Monte Carlo é o programa *Peregrine*, que utiliza o método de Monte Carlo para a terapia de radiação em tumores. Este método permite a solução estocástica para as equações de transporte complexas da radiação 3D que são difíceis de resolver analiticamente. Dentre as vantagens da utilização do método de Monte Carlo, podemos ainda citar que permite cálculos exatos da dose para todos os materiais, todas as geometria (apresentadas pelos diferentes órgãos, por exemplo), além de apresentar informações detalhadas sobre como a dose é depositada na matéria.

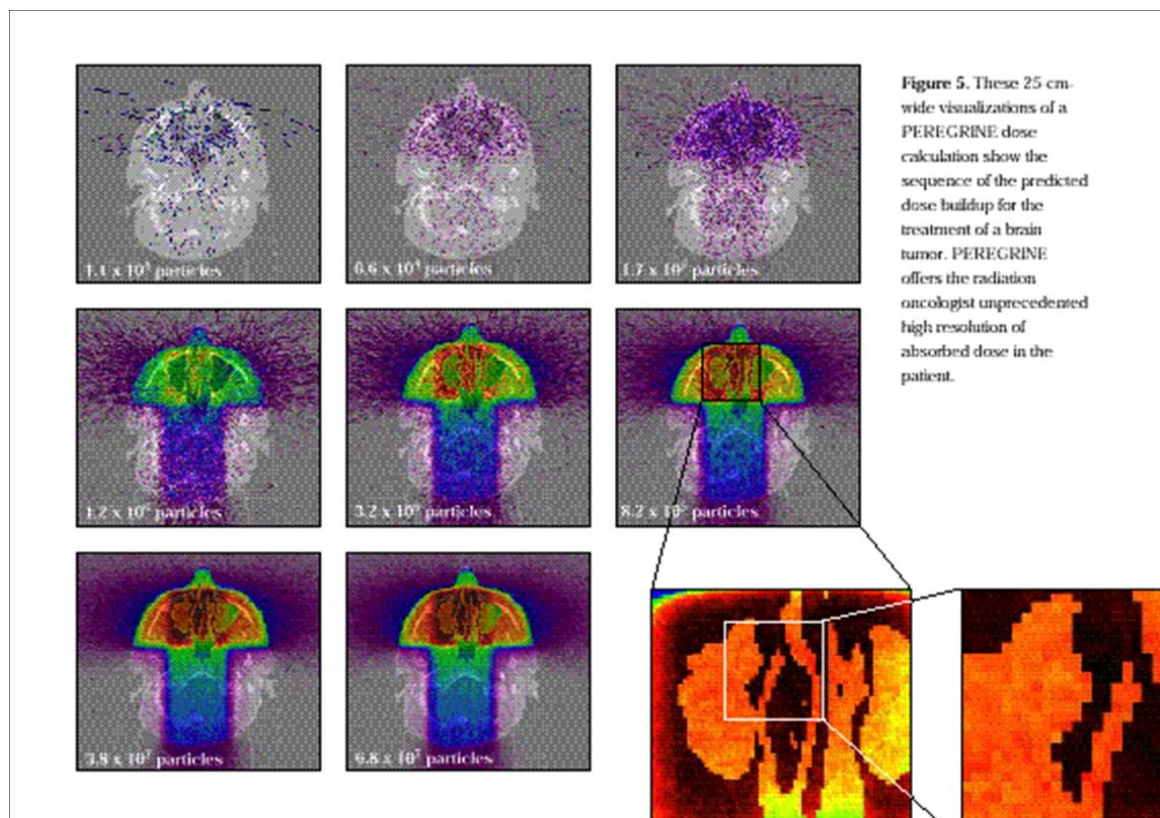


Figura 5: Imagem do funcionamento do programa *Peregrine*.

## 2.4 Método dos Trapézios

O método do trapézio consiste em dividir a área a ser calculada em vários trapézios. A soma da área de cada trapézio aproxima-se do resultado da integral. Logo, quanto maior o número de trapézios, mais próximo do valor real da integral será o resultado.

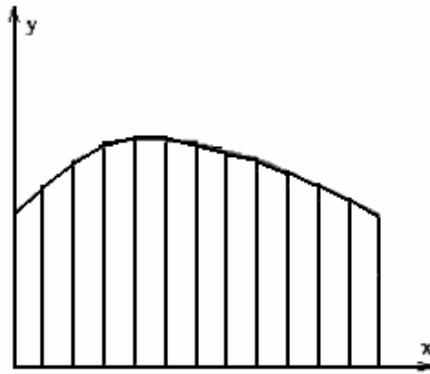


Figura 6: Representação do método de trapézios.

Algebricamente, temos que:

$$\int_a^b f(x)dx = \frac{b-a}{2n} \left[ y_0 + 2 \left[ \sum_{k=1}^{n-1} y_k \right] + y_n \right]$$

Onde  $a$  e  $b$  representam os extremos de valores para  $x$ . Logo,  $b-a$  representa a distância entre os dois pontos.

### 3. Resultados Numéricos

Inicialmente, foi captada a seguinte imagem de uma tomografia computadorizada de um pulmão com a presença de um tumor.

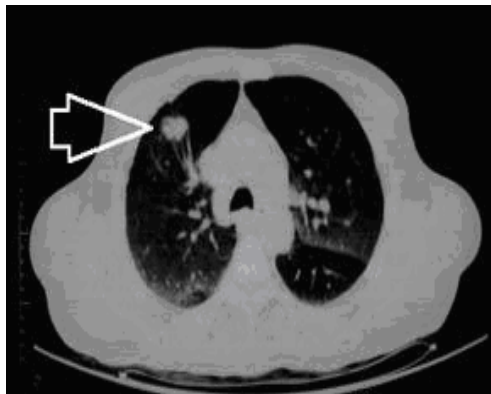


Figura 7: Imagem utilizada para a realização do trabalho.

A partir disso, a imagem foi processada por um programa de edição gráfica, sendo escolhidos pontos na fronteira da região a ser considerada. Com os dados obtidos, representou-se a imagem em um sistema de computação algébrica, Maple.

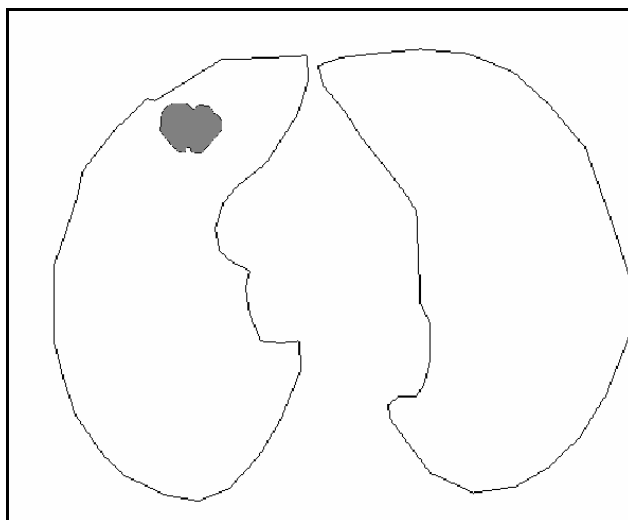


Figura 8: Representação da imagem em um sistema de computação algébrica.

Com os pontos (dados) obtidos foi possível aplicar a interpolação mediante *splines* cúbicos e, a partir das funções encontradas, empregaram-se os demais métodos. Inicialmente foi aplicado o Teorema de Green, e os resultados obtidos foram, aproximadamente,  $2607.858 \text{ pixels}^2$ .

Aplicou-se, também, o Método de Monte Carlo, e os resultados encontrados foram praticamente os mesmos do que os obtidos mediante o Teorema de Green.

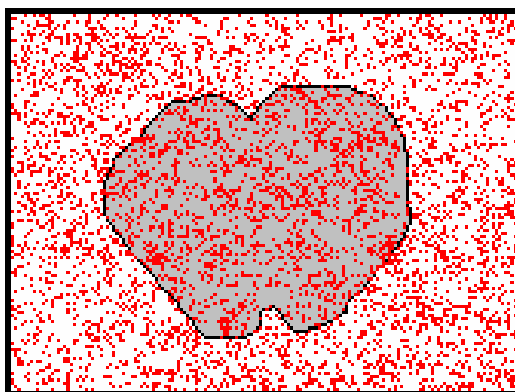


Figura 9: Representação da aplicação do Método de Monte Carlo.

Com a aplicação do método de Monte Carlo, foram encontrados os seguintes resultados:

	2564.662500
	2615.955750
	2513.369250
	2735.640000
	2650.151250
$M :=$	2633.053500
	2667.249000
	2479.173750
	2444.978250
	2598.858000
	2684.346750

Tabela 1: Resultados encontrados com a aplicação do Método de Monte Carlo.

Devido ao método ser de simulação para obter aproximações, encontrou-se essa tabela de resultados para a área do tumor em pixels quadrado. Então, fazendo a média aritmética entre os valores encontrados, temos, de maneira aproximada, que a área correspondente é  $2598.858 \text{ pixels}^2$ .

Para a aplicação do método dos Trapézios, deve-se calcular a soma das áreas dos trapézios formados pelos pontos superiores e subtrair das encontradas com os pontos inferiores. Através da utilização deste método, encontrou-se um valor próximo aos encontrados com os demais métodos:  $2657.848750 \text{ pixels}^2$ . Esta mínima diferença deve-se a particularidades dos métodos.

Para empregar os diferentes métodos para o cálculo da área do tumor foi utilizado o auxílio do sistema de computação algébrica Maple.

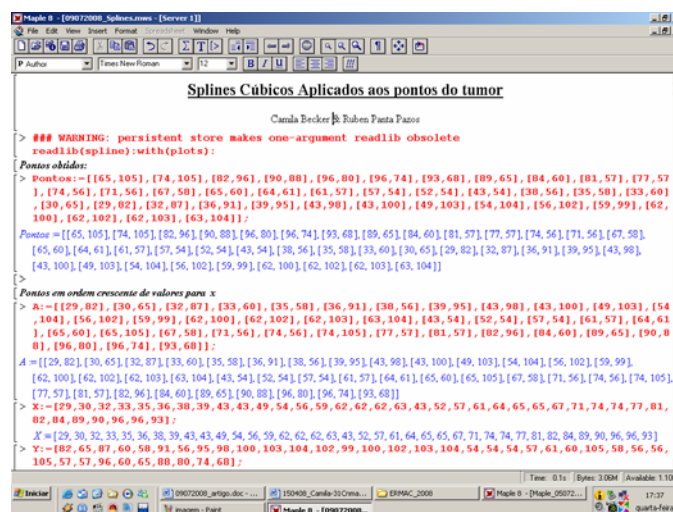


Figura 10: Folha de trabalho da resolução da interpolação mediante *Splines* cúbicos.

#### 4. Conclusões

A combinação entre métodos diversos, visando a determinação de área, permite a geração de um método híbrido para diferentes tipos de regiões: com fronteira poligonal, com fronteira de curvas suaves, ou ainda com fronteira de tipo irregular (Lifschitziano).

O uso de sistemas de computação algébrica da maior eficácia ao método, colocando-o ao alcance de setores acadêmicos mais amplos.

Através da utilização do Teorema de Green e do Método de Monte Carlo, obtiveram-se os mesmos resultados para o cálculo da área do tumor analisado.

Utilizando os métodos propostos e, uma seqüência de tomografias de um mesmo paciente, é possível verificar a situação do tamanho do tumor (aumento, estagnação ou diminuição).

#### Agradecimentos

Os autores agradecem à Universidade de Santa Cruz do Sul (UNISC) pelo apoio financeiro. Além disso, a primeira autora agradece de forma especial a CAPES pelo destacável apoio.

#### Referências Bibliográficas:

ANTON, H.: **Cálculo, um novo horizonte**, Porto Alegre, RS, vol 2, Editora Bookman, (2002).

CUNHA, C.: **Métodos Numéricos**, Editora Unicamp, Campinas, SP, (2000).

DURHAM, A.: **Método Monte Carlo**. Disponível em: <http://www.ime.usp.br/>.

RABELO, A. B. e MANSO, F. F.: **O Planímetro e o Teorema de Green**. Disponível em: <http://www.mat.ufmg.br/comed/2004/e2004/planimetro.pdf> .

SIANTAR, C. H. **PEREGRINE Program**. Disponível em: <http://www.nuc.berkeley.edu/courses/classes/NE39/Vujic-cancer.ppt>

SPERANDIO, D., MENDES, J.T., e SILVA, L.H.M: **Cálculo Numérico: características matemáticas e computacionais dos métodos numéricos**. São Paulo: Prentice Hall, 2003.

**PEREGRINE™ Takes Aim at Cancer Tumors**. Disponível em: [https://www.llnl.gov/str/pdfs/10\\_99.6.pdf](https://www.llnl.gov/str/pdfs/10_99.6.pdf) .

# AVALIAÇÃO DA QUALIDADE DE UM SERVIÇO DE TRANSMISSÃO DE VOZ SOBRE IP USANDO UM SISTEMA DE REGRAS NEBULOSAS

Alex Sandro de Castilho, Marcos Eduardo Valle.

Universidade Estadual de Londrina - UEL, alexcastilho22@gmail.com; Universidade Estadual de Londrina - UEL, valle@uel.br

**Resumo:** Tecnologias que usam a internet para a transmissão de voz, conhecidas como VoIP, oferecem uma alternativa atraente e econômica para os sistemas de telefonia convencionais. Nesse trabalho, desenvolvemos um modelo baseado em sistemas de regras nebulosas para avaliar a qualidade de um serviço VoIP. Especificamente, nosso modelo fornece uma nota para a qualidade do serviço considerando apenas a taxa de pacotes perdidos e o atraso durante a transmissão. Essa nota pode ser usada para o monitoramento e/ou controle do serviço de modo que alguns requisitos referentes a qualidade sejam satisfeitos.

**Palavras-chave:** Qualidade de serviço, transmissão de voz sobre IP, conjuntos nebulosos, sistemas de regras nebulosas.

## Introdução

Atualmente, os serviços de telecomunicações não estão mais restritos apenas aos sistemas de telefonia convencionais. Hoje, temos também a opção de usar a internet como meio de comunicação em tempo real. Visto que a internet emprega um protocolo especial chamado IP (do inglês *internet protocol*), denominamos VoIP (do inglês *Voice over IP*) qualquer tecnologia que usa a internet para a comunicação de voz [ANATEL].

É importante lembrar, entretanto, que a internet foi projetada para o tráfego de dados genéricos, e não necessariamente para a transmissão de voz em tempo real. Conse-

quentemente, a qualidade dos serviços VoIP pode ser inferior aos serviços convencionais de telefonia, que foram projetados especificamente para a transmissão de voz [Manousos et al., 2005]. Nesse contexto, surge a necessidade de medir a qualidade de um serviço VoIP [Sun and Ifeachor, 2006]. Essa medida pode ser usada para o monitoramento e/ou controle do serviço de modo que alguns requisitos sejam satisfeitos. Nesse artigo usaremos um sistema de regras nebulosas para estimar a *qualidade de um serviço* (QoS, quality of service) VoIP.

Muitos fatores podem influenciar a qualidade de um serviço VoIP, incluindo a *taxa de pacotes perdidos* (ou simplesmente a *perda de pacotes*),  *jitter de atraso* (ou simplesmente o *atraso*) e o *codec* empregado para a codificação e decodificação do sinal de voz. Nesse trabalho, iremos estimar a qualidade do serviço VoIP considerando apenas a perda de pacotes e o atraso. O codec será o mesmo durante todos os experimentos.

Existe na literatura basicamente dois tipos de métodos para avaliar a qualidade de um serviço VoIP: *métodos objetivos* e *métodos subjetivos* [Ferrari, 2007, Sun and Ifeachor, 2006, Mohamed et al., 2001]. Os métodos objetivos avaliam a qualidade do serviço VoIP utilizando dados da rede e algumas informações do sistema como o codec empregado. O *modelo-E* é um exemplo de método objetivo [ITU, 2005]. Em contraste, os métodos subjetivos simulam a opinião dos usuários sobre a qualidade do serviço VoIP [ITU, 1996]. É importante observar que a opinião dos usuários pode variar dependendo da língua empregada e dos serviços de telefonia que os usuários estão acostumados a utilizar [Mohamed et al., 2001]. O método que utilizaremos nesse trabalho para analisar a qualidade do sistema VoIP pode ser classificado como um método subjetivo pois foi projetado a partir da nossa concepção de como deve ser a qualidade do serviço VoIP conhecidos a perda e o atraso da rede.

Finalmente, gostaríamos de lembrar que a estimativa da QoS deve ser efetuada em tempo real na ordem de milissegundos. Portanto, nesse trabalho buscamos um algoritmo que efetue poucas operações de ponto flutuante (flops) [Burden and Faires, 2004].

Optamos assim por um método que efetua, na grande maioria, operações de máximos e mínimos.

### Conceitos Básicos da Teoria dos Conjuntos Nebulosos

A *Teoria dos Conjuntos Nebulosos* (ou Conjuntos Fuzzy) foi introduzida em 1965 pelo matemático Loft Zadeh no intuito de representar matematicamente conceitos vagos empregados frequentemente na linguagem cotidiana [Zadeh, 1965, Zadeh, 1979]. Conjuntos nebulosos foram usados com sucesso para resolver diversos problemas com aplicações científicas e/ou na área de engenharia [Barros e Bassanezi, 2006, Kosko, 1992, Lee, 2005, Pedrycz and Gomide, 2007]. Sobretudo, a teoria dos conjuntos nebulosos oferece uma ferramenta eficiente para a manipulação de termos expressos em palavras [Zadeh, 1996]. Nesse trabalho, usaremos conjuntos nebulosos para modelar e manipular conceitos vagos como a noção de “serviço ruim”, “perda imperceptível” e “atraso inaceitável”.

Um conjunto nebuloso, definido sobre um universo de discurso  $\mathbf{X}$ , é caracterizado por uma função de pertinência que atribui um valor no intervalo  $[0, 1]$  para cada elemento  $x \in \mathbf{X}$ . Em outras palavras, um conjunto nebuloso é descrito por uma função  $A : \mathbf{X} \rightarrow [0, 1]$ . O valor  $A(x)$  representa o grau de compatibilidade de  $x \in \mathbf{X}$  com o conceito representado por  $A$ , i.e.,  $A(x)$  representa o quanto  $x$  é  $A$ . As Figuras 1, 2 e 3 apresentam exemplos de conjuntos nebulosos definidos sobre o intervalo unitário, i.e., conjuntos nebulosos onde o universo de discurso é  $\mathbf{X} = [0, 1]$ . Especificamente, essas figuras apresentam os conjuntos nebulosos que usamos para avaliar a qualidade de um serviço VoIP.

A Figura 1 apresenta os três conjuntos nebulosos que foram usados para descrever a taxa de pacotes perdidos, ou simplesmente a perda de pacotes. Esses conjuntos nebulosos representam os conceitos de *perda imperceptível*, *perda aceitável* e *perda inaceitável* de pacotes e foram denotados por  $A^1$ ,  $A^2$  e  $A^3$ , respectivamente. Assim, uma perda de pacotes correspondente à 0.6 possui compatibilidade 0 com o conceito de perda imperceptível,



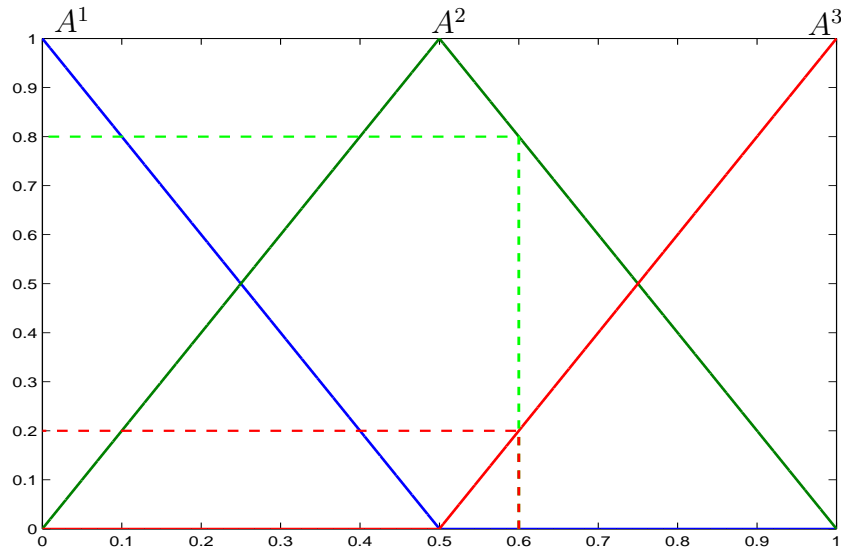


Figura 1: Conjuntos nebulosos usados para modelar a taxa de pacotes perdidos.

0.8 com o conceito de perda aceitável e 0.2 com os conceito de perda inaceitável. De um modo similar, a Figura 2 apresenta conjuntos nebulosos  $B^1$ ,  $B^2$  e  $B^3$  que representam os conceitos de *atraso imperceptível*, *aceitável* e *inaceitável*, respectivamente. Desse modo, um atraso de 0.2 é 0.6 imperceptível, 0.4 aceitável e nada compatível com o conceito inaceitável. Finalmente, a Figura 3 apresenta os conjuntos nebulosos que descrevem a qualidade do serviço VoIP. Especificamente, essa figura apresenta os conjuntos nebulosos  $C^1, C^2, \dots, C^5$  que representam os conceitos de serviço *excelente*, *bom*, *regular*, *ruim* e *péssimo*, respectivamente.

### Operações com Conjuntos Nebulosos

Muitas operações com conjuntos nebulosos são efetuadas usando o conceito de conjunção nebulosa e disjunção nebulosa. Conjunções e disjunções nebulosas generalizam os conceitos clássicos de conjunção e disjunção, i.e., os operadores “e” e “ou” da lógica booleana [Klir and Yuan, 1995, Pedrycz and Gomide, 2007, Zadeh, 1965].

Uma *conjunção nebulosa* é uma operação binária  $C_{\mathcal{F}} : [0, 1] \times [0, 1] \rightarrow [0, 1]$  cres-

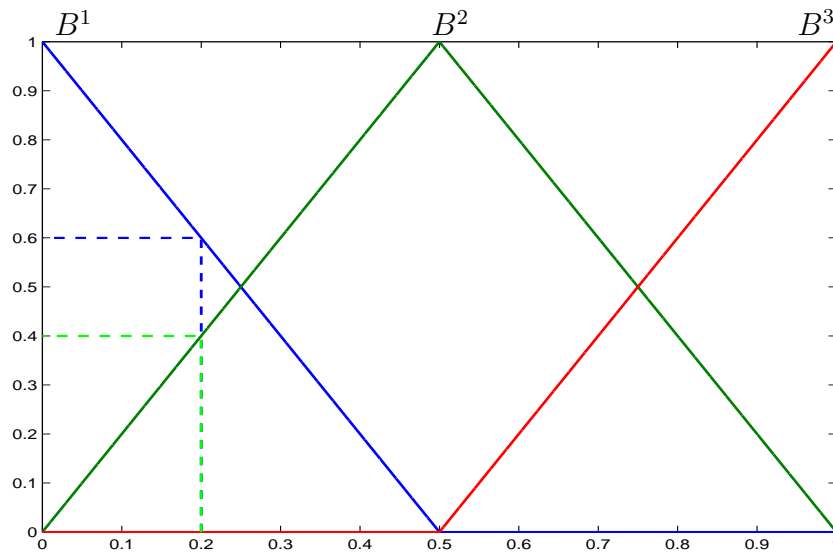


Figura 2: Conjuntos nebulosos usados para modelar o *jitter* de atraso da rede.

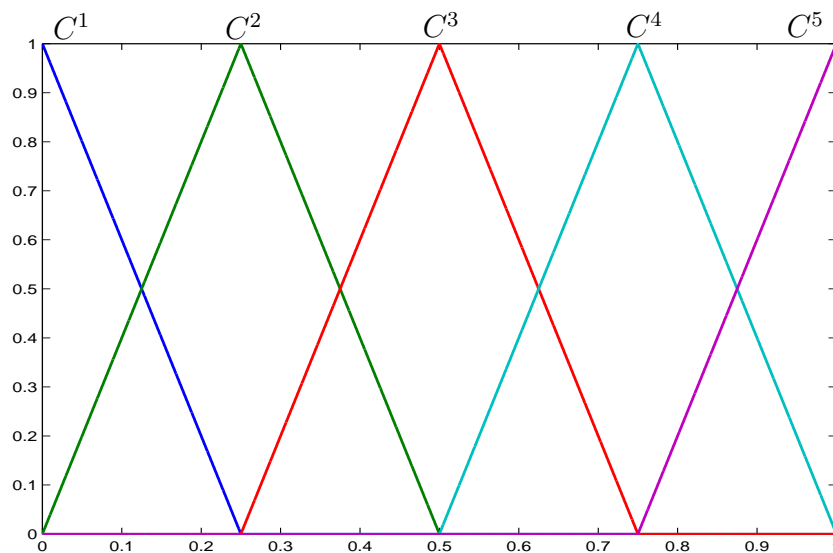


Figura 3: Conjuntos nebulosos usados para modelar a qualidade do serviço VoIP.

cente em ambos os argumentos que satisfaz as equações  $C_{\mathcal{F}}(0, 0) = C_{\mathcal{F}}(1, 0) = C_{\mathcal{F}}(0, 1) = 0$  e  $C_{\mathcal{F}}(1, 1) = 1$  [Valle, 2007, Valle and Sussner, 2008]. O mínimo é um exemplo de conjunção nebulosa comutativa e associativa. Nesse trabalho denotaremos o mínimo usando o símbolo “ $\wedge$ ”. Lembre-se que esse símbolo é usado para representar o conectivo “e” na lógica clássica [Barros e Bassanezi, 2006, Nguyen and Walker, 2000]. Em particular, o mínimo entre vários elementos  $a_1, \dots, a_n$  será denotado por  $\bigwedge_{i=1}^n a_i$ .

De um modo similar, uma *disjunção nebulosa* é uma operação binária  $D_{\mathcal{F}} : [0, 1] \times [0, 1] \rightarrow [0, 1]$  crescente em ambos os argumentos que satisfaz as condições  $D_{\mathcal{F}}(0, 0) = 0$  e  $D_{\mathcal{F}}(1, 0) = D_{\mathcal{F}}(0, 1) = D_{\mathcal{F}}(1, 1) = 1$ . A operação de máximo, denotada nesse artigo através do símbolo “ $\vee$ ”, é um exemplo de disjunção nebulosa.

Nesse artigo empregaremos apenas o mínimo e o máximo como conjunção e disjunção nebulosa. Antes de prosseguir, entretando, gostaríamos de apresentar algumas observações sobre essas duas operações.

Do ponto de vista matemático, a classe de todos os conjuntos nebulosos com as operações de máximo e mínimo, aplicadas pontualmente, representa um *reticulado completo* [Birkhoff, 1993]. Nesse trabalho não discutiremos as propriedades dessas estruturas algébricas. Todavia, gostaríamos de mencionar que reticulados completos formam a base para muitas técnicas de processamento e análise de imagens e inteligência computacional [Banon et al, 2007, Heijmans, 1994, Kaburlasos, 2006, Ritter and Wilson, 2001, Kaburlasos and Ritter, 2007]. Do ponto de vista computacional, as operações de máximo e mínimo não geram erros de arredondamento e podem ser efetuadas de forma rápida e eficiente [Ritter and Wilson, 2001].

## Sistemas de Regras Nebulosas

Nesse trabalho utilizaremos sistemas de regras nebulosas para estimar a qualidade de um serviço VoIP. Uma regra nebulosa é qualquer sentença da forma *se-então* onde os antecedentes e/ou conseqüentes são conjuntos nebulosos [Barros e Bassanezi, 2006, Klir

SRNs	$B^1$	$B^2$	$B^3$
$A^1$	$C^5$	$C^4$	$C^3$
$A^2$	$C^4$	$C^3$	$C^2$
$A^3$	$C^3$	$C^2$	$C^1$

Tabela 1: Sistema de regras usado para avaliar a qualidade do serviço VoIP.

and Yuan, 1995, Nguyen and Walker, 2000, Pedrycz and Gomide, 2007]. A seguinte sentença é um exemplo de regra nebulosa se considerarmos os conjuntos nebulosos definidos nas Figuras 1, 2 e 3:

Se a *perda é imperceptível* e o *atraso é imperceptível*, então o *serviço é ótimo*. (1)

Várias regras nebulosas formam um *Sistema de Regras Nebulosas* (SRNs). Note que SRNs podem ser construídos a partir da nossa concepção de como o sistema deve funcionar [Klir and Yuan, 1995, Pedrycz and Gomide, 2007]. No nosso caso, elaboramos um SRNs para avaliar a qualidade do serviço VoIP usando nossa percepção sobre a taxa de pacotes perdidos e o *jitter* de atraso da rede.

É importante observar que um SRNs com dois antecedentes e um conseqüente pode ser expresso de forma compacta usando uma tabela onde os conjuntos nebulosos do antecedente formam as linhas e colunas e os elementos no interior da tabela representam os conseqüentes. Por exemplo, nesse trabalho consideramos o SRNs apresentado na Tabela 1. Aqui, as linhas e as colunas correspondem aos conjuntos nebulosos definidos nas Figuras 1 e 2, respectivamente. Os conjuntos nebulosos no interior da tabela correspondem aos conjuntos nebulosos apresentados na Figura 3. Note que a interação entre a primeira linha e a primeira coluna, i.e., a interação de  $A^1$  e  $B^1$ , resulta no conjunto nebuloso  $C^5$ . Em outras palavras, temos a regra apresentada em (1).

Matematicamente, podemos representar um SRNs como segue onde  $A^\xi$ ,  $B^\xi$  e  $C^\xi$

são conjuntos nebulosos em  $\mathbf{X}$ ,  $\mathbf{Y}$  e  $\mathbf{Z}$ , respectivamente:

$$\text{Se } x \text{ é } A^\xi \text{ e } y \text{ é } B^\xi, \text{ então } z \text{ é } C^\xi, \quad \text{para } \xi = 1, \dots, k. \quad (2)$$

Dados elementos  $x \in \mathbf{X}$  e  $y \in \mathbf{Y}$ , podemos deduzir (ou inferir) um conjunto nebuloso  $C$  em  $\mathbf{Z}$  como segue.

Primeiro, note que os valores  $A^\xi(x)$  e  $B^\xi(y)$  representam o quanto  $x$  é  $A^\xi$  e o quanto  $y$  é  $B^\xi$ , para  $\xi = 1, \dots, k$ . Desse modo, o valor  $\lambda^\xi = A^\xi(x) \wedge B^\xi(y)$  avalia a veracidade da sentença “ $x$  é  $A^\xi$  e  $y$  é  $B^\xi$ ”. Se  $\lambda^\xi = 0$ , então a regra  $\xi$  não deve contribuir para a dedução do conjunto nebuloso  $C$ . Por outro lado, se  $\lambda^\xi = 1$ , então o conjunto nebuloso  $C$  deduzido pelo SRNs deve conter o conseqüente da regra  $\xi$ , i.e.,  $C^\xi$  deve estar contido em  $C$ . Valores  $\lambda^\xi$  entre 0 e 1 devem contribuir parcialmente para o cálculo de  $C$ . Essas observações podem ser capturadas limitando os valores do conseqüente  $C^\xi$  à  $\lambda^\xi$ , para todo  $\xi = 1, \dots, k$ . Em outras palavras, definimos  $\lambda^\xi \wedge C^\xi(z)$  como sendo a contribuição da regra  $\xi$  para o conjunto nebuloso  $C$  avaliado em  $z \in \mathbf{Z}$ . Finalmente, como o processo de dedução deve levar em consideração todas as regras, definimos o conjunto nebuloso  $C$  como segue para todo  $z \in \mathbf{Z}$ :

$$C(z) = \bigvee_{\xi=1}^k (\lambda^\xi \wedge C^\xi(z)) = \bigvee_{\xi=1}^k (A^\xi(x) \wedge B^\xi(y) \wedge C^\xi(z)) . \quad (3)$$

Por exemplo, considere um valor de perda igual a 0,6 e um atraso de 0.2. Introduzindo esses dados no SRNs apresentado na Tabela 1 com os conjuntos nebulosos das Figuras 1, 2 e 3, obtemos o conjunto nebuloso  $C$  representado na Figura 4.

O processo descrito acima é chamado *método de inferência de Mamdani* [Klir and Yuan, 1995, Nguyen and Walker, 2000, Pedrycz and Gomide, 2007]. Outros métodos de inferência incluem os métodos de Larsen, Kosko e Yager [Kosko, 1992, Pedrycz and Gomide, 2007, Yager and Filev, 1994].

Finalmente, observe que o método de inferência de Mamdani apresenta como saída um conjunto nebuloso  $C$  em  $\mathbf{Z}$ . Podemos utilizar um *método de defuzzificação* para determinar um valor  $z^* \in Z$  que representa o conjunto nebuloso  $C$  [Pedrycz and Gomide,

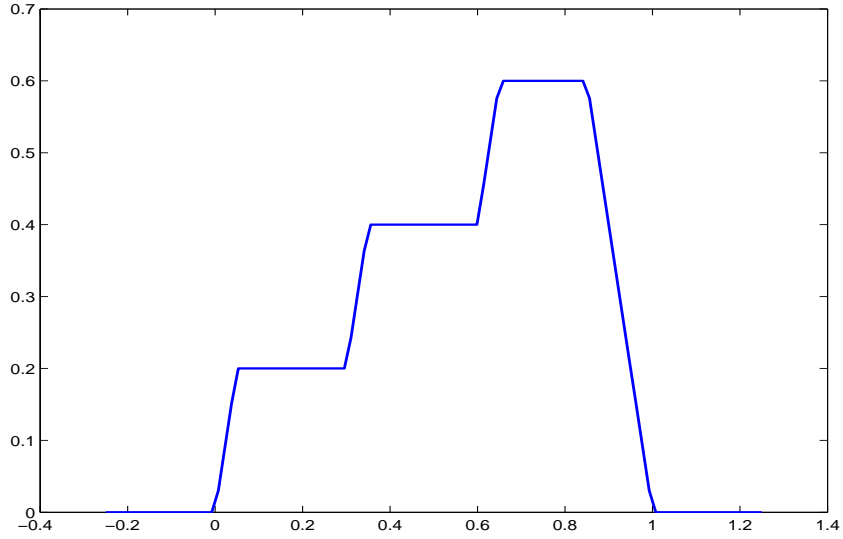


Figura 4: Função de pertinência do conjunto nebuloso  $C$  deduzido pelo SRNs da Tabela 1 quando uma perda de 0.6 e um atraso de 0.2 são apresentados.

2007]. Neste trabalho, adotamos o *centro de área* como método de defuzzificação. O centro de área define  $z^*$  através da seguinte equação onde  $\{z_1, \dots, z_n\}$  é uma partição do universo de discurso  $\mathbf{Z}$ :

$$z^* = \frac{\sum_{j=1}^n C(z_j)z_j}{\sum_{j=1}^n C(z_j)}. \quad (4)$$

Por exemplo, o método de defuzzificação dado pela Equação 4 aplicado ao conjunto  $C$  da Figura 4 fornece o valor 0.57. Esse valor indica a qualidade do serviço VoIP para uma taxa de 0.6 pacotes perdidos e um *jitter* de atraso de 0.2.

É importante observar que um SRNs com o método de inferência de Mamdani e o centro de área como técnica de defuzzificação representa um aproximador universal, i.e., esse processo pode ser usado para aproximar uniformemente qualquer função contínua definida em um compacto [Bando, 2002, Castro and Delgado, 1996].

## Experimentos Computacionais e Considerações Finais

Para avaliar o desempenho do modelo baseado no SRNs da Tabela 1, consideramos

os dados adquiridos por Ferrari [Ferrari, 2007]. Esses dados contém 47 amostras que foram coletadas usando uma variação do sistema QVoIP, que emprega o codec G. 711. Para cada amostra, temos a taxa de pacotes perdidos, o *jitter* de atraso e uma nota atribuída para a qualidade do serviço<sup>1</sup>. As notas foram obtidas calculando a média das notas fornecidas por 7 pessoas. É importante observar que todas as variáveis, i.e., perda, atraso e nota, foram confinadas ao intervalo  $[0, 1]$ , por meio de uma transformação afim.

O *Erro Quadrático Médio* (EQM) e *Erro Percentual Médio Normalizado* (EPMN) obtido considerando as 47 amostras foram 0.008 e 10%, respectivamente. Lembre-se que essas medidas são calculadas através das seguintes equações onde  $N = 47$  refere-se ao número de amostras, os valores  $q_1, \dots, q_N$  denotam a média das notas fornecidas pelas 7 pessoas e  $\hat{q}_1, \dots, \hat{q}_N$  correspondem aos valores fornecidos pelo SRNs.

$$EQM = \frac{1}{N} \sum_{i=1}^N (q_i - \hat{q}_i)^2 \quad \text{e} \quad EPMN = \frac{100}{N} \times \left( \sum_{i=1}^N \frac{|q_i - \hat{q}_i|}{q_i} \right). \quad (5)$$

Realizamos também experimentos considerando uma *rede neural artificial*. Especificamente, utilizamos o *Perceptrons de Múltiplas Camadas* com uma única camada oculta com 5 neurônios artificiais e treinado com o *Algoritmo de Retropropagação* tradicional [Haykin, 1999, Hassoun, 1995, Ripley, 1996]. É importante observar que esse é um dos modelos de rede neural artificial mais empregado na literatura. Sobretudo, vários pesquisadores utilizaram essa rede neural artificial para avaliar a qualidade de um serviço VoIP [Ferrari, 2007, Mohamed et al., 2001, Sun and Ifeachor, 2006].

Para avaliar o desempenho da rede neural artificial, dividimos aleatoriamente as 47 amostras em 3 conjuntos [Haykin, 1999, Ripley, 1996]. Precisamente, 15 amostras foram usadas para treinamento (i.e., para determinar os valores das conexões sinápticas da rede neural), 10 amostras foram usadas para validação (i.e., foram usadas como critério de parada do algoritmo de retropropagação) e 22 amostras foram usadas para teste (para avaliar realmente a performance da rede neural). O EQM e o EPMN fornecidos pela rede

---

<sup>1</sup>Especificamente, as notas para a QoS correspondem à uma normalização do *Mean Opinion Score* (MOS). [ITU, 1996]

neural após 100 simulações foram 0.003 e 4%, respectivamente. Note que esses valores são menores que os erros fornecidos pelo SRNs. Entretanto, algumas observações podem ser feitas com respeito às redes neurais artificiais e o SRNs.

Redes neurais artificiais são modelos que “aprendem” com dados. Portanto, uma rede neural terá uma boa performance somente se uma grande quantidade de dados estiverem disponíveis [Hassoun, 1995, Ripley, 1996]. No nosso caso, temos apenas 47 amostras que foram divididas em 3 conjuntos. Logo, temos poucos dados. Além disso, como muitos pesquisadores afirmam, redes neurais artificiais representam “caixas-pretas”, i.e., não conseguimos extrair muitas informações sobre esses modelos [Pedrycz and Gomide, 2007].

Em contraste, SRNs são projetados partindo da nossa concepção de como deve ser a qualidade do serviço VoIP conhecidos a taxa de pacotes perdidos e o *jitter* de atraso da rede. Portanto, não precisam de um grande banco de dados, como é o nosso caso. Sobretudo, SRNs são modelos transparentes e de fácil compreensão. Pretendemos coletar mais dados sobre a qualidade do serviço VoIP no futuro. Esses dados serão usados para ajustar (ou refinar) os conjuntos nebulosos que definem os antecedentes e os conseqüentes do SRNs.

### **Agradecimentos**

Os autores gostariam de agradecer Hevelin Ferrari pelos dados numéricos usados nos experimentos computacionais. Esse trabalho teve apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES).

### **Referências Bibliográficas:**

ANATEL - Agência Nacional de Telecomunicações, Ministério das Comunicações, Governo Brasileiro. **Página Eletrônica da Agência Nacional de Telecomunicações.** Disponível no endereço: <http://www.anatel.gov.br/>. Acessado em 30 de Junho de 2008.



- ITU - International Telecommunication Union. **Methods for Subjective Determination of Transmission Quality**. ITU-T Recommendation P.800. Series P: Telephone Transmission Quality; Methods for Objective and Subjective Assessment of Quality, 1996.
- ITU - International Telecommunication Union. **The E-model, a computational model for use in transmission planning**. ITU-T Recommendation G.107. Series G: Transmission Systems and Media, Digital Systems and Networks, 2005.
- Bando, F. **Sistemas fuzzy e aproximação universal**. Dissertação de mestrado, Universidade Estadual de Campinas - Instituto de Matemática, Estatística e Computação Científica, 2002.
- Banon, G., Barrera, J., and Braga-Neto, U. **Mathematical Morphology and its Applications to Signal and Image Processing**. INPE, São José dos Campos, 2007. Proceedings of the 8th International Symposium on Mathematical Morphology.
- Barros, L. e Bassanezi, R. **Tópicos de Lógica Fuzzy e Biomatemática**. Editora do Instituto de Matemática, Estatística e Computação Científica (IMECC-Unicamp), Campinas, 2006.
- Birkhoff, G. **Lattice Theory**. American Mathematical Society, Providence, 3 edition, 1993.
- Burden, R. L. and Faires, J. D. **Numerical Analysis**. Brooks Cole, 8th edition, 2004.
- Castro, J. L. and Delgado, M. **Fuzzy systems with defuzzification are universal approximators**. IEEE Transactions on Systems, Man, and Cybernetics, Part B, 26(1):149 – 152, 1996.
- Ferrari, H. V. **Análise e classificação de transmissões de voz sobre IP**. Trabalho de conclusão de curso de graduação. Universidade Estadual de Londrina, 2007.
- Hassoun, M. H. **Fundamentals of Artificial Neural Networks**. MIT Press, Cambridge, MA, 1995.
- Haykin, S. **Neural Networks: A Comprehensive Foundation**. Prentice Hall, Up-

per Saddle River, NJ, 1999.

Heijmans, H. **Morphological Image Operators**. Academic Press, New York, NY, 1994.

Kaburlasos, V. and Ritter, G. **Computational Intelligence Based on Lattice Theory**. Springer-Verlag, Heidelberg, Germany, 2007.

Kaburlasos, V. G. **Towards a Unified Modeling and Knowledge-Representation based on Lattice Theory**, volume 27 of *Computational Intelligence and Soft Computing Applications Series (Studies in Computational Intelligence)*. Springer-Verlag New York, Inc., Secaucus, NJ, USA, 2006.

Klir, G. J. and Yuan, B. **Fuzzy Sets and Fuzzy Logic: Theory and Applications**. Prentice Hall, Upper Saddle River, NY, 1995.

Kosko, B. **Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence**. Prentice Hall, Englewood Cliffs, NJ, 1992.

Lee, K. H. **First Course on Fuzzy Theory and Applications**, volume 27 of *Advances in Soft Computing*. Springer, Berlin, Germany, 2005.

Manousos, M., Apostolacos, S., Grammatikakis, I., Mexis, D., Kagklis, D., and Sykas, E. **Voice-quality monitoring and control for VoIP**. *IEEE Internet Computing*, 9(4):35–42, 2005.

Mohamed, S., Cervantes-Pérez, F., and Afifi, H. **Integrating networks measurements and speech quality subjective scores for control purposes**. In *Proceedings of the Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM 2001)*., volume 2, pages 641 – 649, Anchorage, Alaska, 2001.

Nguyen, H. T. and Walker, E. A. **A First Course in Fuzzy Logic**. Chapman & Hall/CRC, Boca Raton, 2 edition, 2000.

Pedrycz, W. and Gomide, F. **Fuzzy Systems Engineering: Toward Human-Centric Computing**. Wiley-IEEE Press, New York, 2007.

Ripley, B. D. **Pattern Recognition and Neural Networks**. Cambridge University

Press, Cambridge, 1996.

Ritter, G. X. and Wilson, J. N. **Handbook of Computer Vision Algorithms in Image Algebra**. CRC Press, Boca Raton, 2 edition, 2001.

Sun, L. and Ifeachor, E. **Voice quality prediction models and their application in VoIP networks**. *IEEE Transactions on Multimedia*, 8(4):809–820, 2006.

Valle, M. E. **Fundamentos e Aplicações de Memórias Associativas Morfológicas Nebulosas**. Tese de doutorado, Universidade Estadual de Campinas (UNICAMP), Campinas, Brasil, 2007.

Valle, M. E. and Sussner, P. **A general framework for fuzzy morphological associative memories**. *Fuzzy Sets and Systems*, 159(7):747–768, 2008.

Yager, R. and Filev, D. **Essentials of Fuzzy Modeling and Control**. John Wiley and Sons, 1994.

Zadeh, L. A. **Fuzzy sets**. *Information and Control*, 8(3):338–353, 1965.

Zadeh, L. A. **Fuzzy sets and information granularity**. In Gupta, M. M., Ragade, R. K., and Yager, R. R., editors, *Advances in Fuzzy Set Theory and Applications*, pages 3–18. North Holland, Amsterdam, 1979.

Zadeh, L. A. **Fuzzy logic = computing with words**. *IEEE Transactions on Fuzzy Systems*, 4(2):103–111, 1996.

## **REGISTROS DE REPRESENTAÇÃO SEMIÓTICA E TECNOLOGIA NO ENSINO DE FUNÇÕES**

**Tatiany Mottin Dartora –UTFPR- Universidade Tecnológica Federal do**

**Paraná - Campus Toledo**

**tatiany@utfpr.edu.br**

O presente artigo teve como objetivos: compreender problemas ligados ao ensino/aprendizagem de matemática a partir de representações aplicando o uso das tecnologias; analisar como os registros de representação semiótica pode servir como ferramenta para o ensino de matemática; acompanhar a evolução e desempenho dos alunos na construção de conceitos matemáticos através de registros de representação semiótica e uso de tecnologia de informação. Foram sujeitos da pesquisa 10 alunos do curso técnico e ensino médio em gastronomia da Universidade Tecnológica Federal do Paraná, utilizando a sala de informática desta universidade. Os mesmos realizaram atividades individuais com o uso do software Vrum Vrum, com o intuito de aperfeiçoar o ensino de funções de 1ª grau, através de registros de representação semiótica. Os resultados indicaram que: os alunos compreendem melhor noção de funções com auxílio do computador; para o ensino de função é necessário que o indivíduo compreenda, e faça a significação utilizando diferentes representações do mesmo objeto para obter rendimento no estudo. A interação do enunciado verbal do problema, e demais elementos do software utilizado, faz - se essencial o uso de modos de representações que implica no funcionamento semiótico e cognitivo a fim de ser um excelente processo metodológico de ensino/aprendizagem de matemática.

**Palavras chaves:** Registros de Representação Semiótica, Ensino/aprendizagem, Matemática, Funções, Tecnologia.

**pdfMachine**

**A pdf writer that produces quality PDF files with ease!**

Produce quality PDF files in seconds and preserve the integrity of your original documents. Compatible across nearly all Windows platforms, simply open the document you want to convert, click "print", select the "Broadgun pdfMachine printer" and that's it! Get yours now!

A importância desse assunto é justificada pela necessidade que o ensino de funções é encontrado na Educação Matemática, onde estudos mostram que há dificuldades em relação à significação de conceitos matemáticos no processo de ensino aprendizagem.

Para além da introdução de conceitos e métodos para a organização, a interpretação e a análise de dados, é preciso, no desenvolvimento de habilidades que envolvam desde a leitura como o julgamento de informações semiotizadas, aliados aos métodos científicos.

Esta pesquisa inicia-se com um levantamento bibliográfico no qual selecionei algumas pesquisas sobre educação matemática, ligados diretamente ou indiretamente à representação de registros semióticos no ensino de matemática em estudos de funções de 1º grau.

De acordo com Fiorentini (1994), foi executada uma pesquisa exploratória, objetivando proporcionar maior afinidade com o problema, visando torná-lo mais explícito o que, conseqüentemente levou à construção de hipóteses. Além disso procurei na história e epistemologia da matemática para compreender os diversos elementos e desvendar diferentes significados para a noção de funções usando a tecnologia, sendo este parte do trabalho. Ao analisar bibliografias notei a necessidade de encontrar pressupostos para compreender a epistemologia e a partir disso o ensino de matemática a fim de aprofundar as noções sobre significações no ensino da matemática.

Sendo assim os estudos foram fundamentados nos pressupostos teóricos baseado nos registros de representações semióticos de Raymond Duval, e logo após em outros autores que se baseavam no estudo do mesmo.

Em seguida foram desenvolvidas atividades que foram realizadas com o uso do software Vrum Vrum constituindo uma proposta metodológica qualitativa, que desenvolva através de registros de representação semiótica a construção de conceitos apoiando-se nas tecnologias de Informação.

Diante disso optamos nesta pesquisa por uma abordagem qualitativa, pois de acordo com Minayo (1998) a mesma aborda um processo de significados, motivos, valores e atitudes que corresponde a processos e fenômenos que não podemos demonstrar apenas em números.

Para Miskulin (1999), as novas tecnologias requerem uma nova ênfase no currículo. Este deve oferecer condições para que os alunos se sintam capazes de interpretar e verificar resultados numéricos, tabelas e gráficos, de pensarem sobre o

processo de descrever e depurar programas. Além disso, foi abordado o ensino da Matemática frente às novas tecnologias enfatizando que tal ensino contribui efetivamente para a formação integral do indivíduo como um ser capaz de interpretar, compreender e apreciar o mundo que o cerca.

Assim sendo, apresentou-se uma proposta metodológica alternativa, baseada no uso de software livre e Resolução de problemas, a qual enfatizou o dinamismo de condutas cognitivas de dois sujeitos pertencentes aos Estudos de Caso realizados.

Neste sentido, o objetivo é fazer uso da representação semiótica utilizando ambientes computacionais no processo de ensino aprendizagem da matemática.

Apoiando-se nas ferramentas do estudo da semiótica para facilitar o ensino/aprendizagem na educação matemática, em vários segmentos de ensino, principalmente no ensino de funções afim.

Com o objetivo de identificar elementos que revelassem alguns fragmentos de concepções de alunos, referente à noção de função do 1º grau fundamenta-se este trabalho numa pesquisa baseada em Registros de Representação Semiótica de Raymond Duval.

Segundo Duval (1999) a semiótica é a ciência que estuda as maneiras que o indivíduo significa o que o rodeia. O significado da palavra semiótica vem do grego *semeiotiké* ou "a arte dos sinais". E a ciência geral dos signos que estuda sistema de significações e representações de idéias na qual estuda os signos de todos os objetos, difere-se da lingüística que estuda apenas os signos verbais.

A semiótica se originou em três lugares União Soviética, Europa Ocidental e E.U.A, chamado consciencia semiótica ou consciência da linguagem.

Nos E.U.A , século XIX a semiótica teve origem com Charles Sanders Peirce(1839-1914) doutor em Química em Harvard além de ser matemático, físico e astrônomo contribuiu muita para a psicologia. Seus estudos eram apresentados como categorias universais do signos. A primeira categoria compreende ao acaso que se apresenta á consciência, a segunda corresponde a ação reação onde a consciência busca entender e inteiorizar, o último é o processo de mediação e interpretação dos fenômenos. A semiótica pode ser definida como processo de significações é utilizada para as relações entre uma coisa e seu significado.

Santaella (1999) em seu livro “O que é semiótica” esclarece que semiótica é uma ciência que investiga várias linguagens, verificando seus respectivos significados e

sentidos, desvendando sua ação em termos de signo, seu ser de linguagem, isto é, ação de signo.

Neste sentido Raymond Duval apresenta várias contribuições para o estudo da semiótica principalmente no Ensino de Matemática apresentando a importância da linguagem no desenvolvimento das aprendizagens intelectuais

Os registros são modos simbólico de se representar um objeto matemático onde há diversos tipos de representação para um mesmo objeto. Um objeto matemático não deve ser confundido com a representação que dele se faz, devemos dar importância ao conteúdo representado e não a forma como é representado. Há, portanto, uma diversidade de representação semiótica que Duval agrupa em quatro grandes registros: a língua natural, as escritas algébricas e formais, as figuras geométricas e as representações gráficas.

Segundo Raymond Duval (2003), o ensino de matemática serve para que o aluno desenvolva capacidades de raciocínio, visualização e análise, pois a atividade matemática assume um fator importante pela multiplicidade de dependências da representação semiótica. Os registros de Representação Semiótica vêm sendo empregados em pesquisas com objetivo de discutir a aquisição do conhecimento ou organização de situações de aprendizagem aplicadas com grande importância na matemática, onde se torna possível efetuar algumas funções cognitivas.

Vemos que os alunos vêem como um obstáculo à interpretação de situações problemas que envolvem funções, o papel da representação de funções no ensino da matemática vai além, portanto, de ser aquele ligado a comunicação e organização de dados. O uso deste modo de representação implica num estudo do funcionamento semiótico e cognitivo a fim de se destacar os procedimentos metodológicos que geram aprendizagens matemáticas.

Na matemática há diferentes tipos de representação como: sistemas de numeração, escrita algébrica, língua natural, gráficos entre outros, onde agrupando os mesmos temos como resultado registros de representação semiótica.

Sendo assim Duval divide em registros multifuncionais que diz respeito aos tratamentos não algoritmizáveis. Compreendem a língua natural e em registros monofuncionais cujo tratamento é algoritmizáveis tais como: sistemas de escritas-numéricas, algébricas, cálculo e os de representação discursiva: Gráfico cartesiano.

Portanto a compreensão de objetos matemáticos se dá no momento em que somos capazes de mobilizar diferentes tipos de registros de representação do mesmo objeto.

Acreditamos, portanto, que devemos investir um pouco mais na representação gráfica, bem como nas conversões entre os registros gráfico e algébrico ou outras. Visto que segundo Duval (2003, p. 42) “Conversão é a transformação desta representação em uma representação de um outro registro”.

Onde tratamento é uma transformação que ocorre nas representações semiótica dentro de um mesmo registro, como resolver uma função. E Conversão é uma transformação que consiste em mudar de registro mantendo o mesmo objeto.

Além do que Santaella (1999) aponta que a Semiótica busca o significado e sentido das linguagens, e procura descobrir a ação da mesma em termos de signo. Seu ser de linguagem, ou seja, sua ação de signo.

O estudo das funções dá grande valor à leitura e interpretação da linguagem gráfica que dá significado às variações das grandezas envolvidas, e possibilita análise para prever resultados. Paraná (2007, p. 29) acrescenta que:

A metodologia de ensino requer que o aluno consiga identificar os diferentes tipos de funções e explorar seus significados sejam por meio das abstrações em direção ao concreto ou por meio de sua vivência e experiência que partem do concreto e direciona ao abstrato. Por este aspecto o estudo dos gráficos é importantes instrumentos para tornar mais significativas as resoluções das funções e inequações algébricas e as respectivas análises e interpretações.

No que diz respeito à semiótica Duval (1999), aponta que Pierce foi quem reconheceu a importância dos tipos de signos classificando-os em ícones, símbolos e índice. Além do que nos mostra a importância do estudo da semiótica, ciência que estuda os signos, para a compreensão da linguagem matemática. Duval (1995, p.17) aponta que:

Na matemática a especificidade das representações consiste em que elas são relativas a um sistema particular de signos, à linguagem, à escrita algébrica ou aos gráficos artesanais e elas podem ser convertidas em representações equivalentes num outro sistema semiótico, podendo tomar significações diferentes pelo sujeito que as utiliza.

Atualmente várias escolas possuem laboratório de informática ou alguma outra mídia tecnológica para facilitar o ensino aprendizagem de funções, têm-se disponíveis softwares educativos livres, lembrando que a utilização dos mesmos deverá ser feita para construir uma prática de ensino com qualidade. “Compreende prática matemática

**pdfMachine**

**A pdf writer that produces quality PDF files with ease!**

Produce quality PDF files in seconds and preserve the integrity of your original documents. Compatible across nearly all Windows platforms, simply open the document you want to convert, click “print”, select the “Broadgun pdfMachine printer” and that’s it! Get yours now!



toda atuação ou expressão (verbal, gráfica, etc.) realizada por alguém para resolver problemas matemáticos, comunicar a outros a solução obtida, validá-la ou generalizá-la a outros contextos e problemas”. (Godino y Batanero, 1994, p. 334).

Miskulin (1999) nos apresenta que os alunos devem ter uma boa educação matemática para que tenham oportunidades de crescer sendo importante para qualificação profissional, para isso deve ser inserido um contexto tecnológico no ensino aprendizagem, para que os alunos explorem a tecnologia adquirindo maior conhecimento.

Sendo assim o uso de recursos tecnológicos sejam eles os softwares, a televisão, as calculadoras, os aplicativos da Internet, entre outros, têm favorecido as experimentações matemáticas e potencializado formas de ensinar matemática. E através dos mesmos devemos trabalhar com a semiótica para identificar representação e registros dos objetos e converter em significado, fornecendo para o docente uma maior compreensão do ensino aprendizagem de matemática.

Ampliando esta discussão, Papert (1985, p.50) assevera que:

O educador deve atuar como antropólogo. E, como tal, sua tarefa é trabalhar para entender que materiais dentre os disponíveis são relevantes para o desenvolvimento intelectual. Assim, ele deve identificar que tendências estão ocorrendo no meio em que vivemos. Uma intervenção significativa só acontece quando se trabalha de acordo com essas tendências. Em meu papel de educador-antropólogo eu vejo novas necessidades sendo geradas pela penetração dos computadores na vida das pessoas.

Interagindo com a tecnologia e mídias, o docente traz o universo audiovisual para auxiliar no ensino. Moran (2002) considera que o modelo semiótico encontra-se inseparável do conteúdo e o processo de comunicação.

Contudo se tratando das funções cognitivas Duval (1999) apresenta quatro tipos de funções. A função de transmissão das informações e mensagens entre os individuo faz-se necessário diante de códigos ou ferramentas de comunicação. Além disso, apresenta a função de converter uma representação em outra, essa sendo denominada sistema de representação mobilizada que é a função de tratamento que possibilita conscientização daquilo que até então não havia feito, sendo esta função um trabalho de interiorização, onde o individuo desfruta da consciência máxima onde o individuo desfruta do que aprendeu.

De certa maneira a função de objetivação acontece quando há organização das informações e a representação semiótica desempenhada, uma função de objetivação é

uma função de tratamento, além da função de comunicação. Essas representações semióticas têm dois aspectos: a sua forma (o representante) e o seu conteúdo (o representado). E também a função de identificação a qual provem da análise e identificação das informações.

Estas funções segundo Duval (1999) são fundamentais para o funcionamento cognitivo. A identificação, como já foi dito aqui, é uma função fundamental para o tratamento destas representações. Não menos importante, percebemos o papel de destaque que é dado à representação gráfica nos livros didáticos, uma vez que há um interesse crescente pelo recurso às novas tecnologias da informação. Nos processos de produção, transmissão, recepção e interpretação de signos identificaram-se, aspectos de emissores e receptores, na qual a distância da dimensão a ser interpretada é a semiótica da mensagem. O computador auxilia em novas formas de pensar e agir. Completando o que foi dito Arruda (2006), preocupa-se com a utilização correta das inúmeras tecnologias educacionais para buscar soluções ao problema de comunicação na educação.

As mídias digitais representadas pelos computadores tanto como softwares instalados ou o uso da internet e intranet auxiliam no aprendizado desde que os professores saibam como trabalhar com os registros de representações.

Com as ferramentas tecnológicas, mídias e sistemas culturalmente produzidos o sujeito se desenvolve cognitivamente. Vygotsky (1988).

De acordo com as pesquisas realizadas sobre semiótica são de grande importância o tratamento de informação e do registro de representação semiótica no ensino/aprendizagem de matemática, porém, ainda não chegaram às escolas integralmente. Com a globalização os problemas dão ênfase para reprodução de modelos do que para compreensão conceitual. Essa prática não atende as necessidades da sociedade tecnológica que exigem uma educação mais significativa.

Gasparetto (1999) salienta que os significados não são relevados apenas na leitura dos símbolos porém é importante interpretar e compreender a relação com o mundo e as experiências vividas

Soares, M. E Nehring (2006, p.7) apud Duval (2003) afirma:

Neste sentido, para compreender a linguagem matemática é necessário entender a teoria da linguagem - semiótica - que estuda os sistemas de signos, ou seja, a linguagem formal, porque de acordo com Duval, a diferença entre a atividade cognitiva, requerida pela matemática e aquela requerida em outras áreas do conhecimento, não deve ser procurada nos conceitos. A atividade matemática deve ser

estudada naquilo que ela tem de específico, ou seja, no que a diferencia do trabalho de um botânico, ou de um físico em seu laboratório, pois os objetos matemáticos não são objetos diretamente perceptíveis ou observáveis com a ajuda de instrumentos. Isso os leva a depender das representações semióticas, para comunicação e realização das funções de objetivação (expressão particular) e de tratamento (ex: cálculo).

Para Duval (2003) a teoria de registros de representação semiótica compreende um aprendizado cognitivo, no ensino de matemática se mostra através de símbolos, signos, gráficos, significados permitindo registros de representação diferenciada do mesmo objeto matemático. Em sua pesquisa aponta que para isso a diferentes transformações de representações semiótica: o tratamento e a conversão. O primeiro indica transformações de representações dentro de um mesmo registro. E a conversão consiste em mudar de registro conservando os objetos denotados (ex. passar da escrita algébrica de uma equação à sua representação gráfica).

	REPRESENTAÇÃO DISCURSIVA	REPRESENTAÇÃO NÃO-DISCURSIVA
REGISTROS MULTIFUNCIONAIS: Os tratamentos não são algoritmizáveis.	Língua Natural Associações verbais (conceituais) Forma de raciocinar: <ul style="list-style-type: none"> <li>• argumentação a partir de observações, de crenças...;</li> <li>• dedução válida a partir de definições ou de teoremas.</li> </ul>	Figuras geométricas planas ou em perspectivas (configurações em dimensão 0, 1, 2 ou 3). <ul style="list-style-type: none"> <li>• apreensão operatória e não somente perceptiva;</li> <li>• construção com instrumentos.</li> </ul>
REGISTROS MONOFUNCIONAIS: Os tratamentos são principalmente algoritmos.	Sistemas de escritas: <ul style="list-style-type: none"> <li>• numéricas (binária, decimal, fracionária...);</li> <li>• algébricas;</li> <li>• simbólicas (línguas formais).</li> </ul> Cálculo	Gráficos cartesianos <ul style="list-style-type: none"> <li>• mudanças de sistema de coordenadas;</li> <li>• interpolação, extrapolação.</li> </ul>

Fonte: Duval (2003, p.14)

E ainda faz referência a um sistema de signos que permite cumprir as funções de comunicação, tratamento e objetivação. No qual a semiótica é constituída do pensamento e do conhecimento.

**pdfMachine**

**A pdf writer that produces quality PDF files with ease!**

Produce quality PDF files in seconds and preserve the integrity of your original documents. Compatible across nearly all Windows platforms, simply open the document you want to convert, click "print", select the "Broadgun pdfMachine printer" and that's it! Get yours now!

É indispensável que o professor tenha claro o objeto matemático a ser ensinado, quais os registros de representação semiótica que possibilitam a compreensão, para que os alunos encontrem as representações. Sendo assim procuramos na teoria dos registros de representação semiótica possibilidades para compreensão e interpretação dos problemas da realidade na qual estão inseridos para compreenderem o mundo atual e tecnológico. Neste sentido a função do professor é proporcionar uma aula ativa com o uso de computadores e outras tecnologias, e perceber se está acontecendo o caminho entre o objeto e o signo.

Através da metodologia de Lorenzato (2006), surgiram perguntas da reflexão do educador sobre sua própria prática e sobre a prática do outro, e aquelas geradas a partir da investigação dos estudos precedentes ou da própria literatura. Estudos recentes apóiam-se em conjectura que os professores produzem, na prática, saberes práticos sobre a Matemática escolar, mostram que esses saberes práticos transformam-se continuamente, sobretudo quando realizam uma prática reflexiva ou investigativa. Lorenzato(2006, p.51-52) interroga:

Os professor devem questionar em suas aulas “Para que servirá aos meus alunos aprender esse conteúdo? Quais são os conceitos fundamentais desse conteúdo? De quais meios e estratégias disponho para proporcionar a aprendizagem?”

Sendo assim foram realizados diálogos enquanto realizavam atividades num ambiente interativo e dinâmico despertando o interesse dos mesmos. A educação recebe fortes influências dos avanços produzidos nas áreas de informática, tecnologia educacional.

Além do mais, o software Vrum Vrum é livre de fácil acesso e um programa leve. Produzido por V.Teodoro e F.Clérigo, da Universidade Nova de Lisboa, que possibilita que se trabalhe o entendimento gráfico de deslocamento e velocidade no tempo.

Este programa foi desenvolvido com objetivo de auxiliar o processo de ensino/aprendizagem de conceitos relacionados à função de 1º grau. As atividades exploradas pelo sistema baseiam-se no personagem Bugão que deve se deslocar por um caminho com medidas exposta onde apresentam fatores como tempo posição, trajeto, com o intuito de estimular o interesse do aluno na compreensão de função, pois envolve os estudantes em um processo dinâmico de resolução de problemas.

A fim de enfatiza a valorização do raciocínio utilizado pelo sujeito, quando este resolve desafios e problemas, além disso, valoriza a análise do processo de pensamento,

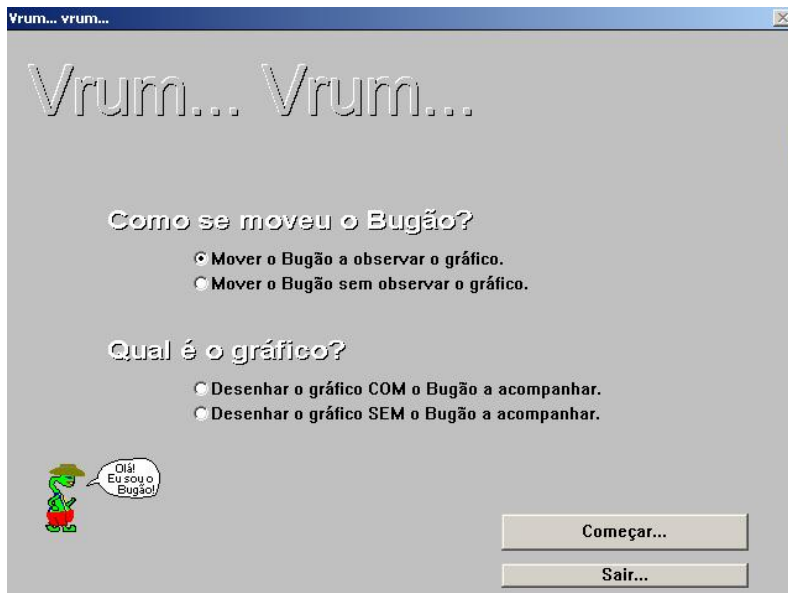
a seqüência de ações do sujeito, quando inserido em situações práticas de resolução de problemas. Pois ao realizar as atividades para se resultado adquire-se grande significação durante o processo do mesmo.

Diante dessa perspectiva, faz-se necessário salientar a importância do software em permitir ao aluno “buscar” suas ações através dos “botões” e “ícones” existentes no sistema. Além disso, permite também ao professor analisar a seqüência dessas ações, nas atividades propostas pelo sistema.

Na concepção da pesquisadora Miskulin (1998), os ambientes computacionais são extremamente úteis e importantes para a exploração e construção de conceitos, porém ressalta que os resultados obtidos dependem muito da intervenção do professor, de como este intervém no processo ensino/aprendizagem.

Nesta perspectiva no laboratório de informática da universidade Tecnológica Federal do Paraná, desenvolvemos um trabalho de ensino/aprendizagem de funções afim com o uso do software Vrum Vrum participaram 10 alunos do 1º Ano do Curso Técnico de Ensino Médio em gastronomia. Realizaram as atividades relacionadas no programa do software apresentado.

Está figura é a pagina de comandos iniciais:



Análise de dados

Nível 1

\* Mover o Bugão sem observar o gráfico

-ver o gráfico

-pôr na posição inicial

pdfMachine

A pdf writer that produces quality PDF files with ease!

Produce quality PDF files in seconds and preserve the integrity of your original documents. Compatible across nearly all Windows platforms, simply open the document you want to convert, click “print”, select the “Broadgun pdfMachine printer” and that’s it! Get yours now!

- Começar o movimento
- ver solução se necessário

Nível 2

\*Relacionar o gráfico ao movimento

\*Desenhar o gráfico com o Bugão e acompanhar

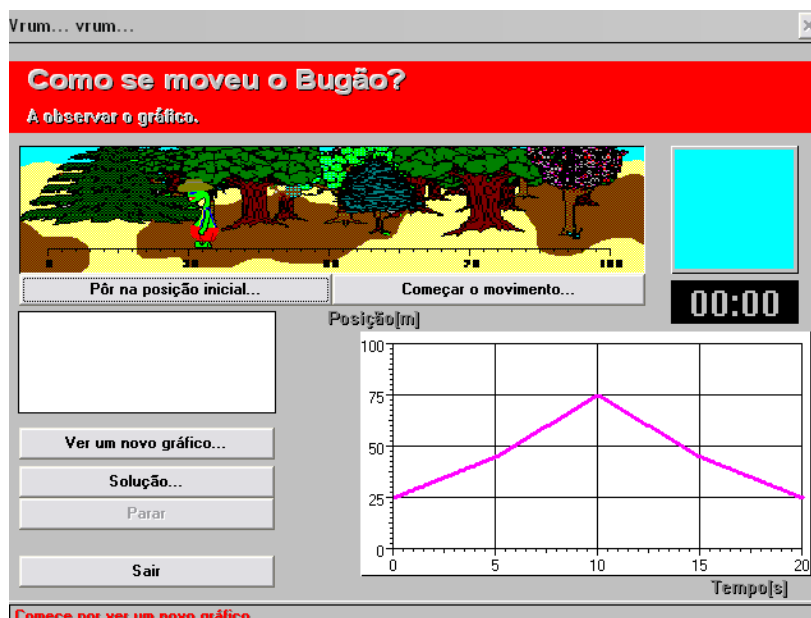
Desenhar o gráfico sem o Bugão e acompanhar

-ver um novo movimento

-repetir o movimento se necessário

-traçar o gráfico

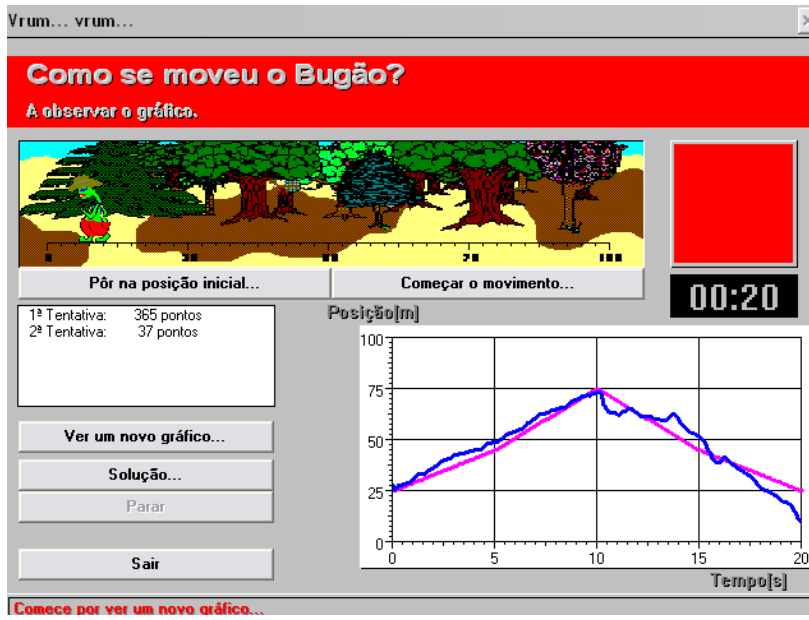
1- Mover o Bugão e relacionar o gráfico



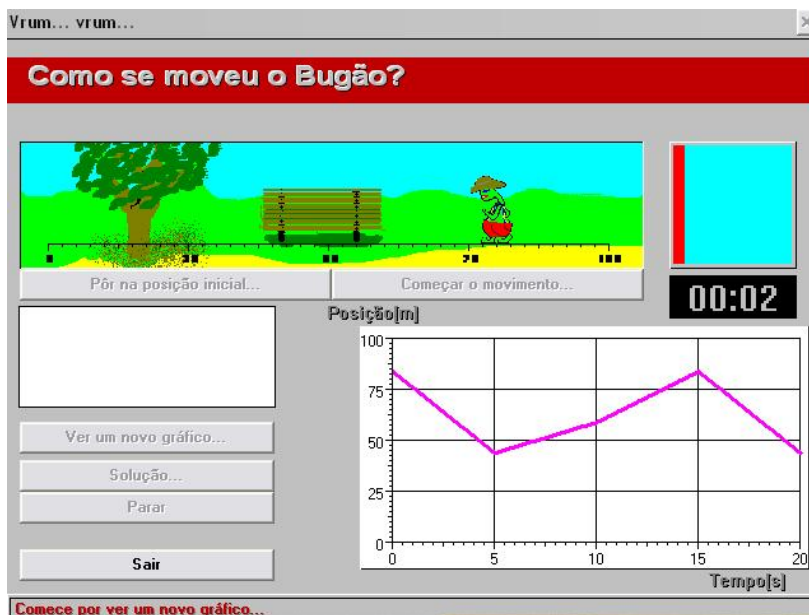
pdfMachine

A pdf writer that produces quality PDF files with ease!

Produce quality PDF files in seconds and preserve the integrity of your original documents. Compatible across nearly all Windows platforms, simply open the document you want to convert, click "print", select the "Broadgun pdfMachine printer" and that's it! Get yours now!



## 2-Mover o Bugão sem relacionar o gráfico



**Obs:** Enquanto o Bugão percorre sua trajetória não aparece o gráfico do percurso que o mesmo faz segundo o comando do aluno

pdfMachine

A pdf writer that produces quality PDF files with ease!

Produce quality PDF files in seconds and preserve the integrity of your original documents. Compatible across nearly all Windows platforms, simply open the document you want to convert, click "print", select the "Broadgun pdfMachine printer" and that's it! Get yours now!



Com essas atividades buscou o entrosamento de todos bem como um roteiro a ser seguido:

1. Qual a relação entre o Bugão e o gráfico?
2. Qual elemento físico está representado?
3. Quais as variáveis envolvidas?
4. Quais as representações de funções que notamos com este software?
5. Qual a influência do software no aprendizado de matemática?
6. Qual o entendimento gráfico de alunos sobre o deslocamento em função da velocidade?
7. Quais as representações feitas pelos alunos?

Neste sentido buscamos fazer com que os alunos compreendam através dos elementos presentes no software formas de linguagens sobre os conceitos de função de 1º grau como linguagem natural, gráfica, tabelas, desenhos.

O quadro abaixo resume o comportamento dos alunos nas questões:

Foram capazes de executar com sucesso as transformações do enunciado, movimento para o gráfico;

Demonstraram uma compreensão correta e solucionaram a situação problema;

Observaram que quanto menor a quantidade de pontos feitos significa melhor desempenho da atividade;

Tentaram melhorar sua pontuação;

Observaram que o movimento relaciona-se ao gráfico, onde, no intervalo em que o



Bugão andava para a direita o gráfico era crescente e quando o Bugão andava a direita o intervalo do gráfico era decrescente;

As representações utilizadas gráfica, de movimento e outras se constituem num importante recurso para a análise de dados e tratamento da informação. Onde os meios de comunicação, sejam eles virtuais ou impressos, são aplicados em quadros, tabelas e/ou gráficos, ou seja, de uma linguagem semiotizada que permite a comunicação de dados e informações.

As noções de funções são validas e constitui grande importância, pois é encontrada diariamente em jornais, revistas, fornecidas para que aconteça a fácil visualização, porém nem sempre o indivíduo sabe fazer a interpretação correta.

Verificou-se nesta pesquisa que para a interpretação das atividades os alunos sentiram a necessidade de fazer a conversão das informações ao lidar com tabelas, gráficos, e demais elementos demonstrados na tela do software Vrum Vrum, para organizar e comunicar dados e informações, o que implica numa análise do funcionamento representacional destes modos de representação.

## CONSIDERAÇÕES FINAIS

Nesta perspectiva constata-se que o ambiente computacional é um fator de grande influência na educação matemática, onde as novas tecnologias estão presentes para auxiliar neste processo. No ensino de funções de 1º grau geralmente poucos alunos manuseiam mais de um registro de representação semiótica, sendo o registro algébrico o mais utilizado. Sabendo que os alunos deverão fazer as conversões entre os registros gráficos, algébricos entre outros para de fato compreenderem os conceitos de funções, suas variáveis, e aplicações. Além do que é importante que aconteça a construção do conceito de funções em diversas situações para que permeie um aprendizado significativo entre os registros de representação semiótica. Como afirma Flores (2006) que o uso deste modo de representação implica num estudo do funcionamento semiótico e cognitivo a fim de se destacar os procedimentos metodológicos que geram aprendizagens matemática.

## 6. Referências Bibliográficas

BORBA, M. C. **A Pesquisa Qualitativa Em Educação Matemática.** In: Anais da 27ª reunião anual da Anped, Caxambu, MG, 21-24 Nov. 2004;

BRANDT, Célia Finck.. **Contribuições dos registros de representação semiótica na conceituação do sistema de numeração.** Disponível em: < <http://www.ppgect.ufsc.br/teses/02/Tese.pdf>> Acesso 05 Jan. 2008.

CARVALHO NETO, C. Z. **Espaços ciberarquitetônicos e a integração de mídias, por meio de técnicas derivadas de tecnologias dedicadas à educação.** 2006. Disponível em: < <http://www.ppgect.ufsc.br/dis/39/resumo.pdf>> Acesso 10 Fev. 2008.

DUVAL, Raymond. **Registros de representações semiótica e funcionamento cognitivo da compreensão em matemática.** In: Aprendizagem em matemática: registros de representação semiótica. Organização de Silvia Dias Alcântara Machado, p.11- 33. Campinas, São Paulo: Papirus, 2003.

FIORENTINI, D. **Alguns modos de ver e conceber o ensino da Matemática no Brasil.** Zetetiké. Campinas: UNICAMP/FE/CEMPEM. Ano 3, n. 4, pp. 1-37, 1995.

FLORES, C. R. **O funcionamento cognitivo e semiótico das representações gráficas: ponto de análise para a aprendizagem matemática.** In: Artigo publicado (em CD) nos anais da 28ª Reunião da Anped. Caxambu, 2005.

GASPARETTO, Maria Marchi. **Didática Comunicativa em Educação Matemática:** Ijuí, RS, Unijuí, Dissertação de Mestrado, 1999.

Kamii, C. e DeClark, G., **Reinventando a Aritmética. Implicações da teoria de Piaget.** 3a. edição. Campinas, Papirus Editora, 1990

MACHADO, Silvia Dias Alcântara (org.). **Aprendizagem em Matemática: Registros de Matemática.** Belo Horizonte: Autêntica, 2001.

MISKULIN, R. G. S. **Concepções teórico-metodológicas sobre a introdução e a utilização de computadores no processo ensino-aprendizagem da geometria.** Tese de Doutorado. Faculdade de Educação, UNICAMP, Campinas, SP, 1999.

MISKULIN, R. S. **Ambientes computacionais utilizados no ensino.** Campinas, SP: Papyrus, 2001.

MORAN, J. M., MASETTO, M.T., BEHRENS, M.A. **Novas tecnologias e mediação pedagógica.** Campinas, SP: Papyrus, 5ª ed., 2002.

PAIS, Luiz Carlos. 2002. **Didática da Matemática; uma análise da influência francesa. Coleção Tendências em Educação Matemática- 2.** ed. Belo Horizonte: Autêntica.

PAPERT, S. (1985) **Logo: Computadores e Educação.** Trad. J. A. Valente, B. Bitelman, A. V.Ripper. São Paulo: Brasiliense (Tradução de: Mindstorms – Children, Computers and Powerful Ideas). **Representação Semiótica.** 1. ed. Campinas SP. Papyrus, 2003

SANTAELLA, L., **O que é Semiótica,** São Paulo: Brasiliense, 1999.

SOARES, M. e NEHRING, C., **O processo de ensinar e aprender Matemática num mundo globalizado e os registros de representação semiótica.** Disponível em <http://ccet.ucs.br/eventos/outros/egem/cientificos/cc65.pdf>. Acesso em 07/03/08.

# **ANÁLISE DA INFLUÊNCIA DO TEOR DE UMIDADE INICIAL DO GRÃO, VELOCIDADE E TEMPERATURA DO AR NA SECAGEM DE SOJA**

Camila Nicola Boeri, Oleg Khatchatourian.

Universidade Regional do Noroeste do Estado do Rio Grande do Sul - UNIJUÍ,  
camilaboeri@hotmail.com; Universidade Regional do Noroeste do Estado do Rio Grande do  
Sul - UNIJUÍ, olegkha@unijui.edu.br

## **Resumo:**

O sistema de beneficiamento e armazenamento de soja requer que esse grão seja colhido sadio e antecipadamente. Desta maneira, a secagem constitui uma das operações de primordial importância entre as técnicas envolvidas na conservação das qualidades desejáveis de produtos de origem vegetal. Este trabalho tem como objetivo obter as curvas de secagem de soja em camada fina, na faixa de temperatura do ar de secagem entre 45 e 90 °C, para os teores de umidade inicial entre 0,13 e 0,32 e velocidades do ar de secagem de 0; 0,5; 0,9; 1,5 e 2,5m/s determinando a influência destes parâmetros no processo. Para tanto, foram realizados experimentos em laboratório e simulações numéricas, onde o modelo matemático utilizado foi o proposto por Khatchatourian et al. (2003), sendo que nesta pesquisa modificou-se a equação que descreve o fluxo de massa, obtendo-se uma boa concordância entre os dados experimentais e os simulados. Conclui-se que a velocidade de secagem depende da temperatura, velocidade do ar, do tempo de exposição a estas condições e dos teores de água inicial e final. Quanto mais elevada a temperatura e o fluxo de ar, maior a taxa de secagem e menor o tempo total de exposição ao ar aquecido.

**Palavras-chave:** secagem de soja, modelagem matemática, simulação numérica.

## **1. Introdução**

A produção de grãos ou sementes de soja requer que o produto seja colhido sadio e antecipadamente, visando minimizar as perdas ocasionadas no campo pelo ataque de insetos, doenças e microrganismos; desta maneira e devido ao teor de

umidade elevado por ocasião da colheita, a secagem constitui uma das operações de primordial importância entre as técnicas envolvidas na conservação das qualidades desejáveis de produtos de origem vegetal.

A secagem de produtos agrícolas pode ser definida como um processo simultâneo de transferência de calor e massa entre o produto e o ar de secagem, que consiste na remoção da umidade excessiva contida no interior do grão por meio de evaporação, geralmente causada por convecção forçada de ar aquecido.

Vários parâmetros têm influência no tempo necessário para reduzir a umidade das sementes ao nível seguro para armazenamento. Os principais fatores a considerar são: temperatura, umidade relativa e vazão do ar, umidade inicial e final e tipo de grão.

De acordo com o exposto, objetivou-se com este trabalho obter as curvas de secagem de soja, em camada fina, com temperaturas do ar de secagem de 45, 60, 70, 80 e 90 °C, para os teores de umidade inicial de 0,13; 0,19; 0,22; 0,28 e 0,32 (base seca) e velocidades do ar de secagem iguais a 0m/s; 0,5m/s; 0,9m/s; 1,5m/s e 2,5m/s; determinar a influência do teor de umidade, da temperatura e da velocidade do ar sobre a secagem de soja através de experimentos e realizar simulações numéricas através do modelo matemático proposto, comparando os resultados simulados com os dados experimentais.

## **2. Determinação do teor de umidade inicial dos grãos**

A determinação do teor de umidade dos grãos foi feita através do método direto seguindo-se as normas determinadas pelo Ministério da Agricultura, submetendo-se uma massa de 15g de grãos em uma estufa a uma temperatura de 105 °C, durante 24 horas. As medidas de massa foram feitas usando-se uma balança analítica digital com precisão  $10^{-3}$ g.

## **3. Equipamento utilizado**

Para a realização dos experimentos de secagem de soja em camada fina, foi utilizado um protótipo de um secador, que consistiu de um tubo metálico com 0,15m de diâmetro, isolado em toda a sua superfície com lã de vidro e lona.

Este equipamento pode ser dividido em dois sistemas: um que gera o fluxo de ar e o outro que realiza a secagem. O primeiro é constituído por um ventilador centrífugo, dutos e uma placa de orifício para medir a vazão do ar. O segundo é composto pelo aquecedor do ar, cilindro de secagem e equipamentos para medir a temperatura do ar.

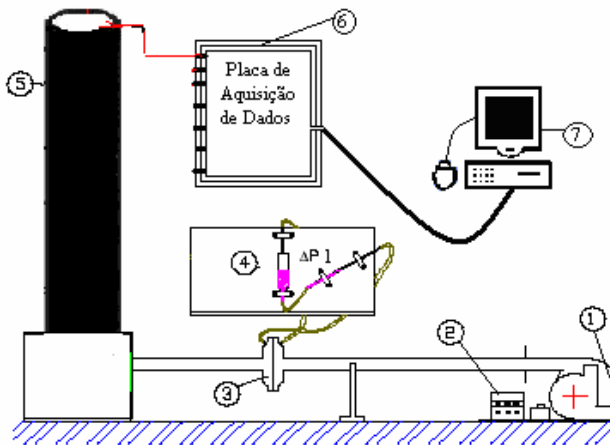


Figura 1: Representação do equipamento de secagem

Para aquecer o ar, foram utilizados seis resistores elétricos com potência de 600W. Para controlar a temperatura do ar de secagem, foi instalado um termopar dentro da câmara de secagem que continha a massa de grãos de soja. Esta câmara era metálica e cilíndrica, com diâmetro de 0,15m, acoplada ao sistema de ventilação/aquecimento do ar, e continha uma tela, sobre a qual era disposta a massa de grãos para ser submetida ao fluxo. A espessura da camada fina de grãos era em torno de 7 mm ou dois grãos.

#### 4. Experimentos de secagem

Após a medida do teor de umidade, a soja foi colocada em embalagens plásticas, fechadas hermeticamente, contendo uma massa de 200g de grãos em cada amostra. O laque era feito com o auxílio de uma máquina lacradora. Depois de lacradas, as amostras eram postas na geladeira a uma temperatura de 4°C, a fim de se conservar o seu teor de umidade.

Vinte e quatro horas antes de iniciar o processo de secagem, a amostra era retirada da geladeira e deixada em contato com o ar ambiente. Após este tempo, era realizado um pré-aquecimento de cinco minutos na estufa, com a mesma temperatura em que seria realizada a secagem. Ao ser retirada da estufa, a massa de grãos era novamente medida.

Em seguida, a amostra era colocada na câmara de secagem e acionada a contagem do tempo. Passado o tempo de secagem desejado, a massa era novamente medida e recolocada no secador, repetindo-se este procedimento para todos os tempos desejados. Para este procedimento, a massa de grãos ficava em torno de 20s fora do

secador, tempo necessário para retirar a amostra, colocá-la na balança, anotar a medida e retorná-la ao secador.

O intervalo de secagem era incrementado de 10 minutos na primeira hora, 20 minutos na segunda hora e a partir da terceira hora, 30 minutos. O experimento era encerrado quando o teor de umidade atingia em torno de 8% (base seca).

O teor de umidade de equilíbrio da soja era calculado para cada condição de umidade relativa do ar e de temperatura de secagem, de acordo com Dalpasquale (1981):

$$X_e = \frac{3,962 \cdot UR_s^{0,4919}}{100 \cdot \ln(T_s)} \quad (1)$$

onde:

$UR_s$  é a umidade relativa do ar de secagem, porcentagem;

$T_s$  é a temperatura do ar de secagem, °C.

Tendo-se os teores de umidade inicial, de equilíbrio e os de umidade instantânea da massa de grãos de soja, obteve-se os valores da razão de umidade para cada processo de secagem, dados pela expressão abaixo:

$$RU = \frac{X - X_e}{X_0 - X_e} \quad (2)$$

onde:

$X$  é o teor de umidade do produto;

$X_e$  é o teor de umidade de equilíbrio;

$X_0$  é o teor de umidade inicial.

## 5. Modelo Matemático

O modelo matemático utilizado nesta pesquisa é baseado no proposto por Khatchatourian et al. (2003). Este modelo foi adaptado para descrever o processo de secagem de grãos de soja num secador de leito fixo e é composto por um sistema de quatro equações diferenciais parciais, hiperbólicas, quase-lineares, acopladas, onde estas equações contemplam as trocas de calor e de massa entre o ar e o grão. Entretanto, apesar de descrever satisfatoriamente as curvas de secagem em relação a influência da temperatura do ar, este modelo não leva em consideração a influência da velocidade do ar e do teor de umidade inicial do grão no processo.

Desta forma, a fim de considerar a influência da velocidade do ar de secagem e do teor de umidade inicial dos grãos, será feito um ajuste do coeficiente de transferência de massa durante o processo de secagem, modificando-o em relação ao modelo proposto por Khatchatourian (2003).

Levando-se em conta as considerações feitas na formulação do problema de secagem de soja num secador de leito fixo, o modelo matemático composto pelo sistema de equações diferenciais parciais (EDP) obtido dos balanços de energia e massa, é formado pelas seguintes equações: Equação de conservação da massa de água em relação ao ar, Equação de conservação de energia em relação ao ar, Equação de conservação da massa de água em relação ao grão e Equação de conservação de energia em relação ao grão. Possui ainda as seguintes equações: Teor de umidade do grão (base seca), Umidade do ar que passa por entre os grãos, Densidade do fluxo de massa e Densidade do fluxo de calor.

### 5.1. Equação de conservação da massa de água no ar

Esta equação determina a variação do teor de umidade do ar em diferentes pontos internos do secador.

$$\frac{\partial Y}{\partial t} + v_a \frac{\partial Y}{\partial z} = \frac{a_{sv} \Phi_m (1 - \varepsilon)}{\varepsilon \rho_a} \quad (3)$$

Onde  $Y$  é o teor de umidade do ar,  $v_a$  é velocidade do ar,  $a_{sv}$  é a razão entre a área da superfície e o volume do grão,  $\varepsilon$  é a porosidade,  $\rho_a$  é a massa específica do ar e  $\Phi_m$  é a densidade do fluxo de massa.

### 5.2. Equação de conservação de energia para o ar

Esta equação determina a energia transferida pelo ar para os grãos.

$$\frac{\partial T_a}{\partial t} + v_a \frac{\partial T_a}{\partial z} = \frac{a_{sv} (1 - \varepsilon)}{\varepsilon \rho_a (C_{pa} + Y C_{pv})} [\Phi_m C_{pv} (T_g - T_a) + \Phi_h] \quad (4)$$

Onde  $T_a$  é a temperatura do ar,  $C_{pa}$  é o calor específico do ar à pressão constante,  $C_{pv}$  é o calor específico do vapor de água à pressão constante,  $T_g$  é a temperatura do grão e  $\Phi_h$  é a densidade do fluxo de calor.



### 5.3. Equação de conservação da massa de água no grão

Esta equação determina a variação do teor de umidade dentro do grão.

$$\frac{\partial X}{\partial t} = - \frac{a_{sv} \Phi_m}{\rho_g} \quad (5)$$

### 5.4. Equação de conservação de energia no grão

Esta equação determina a energia para aquecer o grão e fazer evaporar a sua água.

$$\frac{\partial T_g}{\partial t} = - \frac{a_{sv} \{ \Phi_h + \Phi_m [H_v + T_g (C_{pv} - C_{pw})] \}}{\rho_g (C_{pg} + X C_{pw})} \quad (6)$$

Onde  $T_g$  é a temperatura do grão,  $H_v$  é o calor latente de vaporização,  $C_{pw}$  é o calor específico da água à pressão constante e  $C_{pg}$  é o calor específico do grão.

### 5.5. Teor de Umidade do grão (base seca)

Supondo o esquema de um grão de soja, onde é possível separar a massa de água da massa de matéria seca, o teor de umidade é determinado pela razão entre elas:

$$X = \frac{m_w}{m_{gs}} \quad (7)$$

Onde  $m_w$  é massa de água no grão e  $m_{gs}$  é massa de matéria seca do grão

### 5.6. Umidade do ar que passa por entre os grãos

Considerando o esquema de um determinado volume de ar, a umidade do ar que passa por entre os grãos (umidade do ar existente nos espaços intersticiais) num determinado período de tempo de secagem ou mesmo antes de iniciá-la, pode ser expressa por:

$$Y = \frac{m_{aw}}{m_{as}} \quad (8)$$

Onde  $Y$  é umidade do ar,  $m_{aw}$  é massa de água no ar e  $m_{as}$  é massa de ar seco.

### 5.7. Densidade do fluxo de massa

Com o aquecimento do grão, e o conseqüente aumento da pressão no seu interior, ele rompe o equilíbrio higroscópico e a água, em forma de vapor, tende a sair. A quantidade de massa de água em forma de vapor, que atravessa a superfície do grão por unidade de área e por unidade de tempo é o que determina a densidade do fluxo de massa.

Khatchatourian (2003), usando as relações de Chilton-Colburn (SISSON, 1988), ajustou a equação do fluxo de massa, levando em consideração a influência da velocidade do ar de secagem, obtendo a seguinte relação:

$$\Phi_m = \sqrt{\frac{v_a}{v_0}} M_1(T) \cdot e^{-\sqrt{M_2(T)+M_3(T) \cdot \ln(X-X_e)}} \quad (9)$$

onde:

$\Phi_m$  é a densidade do fluxo de massa e

$$M_1(T) = K_1(X_0 - X_e)$$

$$M_2(T) = K_2 - K_3 \cdot \ln(X_0 - X_e)$$

$$M_3(T) = K_3 \text{ em que } K_1, K_2 \text{ e } K_3$$

são funções temperatura.

### 5.8. Densidade do fluxo de calor

É a quantidade de calor que passa na superfície do grão por unidade de tempo e por unidade de área. A transferência de calor no processo de secagem ocorre, predominantemente, por convecção forçada entre o ar de secagem e a superfície dos grãos. Segundo a lei de Newton-Richmann para transferência de calor convectiva, o fluxo de calor é proporcional à diferença de temperatura entre o ar e o grão, dada por:

$$\Phi_h = h(T_g - T_a) \quad (10)$$

Onde  $\Phi_h$  é a densidade do fluxo de calor,  $h$  é o coeficiente de transmissão de calor,  $T_g$  é a temperatura do grão e  $T_a$  é a temperatura do ar.

Khatchatourian (2003) a partir do coeficiente de transferência de calor proposto por Loncin, utilizado por Courtois e Lasseran (1993) e usando a mesma analogia feita para a transferência de massa, encontrou a seguinte equação para o valor de  $h$ :

$$h = \sqrt{\frac{v_a}{v_0}} (-19,718 + 0,2576T_g + 379,41Y) \quad (11)$$

## 5.9. Modificação do Modelo

O modelo acima descrito foi modificado a fim de ajustar a influência da velocidade do ar de secagem e do teor de umidade inicial dos grãos. A modificação deu-se na equação referente ao fluxo de massa dada por (9).

Assim, o fluxo de massa passa a ser descrito por:

$$\Phi_{me} = (a \cdot v + b) \cdot \left(\frac{X}{X_0}\right)^c \cdot \left(M_1(T) \cdot e^{-\sqrt{M_2(T)+M_3(T) \cdot \ln(X-X_e)}}\right) \quad (12)$$

onde os coeficientes  $a$ ,  $b$  e  $c$  foram obtidos através do método de mínimos quadrados utilizando-se os dados experimentais na minimização do funcional.

Dessa forma, a equação pode ser escrita como:

$$\Phi_{me} = (0,85 \cdot v + 0,15) \cdot \left(\frac{X}{X_0}\right)^{0,856} \cdot \left(M_1(T) \cdot e^{-\sqrt{M_2(T)+M_3(T) \cdot \ln(X-X_e)}}\right) \quad (13)$$

## 6. Solução Numérica do Modelo Matemático

### 6.1. Determinação dos coeficientes através do método de mínimos quadrados

A determinação dos coeficientes  $a$ ,  $b$  e  $c$  da equação (13) que descreve o fluxo de massa foi feita através do Método de Mínimos Quadrados, minimizando-se a funcional dada por:

$$S = \min_{a,b,c} \sum_{i=1}^N \left( \Phi_{me_i}(a,b,c) - \Phi_{\text{exp}_i} \right)^2 \quad (14)$$

sendo  $N$  o número de pontos obtidos experimentalmente ao longo das medidas efetuadas.

Encontraram-se os seguintes valores para os parâmetros desejados:

$$a = 0,85$$

$$b = 0,15$$

$$c = 0,856$$

## 6.2. Método Numérico para solução do Modelo Matemático

A equação que determina o fluxo de massa durante o processo de secagem do modelo matemático desta pesquisa é resolvida numericamente através do Método de Newton, dado por:

$$X^{(k+1)} = X^{(k)} - \text{inv}(J) \cdot \begin{pmatrix} f_1 \\ f_2 \\ f_3 \end{pmatrix} \quad (15)$$

onde:

$$X = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \quad (16)$$

em que  $J$  é a Matriz de Jacobi.

## 7. Curvas de Secagem em Camada Fina

### 7.1. Influência da velocidade do ar no processo de secagem

Nas Figuras 2 e 3 estão apresentadas as curvas de secagem em função da velocidade do ar, mantida a temperatura constante em 45°C para diversos teores de umidade, onde as curvas discretas representam os dados experimentais e as contínuas são os resultados simulados.

Observa-se uma maior retirada de água durante as primeiras horas de secagem. Conforme Hall (1980), após as primeiras horas, a secagem apresentou velocidade decrescente, devido ao transporte interno da água ser menor do que a taxa de evaporação; assim, a transferência de calor não é compensada pela transferência de massa e, portanto, há aumento da temperatura do produto.

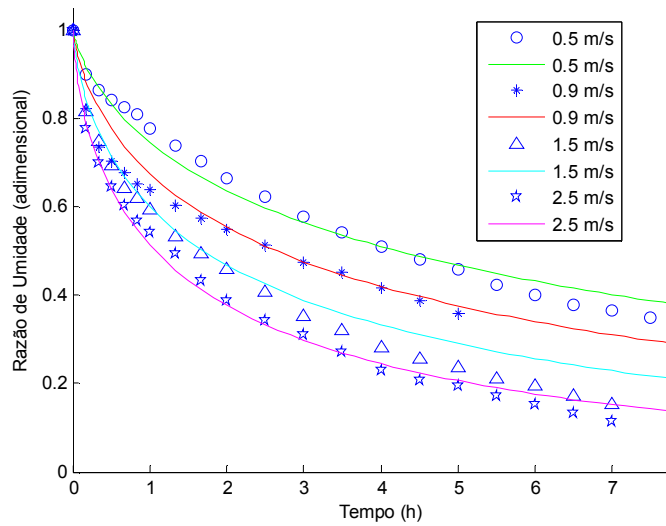


Figura 2: Influência da velocidade do ar.  $T_s = 45^\circ\text{C}$ ,  $X_0 = 0,28$

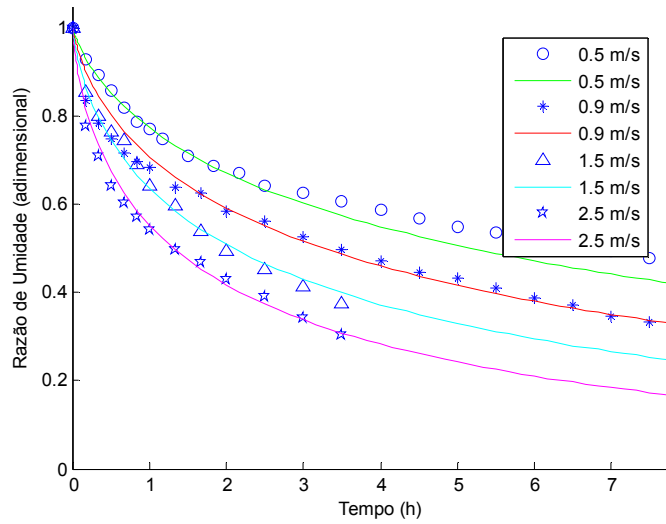


Figura 3: Influência da velocidade do ar.  $T_s = 45^\circ\text{C}$ ,  $X_0 = 0,22$

As Figuras 4 a 7 apresentam as demais curvas de secagem em função da velocidade do ar para os restantes níveis de temperaturas avaliados –  $60^\circ\text{C}$ ,  $70^\circ\text{C}$ ,  $80^\circ\text{C}$  e  $90^\circ\text{C}$  –, onde pode ser observado um comportamento semelhante em todos os casos, conforme descrito anteriormente.

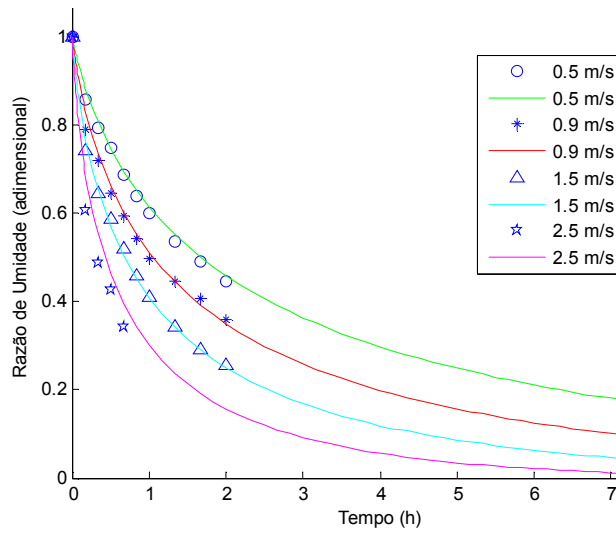


Figura 4: Influência da velocidade do ar.  $T_s = 90^\circ\text{C}$ ,  $X_0 = 0,19$

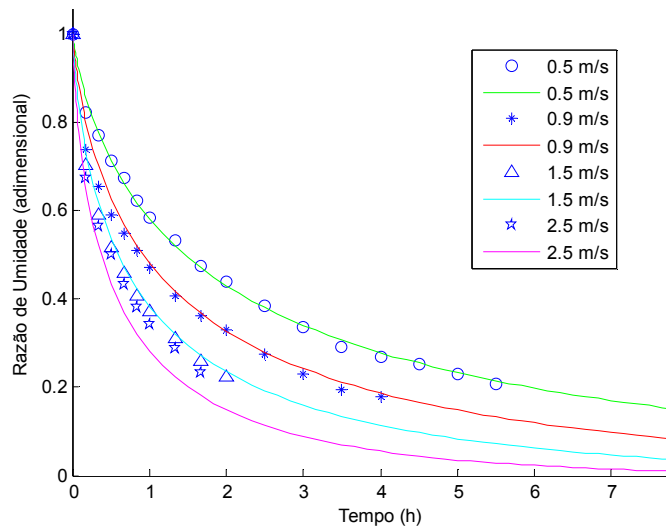


Figura 5: Influência da velocidade do ar.  $T_s = 80^\circ\text{C}$ ,  $X_0 = 0,28$

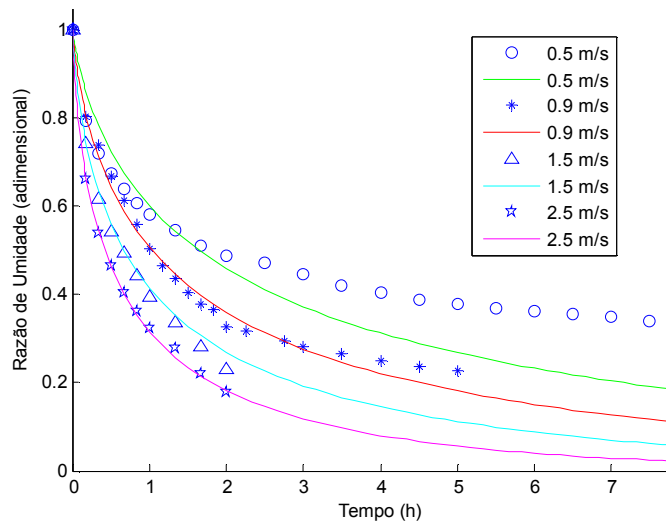


Figura 6: Influência da velocidade do ar.  $T_s = 70^\circ\text{C}$ ,  $X_0 = 0,32$

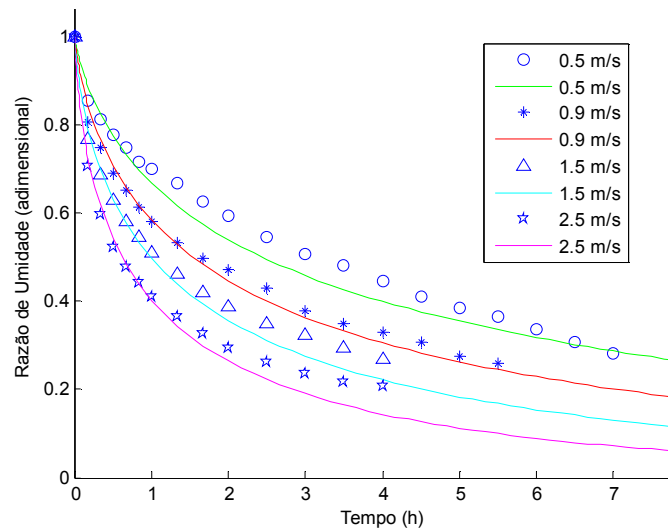


Figura 7: Influência da velocidade do ar.  $T_s = 60^\circ\text{C}$ ,  $X_0 = 0,28$

Na análise dos gráficos acima, fica evidenciada a influência que a velocidade do ar possui no processo de secagem, onde observa-se que este ocorre mais rapidamente para velocidades mais elevadas, comportamento observado para todos os teores de umidade iniciais e temperaturas do ar.

Outro fator importante em relação ao fluxo de ar é que quanto maior a quantidade de ar forçado que passar pela massa de grãos, mais rapidamente à frente de secagem se desloca, diminuindo o tempo de secagem e tornando o processo mais seguro (Queiroz et al., 1985). Isso ocorre pelo aumento do coeficiente de transferência de massa, ou seja, a elevação da velocidade do ar acelera o processo de transferência de massa entre a superfície do grão e o ambiente.

## 7.2. Influência da temperatura do ar de secagem

As Figuras 8 e 9 apresentam as curvas de secagem com a influência da temperatura do ar. Observa-se que a velocidade de secagem aumentou com a elevação da temperatura, ou seja, foi necessário um menor tempo de exposição dos grãos de soja ao ar aquecido para atingir o teor de umidade desejado, devido a maior disponibilidade de energia para a vaporização da água. Isso acontece porque a elevação da temperatura aumenta o coeficiente de difusão de massa. Nota-se também que, no início da secagem, a evaporação de água é mais pronunciada, diminuindo ao longo do tempo, tornando-se nula ao final desta. Neste momento, os efeitos de transferência de massa tornam-se desprezíveis, quando os grãos tendem a alcançar os teores de umidade de equilíbrio.

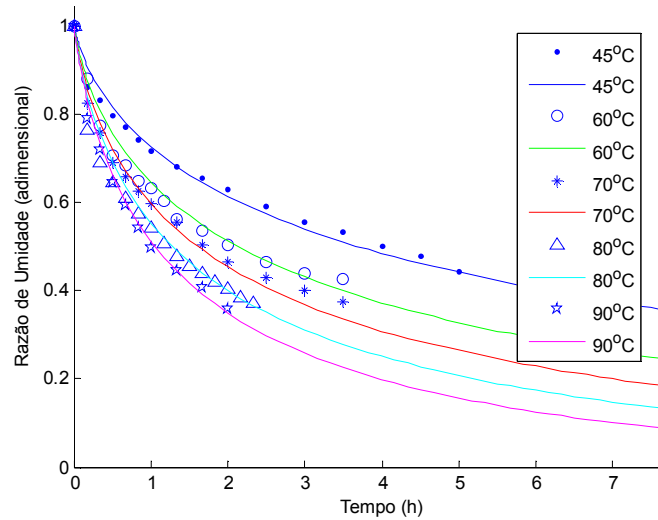


Figura 8: Influência da temperatura do ar.  $v = 0,9 \text{ m/s}$ ,  $X_0 = 0,19$

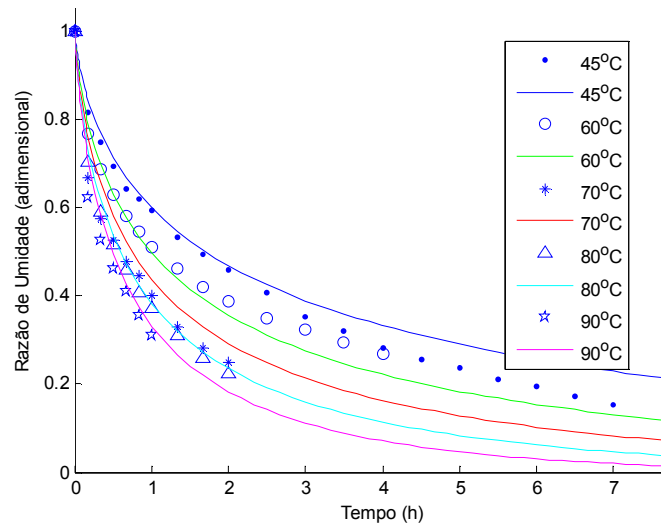


Figura 9: Influência da temperatura do ar.  $v = 1,5 \text{ m/s}$ ,  $X_0 = 0,28$

Observou-se que o maior tempo de secagem ocorreu a  $45^\circ\text{C}$ , e o menor a  $90^\circ\text{C}$ , havendo efeito decrescente no tempo de secagem à medida que ocorre o aumento na temperatura do ar. Isso é ocasionado devido ao aumento do potencial de secagem decorrente do aumento da temperatura do ar, o que faz com que haja diminuição da umidade relativa do ar de secagem.

### 7.3. Influência do teor de umidade inicial dos grãos

As curvas de secagem a seguir indicam uma variação de umidade mais rápida para teores de umidade mais elevados e secagem mais lenta para teores de umidade mais baixos, já que nos teores maiores existe um maior gradiente de umidade entre o



grão e o ar. Outro fator é que, quando o grão possui um alto teor de umidade, grande parte dessa água se concentra nos poros dos grãos, sendo mais fácil sua liberação. Já com baixos teores de umidade, a concentração de água nos poros é menor, sobrando a água que faz parte da composição das moléculas dos grãos, que é mais difícil de ser liberada.

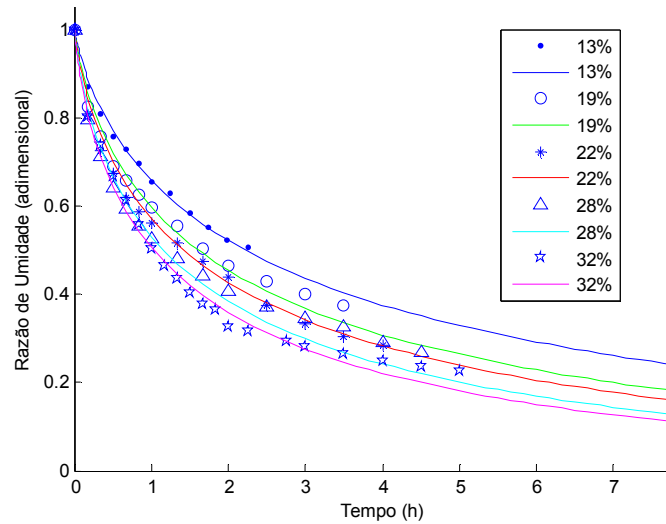


Figura 10: Influência do teor de umidade inicial do grão.  $v = 0,9 \text{ m/s}$ ,  $T_s = 70^\circ\text{C}$

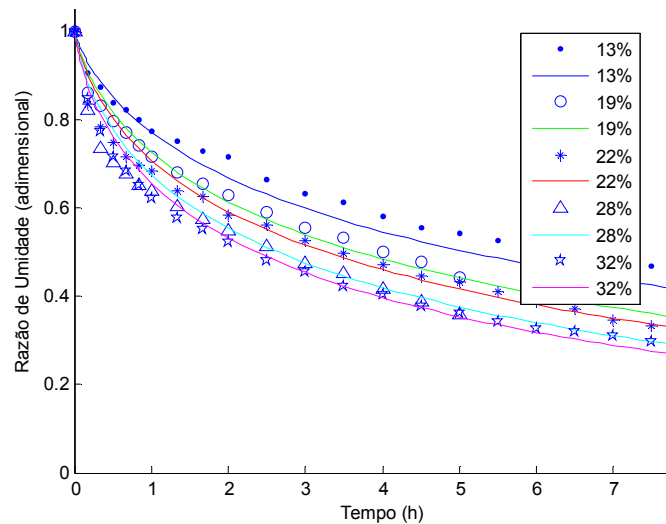


Figura 11: Influência do teor de umidade inicial do grão.  $v = 0,9 \text{ m/s}$ ,  $T_s = 45^\circ\text{C}$

## 8. Conclusões

A metodologia experimental empregada mostrou-se adequada para atingir os objetivos da pesquisa. Com este estudo, verifica-se que a secagem é significativamente

afetada pelos parâmetros velocidade e temperatura do ar de secagem e teor de umidade inicial dos grãos de soja.

Pode-se observar que a velocidade do ar de secagem apresenta relevante influência no processo, notando-se uma maior retirada de água durante as primeiras horas de secagem. Percebe-se que a dinâmica de secagem dos grãos de soja mostra uma clara dependência da velocidade da secagem com a temperatura do ar, característica presente na maioria dos produtos agrícolas. Nota-se que a influência da temperatura do ar na taxa de secagem é maior no princípio do experimento, diminuindo com o tempo de processamento.

Quanto ao modelo matemático utilizado nesta pesquisa, a alteração proposta mostrou-se adequada visto que verificou-se que a velocidade do ar de secagem é um parâmetro muito importante e que se faz indispensável dentro do modelo matemático, conforme mostraram as análises realizadas.

#### **Referências Bibliográficas:**

BROOKER, D.B., BAKKER-ARKEMA, F.W. and HALL, C.W. **Drying cereal grains**, The Avi Publishing Company, Westport, (1974).

COURTOIS, F., LASSERAN, J.C. **A CAD Software to improve the heat energy efficiency of mixed-flow maize dryers**, CESA/FAO Latin American Technical Meeting of Grain Drying a Storage, (1993), 18-22.

DALPASQUALE, V.A. **Drying of soybeans in continuous-flow dryers and fixed-bed drying systems**. Michigan: 1981. Tese (Ph.D), Michigan State University, (1981).

HALL, C.W. **Drying and storage of agricultural crops**. Westport: AVI Publishing Company, (1980).

KHATCHATOURIAN, O., BORGES, P.A.P. AND VIELMO, H.A. **Simulation of Soy Grains Drying Dynamics in Camera of Fixed Bed**. In: COBEM, International Congress of Mechanical Engineering, 17th. Anais. CD/ROM, (2003).

MINISTÉRIO DA AGRICULTURA E REFORMA AGRÁRIA, Brasil, Regras para análise de sementes. Brasília: SNDA/DNDV/CLAV, 365 p., (1992).

QUEIROZ, D.M. et al. **Modelos Matemáticos de Simulação de Secagem de Grãos**. Viçosa: Centreinar, (1985).

# ESTABILIDADE DO SISTEMA DE ENERGIA ELÉTRICA COM FACTS

Meire de Melo Marques, Laurence Duarte Colvara  
Universidade Estadual Paulista Júlio de Mesquita Filho-UNESP, memelmarques@hotmail.com;  
Universidade Estadual Paulista Júlio de Mesquita Filho-UNESP, laurence@dee.feis.unesp.br

## **Resumo:**

Esse trabalho apresenta um estudo sobre a influência de um dispositivo FACTS, o compensador estático de reativos (SVC), em um sistema multimáquinas tendo em vista a análise de estabilidade transitória. A função básica do sistema de controle de um SVC é controlar a tensão do sistema. O SVC gera e absorve potência reativa da rede no ponto de conexão. No trabalho será identificada a contribuição do SVC na composição do torque de sincronização. Para isto considera-se de início o sistema representado com a preservação da completa estrutura da rede de modo a se poder implementar a atuação do SVC. Na seqüência e para fins de análise se faz uma redução da rede conservando as barras internas dos geradores e a barra onde se localiza o SVC. Finalmente, a redução as barras internas proporciona uma quantificação dos efeitos da atuação do SVC nos torques de sincronização entre máquinas do sistema.

**Palavras-chave:** estabilidade transitória, SVC, sistema multimáquinas.

## **Introdução**

Os Sistemas de Energia Elétrica são projetados com o principal objetivo de atender à demanda de potência requerida pelos consumidores, com padrões de confiabilidade e qualidade garantidos. Isso significa operar dentro de limites especificados de tensão e frequência. Deve-se, para tanto, garantir flexibilidade de tais sistemas quanto à presença de faltas ou perturbações, sendo a prestação do serviço garantida, mesmo quando da ocorrência de anomalias.

Nesse trabalho, a inclusão do SVC em um sistema multimáquinas tem em vista a análise do ponto de vista da estabilidade transitória. Faz-se uma redução da rede,

conservando-se as barras internas dos geradores e a barra em que está inserido o dispositivo.

Segundo Hingorani (HINGORANI, 1993), uma alternativa plausível encontrada para o problema do aumento de demanda e frente à impossibilidade de investimentos na construção de novas linhas de transmissão e unidades geradoras seria a utilização de dispositivos FACTS. Os dispositivos em questão se utilizam de eletrônica de potência como forma de acionamento garantindo, com isso, uma operação com maior velocidade, mais continuidade, robustez, flexibilidade e confiabilidade.

Dentre os conversores mais utilizados está o SVC (Static Var Compensator) para o controle de tensão. Utiliza-se esse tipo de dispositivo desde o início dos anos 60. A precisão, a disponibilidade e a rápida resposta permitem um alto desempenho no controle de tensão em regimes permanente e transitório em relação à compensação *shunt* convencional. Também são usados no amortecimento de oscilações de potência, na melhoria da estabilidade transitória e na redução de perdas sistêmicas através da otimização do controle da potência reativa (HINGORANI & GYUGYI, 2000).

A melhoria da estabilidade é de grande importância em sistema de potência. O controle de excitação e o compensador estático de reativos (SVC) têm papel importante em melhorar a estabilidade de sistemas de potência. SVCs são crescentemente aplicados por várias razões pelas utilidades em sistemas de potência modernos com linhas de transmissão longas e fontes remotas de geração, uma vez que podem alcançar rapidamente a regulação precisa de tensão terminal. Instalações de SVC servem para melhorar a estabilidade transitória de sistema de potência que atingido por uma instabilidade (LERCH et al., 1991).

Para perturbações pequenas, um controle contínuo linear já se faz adequado para o amortecimento dessas oscilações. Porém, para perturbações maiores, uma ação mais drástica, a exemplo do controle de tipo de *bang-bang* (WANG et al., 2000).

O SVC é normalmente operado para regular a tensão da linha de transmissão no ponto onde o mesmo é conectado. Se o equipamento é instalado no meio da linha de transmissão e atua de tal forma que a tensão seja mantida constante em todos os níveis de fluxo de potência, então a linha é efetivamente dividida em duas partes iguais (reatância  $X/2$ ). Nessa situação a presença do SVC dobra o valor limite de potência estável do sistema. O ângulo total da linha de transmissão é maior para esse valor de potência. No entanto, a magnitude da tensão é mantida constante dinamicamente, e a estabilidade síncrona é mantida mesmo para ângulos superiores a  $\pi/2$ . Desse modo os

limites de tensão, colapso e limites de estabilidade são melhorados (SONG & JOHNS, 1999).

### Inclusão de um SVC em um Sistema Multimáquinas Análise da Estabilidade Transitória do Sistema

Seja um Sistema de Energia Elétrica em cuja rede de  $m$  nós tenha-se instalado um SVC e há uma linha conectando as barras genéricas  $k$  e  $\ell$ . Para representar essa instalação, cria-se o nó denominado  $n_{k\ell}$  como ilustrado na Figura 1.

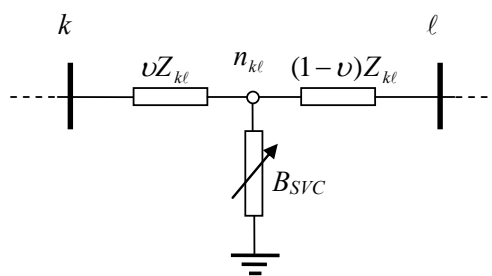


Figura 1 - SVC instalado em um ponto da linha  $k - \ell$ .

Indica-se a instalação no ponto central da linha. Com a inclusão do dispositivo, o nó  $n_{k\ell}$  é adicionado à rede, de modo que então a rede passa a ter  $m+1$  nós. A essa rede, adicionam-se as barras internas dos geradores à matriz de admitâncias. Com isso, obtém-se a dimensão  $n+m+1$ , sendo  $n$ , o número de geradores e  $m$  o número de barras da rede e a  $(n+m+1)$ -ésima barra  $l$  é justamente a barra do SVC, como mostra a Figura 2.

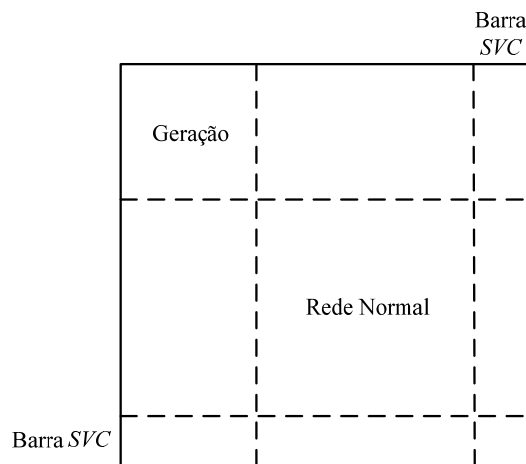


Figura 2 - Matriz de admitâncias adicionado o dispositivo SVC.

Faz-se então uma reordenação de barras, de modo que a  $(n+1)$ -ésima linha e coluna sejam as correspondentes ao nó  $n_{kl}$ . Então, a matriz da rede fica como ilustrado na Figura 3.

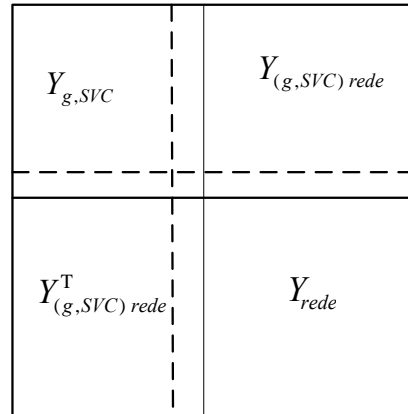


Figura 3 - Reordenação de linhas e colunas da matriz de admitâncias.

Definem-se, no âmbito desse estudo, as barras internas dos geradores e a barra do SVC, como barras de interesse, sendo o enfoque das análises de ora em diante dedicado a essas barras. O desempenho elétrico das demais barras deixa de interessar e a rede é então reduzida as barras de interesse.

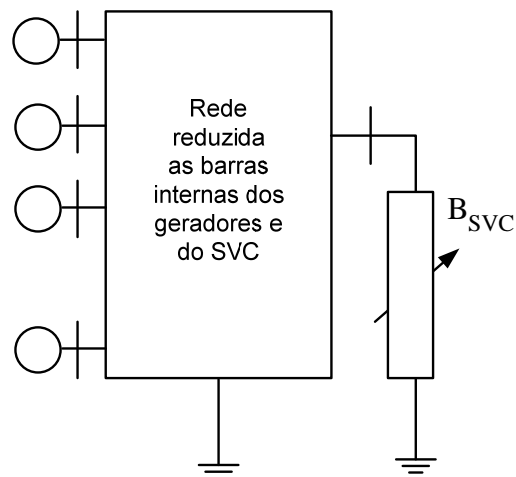


Figura 4 - Rede reduzida as barras internas dos geradores e do SVC.

Efetuando-se a redução da matriz de admitâncias as barras de interesse obtém-se o sistema reduzido conforme a relação (1).

$$\begin{bmatrix} I_g \\ I_r \end{bmatrix} = \begin{bmatrix} Y_{gg} & Y_{gr} \\ Y_{rg} & Y_{rr} \end{bmatrix} \begin{bmatrix} V_g \\ V_r \end{bmatrix} \quad (1)$$

As cargas representadas por admitâncias constantes são incorporadas à rede, de modo que a injeção de corrente nessas barras é nula. Então as equações podem ser reescritas conforme Anderson & Fouad (ANDERSON & FOUAD, 1977).

### ***Redução as barras internas***

Dadas as equações representadas de (2)-(5)

$$I_g = Y_{gg} V_g + Y_{gr} V_r \quad (2)$$

$$0 = Y_{rg} V_g + Y_{rr} V_r \Rightarrow -Y_{rr} V_r = Y_{rg} V_g \quad (3)$$

$$Y_{rr} V_r = -Y_{rg} V_g \quad (4)$$

$$V_r = -Y_{rr}^{-1} \cdot Y_{rg} \cdot V_g \quad (5)$$

Substituindo-se (5) na relação (2)

$$I_g = Y_{gg} V_g + Y_{gr} (-Y_{rr}^{-1} \cdot Y_{rg} V_g) \quad (6)$$

$$I_g = (Y_{gg} - Y_{gr} Y_{rr}^{-1} Y_{rg}) V_g \Rightarrow I_g = Y_{red} V_g \quad (7)$$

$$Y_{red} = Y_{gg} - Y_{gr} Y_{rr}^{-1} Y_{rg} \quad (8)$$

Obtém-se (9)

$$Y_{int} = Y_{g,SVC} - Y_{(g,SVC)rede} Y_{rede}^{-1} Y_{(g,SVC)rede}^T \quad (9)$$

e a matriz reduzida as barras de interesse tem a sua estrutura mostrada na Figura 5.

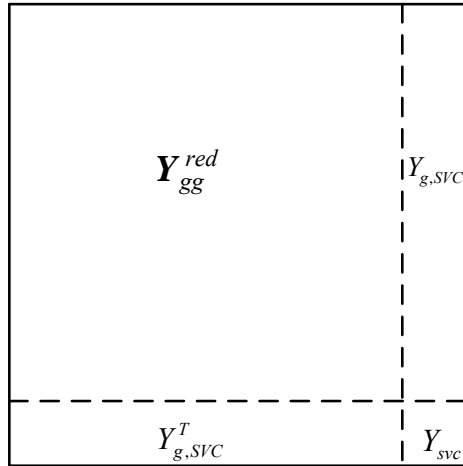


Figura 5 - Matriz de admitâncias reduzida.

A equação (9) representa a rede reduzida às barras de interesse e pode ser representada conforme a relação (10).

$$\begin{bmatrix} \mathbf{I}_g \\ \mathbf{I}_{SVC} \end{bmatrix} = \begin{bmatrix} \mathbf{Y}_{gg}^{red} & \mathbf{Y}_{g,SVC} \\ \mathbf{Y}_{g,SVC}^T & \mathbf{Y}_{SVC} \end{bmatrix} \begin{bmatrix} \mathbf{E} \\ \mathbf{V}_{SVC} \end{bmatrix} \quad (10)$$

Nota-se que o termo  $\mathbf{Y}_{g,SVC}$  é um vetor de dimensão  $n$  (número de máquinas) e seu  $i$ -ésimo elemento genérico  $Y_{g,SVC_i}$  podendo ser entendido como a sensibilidade da corrente terminal do  $i$ -ésimo gerador às variações da tensão sobre o SVC.

Considerando-se que o objetivo final do estudo é a introdução do método da RSP de análise de estabilidade transitória a influência do SVC, que trata a questão da estabilidade máquina-versus-máquina. É necessário para tanto a obtenção da matriz da rede reduzida as barras internas dos geradores. A matriz reduzida final, relacionando exclusivamente as barras internas dos geradores é obtida partindo-se de (11).

$$\mathbf{Y}_{svc}^{-1} = \mathbf{Z}_{svc} = (\mathbf{Z}_{svc}^0 + \Delta\mathbf{Z}_{svc}) \quad (11)$$

Em (11), o termo  $\mathbf{Y}_{SVC}^0$  ( $= \mathbf{Z}_{SVC}^0^{-1}$ ) representa a admitância equivalente do SVC na condição de operação normal (equilíbrio) e o termo  $\Delta\mathbf{Z}_{SVC}$ , correspondente à



variação (em série) da impedância para alteração da admitância do SVC (considerada em paralelo). A matriz reduzida as barras internas dos geradores é expressa por (12).

$$\mathbf{Y}_{\text{redf}} = \mathbf{Y}_{\text{red}}^0 + \Delta \mathbf{Y}_{\text{svc}}^* \quad (12)$$

sendo

$$\mathbf{Y}_{\text{red}}^0 = [\mathbf{Y}_{\text{gg}}^{\text{red}} + \Delta \mathbf{Y}_{\text{svc}}^0] \quad (13)$$

com

$$\Delta \mathbf{Y}_{\text{svc}}^0 = -\mathbf{Z}_{\text{svc}}^0 [\mathbf{Y}_{\text{g,svc}} \mathbf{Y}_{\text{g,svc}}^T] \quad (14)$$

Destaca-se que o termo  $\mathbf{Y}_{\text{red}}^0$  é a matriz admitância de barra da rede reduzida as barras internas dos geradores, incorporando o SVC na situação de regime permanente. Pode-se dizer que é a matriz reduzida considerando a compensação estática (fixa ou constante). Mas, se além da compensação paralela constante, tem-se uma parcela da admitância (ou impedância) variante segundo uma lei de controle (não considerada nesse momento) tem-se adicionalmente uma variação da matriz de admitâncias reduzida expressa por:

$$\Delta \mathbf{Y}_{\text{svc}}^* = -\Delta \mathbf{Z}_{\text{svc}} [\mathbf{Y}_{\text{g,svc}} \mathbf{Y}_{\text{g,svc}}^T] \quad (15)$$

Note-se que cada elemento  $\Delta Y_{\text{svc}ij}^*$  de  $\Delta \mathbf{Y}_{\text{svc}}^*$  é a variação da admitância de transferência entre as máquinas  $i$  e  $j$  e é dada por:

$$\Delta Y_{\text{svc}ij}^* = -[Y_{\text{g,svc}i} Y_{\text{g,svc}j}] \Delta Z_{\text{svc}}$$

Sendo  $Y_{\text{g,svc}i} Y_{\text{g,svc}j} = \text{cte}, \forall i, j$  têm-se que cada admitância de transferência entre máquinas do sistema depende diretamente das variações da impedância do SVC, por intermédio do coeficiente:

$$K_{\text{svc}ij} = -Y_{\text{g,svc}i} Y_{\text{g,svc}j} \quad (16)$$

O coeficiente  $K_{svc\ ij}$  pode ser identificado como o fator de sensibilidade da capacidade de sincronização do par de máquinas ( $i,j$ ) em relação às variações da impedância do SVC, como se descreve a seguir.

Desconsiderada a condutância de transferência e tomando, a título de ilustração, a potência transferida entre as máquinas  $i$  e  $j$ , na sua forma mais simples, como:

$$P_{eij} = E_i B_{ij} E_j \text{sen}(\delta_{ij}) \quad (17)$$

têm-se, com atuação de um SVC instalado em alguma barra da rede (não necessariamente em  $i$  ou  $j$ ) a potência transferida dada por:

$$P_{eij} = E_i E_j (B_{ij}^0 + \Delta B_{ij}) \text{sen}(\delta_{ij}) \quad (18)$$

sendo  $\Delta B_{ij}$  dependente diretamente das variações da impedância do SVC. A princípio, tal dependência pode ensejar o controle do fluxo de potência entre máquinas e, mediante uma lei de controle a ser definida, atuar sobre o desempenho transitório do sistema. Mas, nesse momento, isto é apenas uma conjectura prospectiva, e no restante desse estudo se fará considerações no sentido de analisar o desempenho do sistema com o SVC e não no de sintetizar um controlador.

### O Método da RSP para Sistemas com SVC

A Região de Sincronização Positiva é definida como a região em torno de um ponto de operação do sistema onde os torques no eixo da máquina síncrona (MS) são forças restaurativas do equilíbrio e é dado por  $f(\sigma, e)$  em que  $f(\sigma, e) = BV[(E^0 + e)\text{sen}(\sigma + \delta^0) - E^0 \text{sen}\delta^0]$ . Fora dessa região, nenhuma ação do sistema livre leva o sistema ao equilíbrio, sendo então esse conceito usado para definir a instabilidade do SEE (COLVARA, 1988).

Considerando as máquinas descritas pelo modelo de dois eixos, com atuação dos Reguladores Automáticos de Tensão, os desvios da potência (ou torque) líquida(o) transmitida entre as máquinas do par  $k=(i,j)$  através da rede reduzida as barras internas é expressa pela função.

$$f_k(\sigma_k, e_i, e_j) = Y_{ij} [E'_i E'_j \text{sen}(\sigma_k + \delta_{ij}^0 - \alpha_{ij} - \beta_i + \beta_j) - E_i^0 E_j^0 \text{sen}(\delta_{ij}^0 - \alpha_{ij} - \beta_i + \beta_j)] \quad (19)$$

em que 
$$\sigma_k = (\delta_{ij} + \alpha_{ij}) - (\delta_{ij}^0 + \alpha_{ij}) = \delta_{ij} - \delta_{ij}^0 \quad (20)$$

As tensões internas  $E'_i$  e  $E'_j$  bem como os ângulos  $\delta, \beta$  e  $\alpha$  são como mostrados nas Figuras 6 e 7, respectivamente.

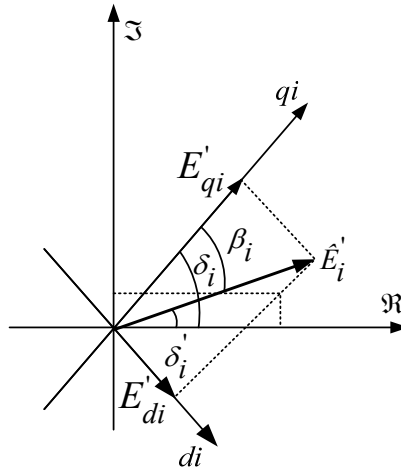


Figura 6 - Diagrama fasorial para a máquina  $i$ .

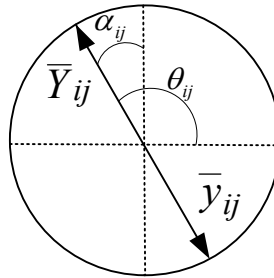


Figura 7 - Ângulos das admitâncias próprias dos elementos da rede ( $y_{ij}$ ) e das admitâncias de transferência ( $\theta_{ij}$ ).

$$\bar{Y}_{ij} = Y_{ij} \angle \theta_{ij} = -\bar{y}_{ij} \quad (21)$$

$$\theta_{ij} = \frac{\pi}{2} + \alpha_{ij} \quad (22)$$

Introduzindo a variação de  $Y_{ij}$  devida ao SVC, vem

$$\begin{aligned}\bar{Y}_{ij} &= \bar{Y}_{ij}^0 + \bar{K}_{SVCij} \Delta \bar{Z}_{SVC} \\ &= (Y_{ij}^0 + \Delta Y_{ij}) \angle \theta_{ij}^0 + \Delta \theta_{ij}\end{aligned}\quad (23)$$

Observando a Figura 7 nota-se que se pode escrever  $\Delta \alpha_{ij} = \Delta \theta_{ij}$  e a equação do torque mútuo pode ser reescrita como

$$\begin{aligned}f_k(\sigma_k, e_i, e_j) &= \\ &= (Y_{ij}^0 + \Delta Y_{ij}) E_i' E_j' \text{sen}(\sigma_k + \delta_{ij}^0 - \alpha_{ij}^0 - \Delta \alpha_{ij} - \beta_i + \beta_j) - E_i^0 E_j^0 \text{sen}(\delta_{ij}^0 - \alpha_{ij}^0 - \beta_i^0 + \beta_j^0)\end{aligned}\quad (24)$$

Então a inserção do SVC modifica a RSP pela alteração de suas fronteiras em função das variações da admitância de transferência. Essas modificações devem ser levadas em conta nos cálculos relativos ao monitoramento da trajetória do sistema em relação às fronteiras da RSP.

### Sistema de 4 máquinas

O sistema exemplo considerado é encontrado em (KUNDUR,1994) e é aqui apresentado por meio de seu diagrama unifilar e resultados de fluxo de potência do caso base.

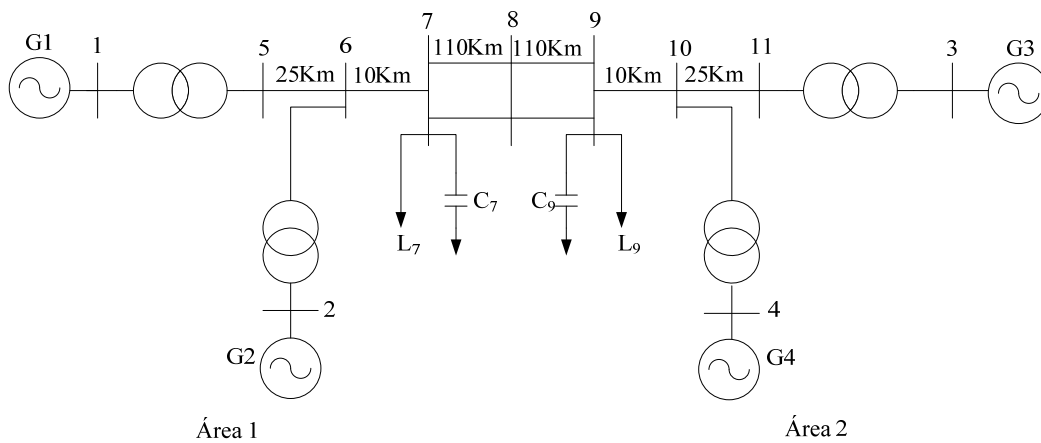


Figura 8 - Sistema simples de duas áreas.

## FLUXO DE POTÊNCIA

**Nº de Barras**    **Nº de Linhas**    **Nº Máx de Interações**    **Tolerância**  
 12                      11                      20                      .0010

Tabela 1 - Dados de Linha.

Linha	Início	Término	$Z_{serie}$		$Y_{shunt}$	
1	1	5	.0000	.1500	.0000	.0000
2	2	6	.0000	.1500	.0000	.0000
3	3	11	.0000	.1500	.0000	.0000
4	4	10	.0000	.1500	.0000	.0000
5	5	6	.0225	.2250	.0000	.0049
6	6	7	.0090	.0900	.0000	.0019
7	7	8	.0000	.1400	.0000	.0214
8	8	12	.0495	.4250	.0000	.0107
9	12	9	.0495	.4250	.0000	.0107
10	9	10	.0099	.0990	.0000	.0019
11	10	11	.0225	.2250	.0000	.0049

Tabela 2 - Especificação de Potência e de Tensão.

Barra	Nome	Tipo	Potência nas Barras		Tensão de Barra	
			Ativa	Reativa	Módulo	Arg(GRA)
1	Barra 1	3	Não Esp.	Não Esp.	1.0300	.0000
2	Barra 2	2	.7778	.0000	1.0100	.0000
3	Barra 3	2	.7989	.0000	1.0300	.0000
4	Barra 4	2	.7778	.0000	1.0100	.0000
5	Barra 5	1	.0000	.0000	.0000	.0000
6	Barra 6	1	.0000	.0000	.0000	.0000
7	Barra 7	1	-1.0744	.1111	.0000	.0000
8	Barra 8	1	.0000	.0000	.0000	.0000
9	Barra 9	1	-1.9633	.2778	.0000	.0000
10	Barra 10	1	.0000	.0000	.0000	.0000
11	Barra 11	1	.0000	.0000	.0000	.0000
12	Barra SVC	1	.0000	.1444	.0000	.0000

Barras:            TIPO 1 – Barra (PQ)  
                       TIPO 2 – Barra (PV)  
                       TIPO 3 – Barra de Folga

3. O sistema convergiu com 11 iterações e os resultados estão ilustrados na Tabela

Tabela 3 - Resultados Obtidos.

Barra	Nome	Tensão		Módulo	Arg	Potências nas Barras	
						Ativa	Reativa
1	Barra 1	1.0300	0.0000	1.0300	0.0000	.7771	.1802
2	Barra 2	.9956	-1.700	1.0100	-9.6891	.7778	.1999
3	Barra 3	.9380	-4.255	1.0300	-24.3998	.7989	.1731
4	Barra 4	.8320	-0.5726	1.0100	-34.5342	.7778	.1706
5	Barra 5	1.0038	-0.1132	1.0101	-6.4323	.0000	.0000
6	Barra 6	.9469	-0.2789	.9871	-16.4098	.0000	.0000
7	Barra 7	.8881	-0.4074	.9771	-24.6442	-1.0744	.1111
8	Barra 8	.8656	-0.4673	.9837	-28.3608	.0000	.0000
9	Barra 9	.6297	-0.7590	.9862	-50.3191	-1.9633	.2778
10	Barra 10	.7457	-0.6534	.9914	-41.2258	.0000	.0000
11	Barra 11	.8670	-0.5210	1.0115	-31.0049	.0000	.0000
12	Barra SVC	.7706	-0.6371	.9999	-39.5811	.0000	.1444

Tabela 4 - Dados de Máquinas.

Maquina	Barra	XD	Xld	M	D	Tld0	Tlq0	KR	TR
1	1	1.800	0.300	0.345	.000	8.0	0.40	25.00	0.10
2	2	1.800	0.300	0.345	.000	8.0	0.40	25.00	0.10
3	3	1.800	0.300	0.328	.000	8.0	0.40	25.00	0.10
4	4	1.800	0.300	0.328	.000	8.0	0.40	25.00	0.10

Maquina	Barra	Efdmax	Efdmin	Tg	Tm	Ag	Bg	Alf	Bet
1	1	5.00	-5.00	1.00	1.00	0.000	0.000	0.000	0.000
2	2	5.00	-5.00	1.00	1.00	0.000	0.000	0.000	0.000
3	3	5.00	-5.00	1.00	1.00	0.000	0.000	0.000	0.000
4	4	5.00	-5.00	1.00	1.00	0.000	0.000	0.000	0.000

$$Y = Y_{red}^0 + K_{SVC} \Delta Z_{SVC}$$

$$Y = \begin{bmatrix} 0.1467 - 1.0093i & 0.1696 + 0.7009i & 0.0590 + 0.0605i & 0.0916 + 0.0878i \\ 0.1696 + 0.7009i & 0.2894 - 1.1794i & 0.0916 + 0.0878i & 0.1418 + 0.1272i \\ 0.0590 + 0.0605i & 0.0916 + 0.0878i & 0.2023 - 1.0544i & 0.2508 + 0.6305i \\ 0.0916 + 0.0878i & 0.1418 + 0.1272i & 0.2508 + 0.6305i & 0.4077 - 1.2891i \end{bmatrix} +$$

$$+ \begin{bmatrix} 0.3303 - 0.0961i & 0.3482 - 0.1033i & 0.2999 - 0.1466i & 0.3157 - 0.1564i \\ 0.3482 - 0.1033i & 0.3669 - 0.1109i & 0.3157 - 0.1564i & 0.3324 - 0.1669i \\ 0.2999 - 0.1466i & 0.3157 - 0.1564i & 0.2624 - 0.1899i & 0.2759 - 0.2019i \\ 0.3157 - 0.1564i & 0.3324 - 0.1669i & 0.2759 - 0.2019i & 0.2901 - 0.2146i \end{bmatrix} \Delta Z_{SVC}$$

$$Y_{red}^0 = \begin{bmatrix} 1.2824 \angle -81.45^\circ & 0.9036 \angle 77.59^\circ & 0.1323 \angle 49.21^\circ & 0.1396 \angle 48.91^\circ \\ 0.9036 \angle 77.59^\circ & 1.3086 \angle -80.77^\circ & 0.1396 \angle 48.90^\circ & 0.1474 \angle 48.60^\circ \\ 0.1323 \angle 49.21^\circ & 0.1396 \angle 48.90^\circ & 1.3683 \angle -77.88^\circ & 0.8609 \angle 69.90^\circ \\ 0.1396 \angle 48.91^\circ & 0.1474 \angle 48.60^\circ & 0.8609 \angle 69.90^\circ & 1.4066 \angle -76.98^\circ \end{bmatrix}$$

$$K_{SVC} = \begin{bmatrix} 0.3440 \angle -16.22^\circ & 0.3631 \angle -16.52^\circ & 0.3338 \angle -26.05^\circ & 0.3524 \angle -26.36^\circ \\ 0.3631 \angle -16.52^\circ & 0.3833 \angle -16.82^\circ & 0.3524 \angle -26.35^\circ & 0.3719 \angle -26.66^\circ \\ 0.3338 \angle -26.05^\circ & 0.3524 \angle -26.35^\circ & 0.3239 \angle -35.89^\circ & 0.3419 \angle -36.19^\circ \\ 0.3524 \angle -26.36^\circ & 0.3719 \angle -26.66^\circ & 0.3419 \angle -36.19^\circ & 0.3609 \angle -36.49^\circ \end{bmatrix}$$

### Simulações:

Falta trifásica na barra 009, eliminada por desconexão da linha 008-009, com duração de  $t_{ch} = 0,074s$ .

### 1. Sem SVC

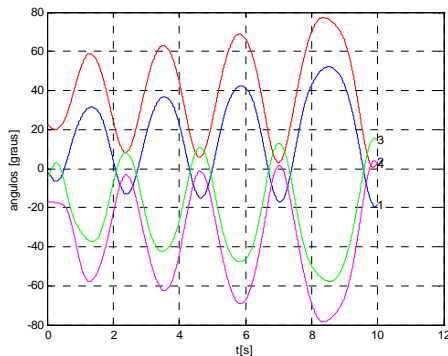


Figura 9 - Ângulos das máquinas (referência centro inercial) sem SVC.

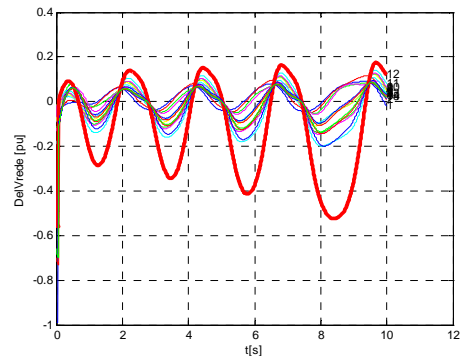


Figura 10 - Tensões de barra (em destaque a tensão da barra 8) sem SVC.

## 2. Com SVC instalado na barra 008, com controle proporcional de tensão

$$\Delta B_{SVC} = K_{SVC} \Delta V$$
$$K_{SVC} = 4,5$$

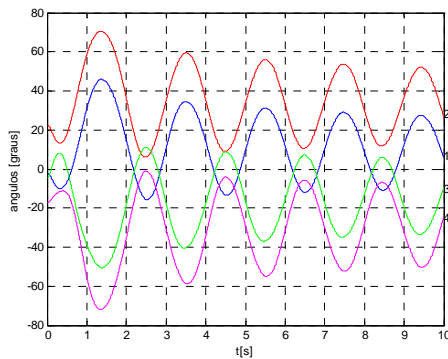


Figura 11 - Ângulos das máquinas (referência centro inercial) com SVC instalado na barra 008.

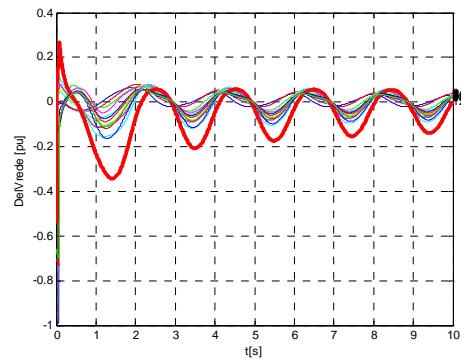


Figura 12 - Tensões de barra (em destaque a tensão da barra 8) com SVC instalado na barra 008.

### Conclusão

Desenvolveu-se uma formulação para a consideração do dispositivo FACTS SVC em estudo de estabilidade transitória de sistema de energia elétrica multimáquinas, em que se explicita a contribuição das variações da impedância do dispositivo na admitância de transferência entre máquinas do sistema e conseqüentemente na potência sincronizante entre elas.

Alguns testes prospectivos foram realizados em um sistema de 4 máquinas visando observar o comportamento do fator de contribuição das variações da impedância do SVC a cada par de máquinas do sistema, ou, visto pelo outro lado, o fator de sensibilidade de cada par de máquinas às variações da impedância do SVC.

A formulação da admitância de transferência – e, portanto a capacidade de sincronização entre máquinas do sistema – dependente diretamente das variações da impedância do SVC parece abrir um caminho para estabelecer controles eficazes diante de perturbações transitórias do sistema via uma adequada lei de controle, bem como considerar o dispositivo em procedimento de avaliação da estabilidade por método automático.



Esta última alternativa será enfocada nos próximos passos do desenvolvimento deste trabalho, pela inclusão dos efeitos do SVC na metodologia de análise de estabilidade baseada no conceito da região de sincronização positiva, uma vez que se identifica com relativa facilidade os efeitos do dispositivo sobre as fronteiras da região, além, evidentemente, de sua influência sobre a trajetória do sistema.

### **Referências Bibliográficas:**

ANDERSON, P.M. and FOUAD, A.A.: **Power system control and stability**, The Iowa State University Press, (1977).

COLVARA, L.D.: **Análise de estabilidade transitória de sistemas de energia elétrica com reguladores automáticos de tensão usando funções de lyapunov**, Tese Faculdade de Engenharia Elétrica da Universidade do Rio de Janeiro, RJ, (1988).

HINGORANI, N.: **Flexible AC transmission IEEE spectrum, april**, (1993), 40-45.

HINGORANI, N. and GYUGYI, L.: **Understanding FACTS concepts and technology of flexible AC transmission systems**, 1. ed. The institute of electrical and electronics engineers. Inc New York, (2000).

KUNDUR.: **P. Power system stability and control**. New York: Mc Graw-Hill, (1994).

LERCH, E., POVH, D. and XU, L.: **Advanced SVC control for damping power system oscillations**, IEEE Transactions Power Systems; (1991), PWRS-6:524-31.

SONG, Y.H. and JOHNS, A.T.: **Flexible ac transmission systems (FACTS) IEE, UK October**, (1999).

WANG, Y., TAN, Y.L. and GUO, G.: **Robust nonlinear coordinated generator excitation and SVC control for power systems**. Electrical Power & Energy Systems vol. 22, (2000), p187-195.

# MODELAGEM E SIMULAÇÃO COMPUTACIONAL DE UM ATUADOR HIDRÁULICO CONSIDERANDO O EFEITO DO ATRITO DINÂMICO

**Fabiane Eloisa Morandini Miotto, Antonio Carlos Valdiero,  
Vanessa Dilda, Pedro Luís Andrighetto**

Universidade Regional do Noroeste do Estado do RS - UNIJUI,  
Departamento de Tecnologia, Campus Panambi  
fabianemiotto@gmail.com, valdiero@unijui.edu.br,  
vanedilda@yahoo.com.br, pedro@unijui.edu.br

## **Resumo:**

Este trabalho apresenta a modelagem matemática e a simulação computacional dos efeitos do atrito em um atuador hidráulico. O atrito é uma das principais não linearidades que perturbam o controle de atuadores hidráulicos, podendo inclusive causar instabilidade e degradação do desempenho do sistema que é descrito através de um modelo matemático não linear de 5ª ordem. O modelo dinâmico é simulado computacionalmente por meio da implementação do diagrama de blocos no MatLab/Simulink. Os resultados das simulações ilustram as características do atrito e são importantes para definição das estratégias de controle preciso, compensação de atrito e caracterização e seleção dos atuadores hidráulicos.

**Palavras-chave:** atuador hidráulico, atrito, simulação computacional.

## **1 - Introdução:**

Neste trabalho apresenta-se a simulação computacional de um modelo matemático para atuadores hidráulicos considerando o atrito dinâmico.

Ainda no século passado, Merrit (1967) já apontava a importância e as vantagens da aplicação de sistemas hidráulicos, tais como a excelente relação torque/dimensão e a resposta rápida aos comandos de partidas, paradas ou inversões de velocidade sem danos às partes mecânicas.

De acordo com Linsingen (2003), em função dos esforços de desenvolvimento tecnológico, os sistemas hidráulicos encontram aplicação em praticamente todos os ramos de atividade, da extração mineral à indústria aeroespacial, bem como em aplicações de uso cotidiano, como em veículos de transporte e passeio, equipamentos odontológicos e médico-hospitalares, construção civil, etc.

Christensen et al. (2000) compara as diferentes tecnologias de transmissão de potência e mostra que os sistemas hidráulicos são competitivos nas aplicações com potência ou forças altas e onde são necessários atuadores relativamente pequenos com flexibilidade de instalação.

Atuadores hidráulicos estão entre os sistemas de atuação mais utilizados para uma variedade de aplicações da geração de força e posicionamento. Porém, exibem não linearidades significantes em suas dinâmicas.

O atrito é uma das principais não linearidades que perturbam o controle de atuadores hidráulicos. Nestes, o atrito ocorre principalmente entre as superfícies de contato nas vedações da haste com o cilindro, mas também nas paredes do cilindro com o êmbolo.

Canudas-de-Wit e Lischinski (1997) afirmam que o atrito ocorre em todas as máquinas que incorporam peças com movimento relativo, causando erros típicos de regime permanente em controle de posição e atrasos no seguimento podendo inclusive causar instabilidade. As características dinâmicas do atrito também causam degradação do desempenho do sistema e precisam ser previstas ou compensadas adequadamente para reduzir seus efeitos. Uma pesquisa na bibliografia recente demonstra que o atrito tem sido objeto de constante investigação por parte da comunidade científica internacional. Aho et al. (2007) comenta que em controle de mecanismos, o atrito tem sido reconhecido como fonte para gerar erros, ciclos limite, e movimentos stick-slip, entre outros fenômenos de atrito, que degradam o desempenho do mecanismo controlado. No controle de mecanismos de alta-precisão, atrito é o principal obstáculo a ser enfrentado porque é difícil identificá-lo e não há nenhum modelo universal para descrevê-lo. Marton et al. (2007) apresenta uma aproximação moderna para modelagem de atrito stick-slip. O modelo distingue claramente os regimes de alta e baixa velocidade. Um método de identificação não linear para os parâmetros modelo de atrito também é apresentado. Cornejo (2007) introduz uma modificação para o modelo de atrito dinâmico LuGre. Dumitriu (2007) menciona que o atrito é a maior causa do erro de posição e justifica a necessidade de ser bem modelado e compensado para o controle preciso de servo sistemas.

As características dinâmicas do atrito são a causa da degradação do desempenho do sistema e precisam ser previstas ou compensadas adequadamente para reduzir seus efeitos.

O modelo de atrito adotado neste estudo é baseado no modelo de LuGre, conforme Valdiero (2005) e é escrito na forma de equações diferenciais.

A próxima seção deste trabalho traz uma descrição do atuador hidráulico e seu funcionamento. A seção 3 mostra características do atrito e efeitos de degradação causados por ele. Na seção 4 é apresentada a modelagem matemática de um atuador hidráulico, considerando o atrito. Os resultados de simulação são mostrados na seção 5. Por fim, apresentam-se as conclusões.

## 2 – Descrição do atuador hidráulico:

O sistema de atuador hidráulico, adotado na modelagem do atrito, é formado pelo conjunto de uma válvula direcional proporcional tipo carretel de 4 vias e um cilindro hidráulico, de haste dupla, conforme mostrado na Figura 1, destacando-se as regiões com maior presença de atrito.

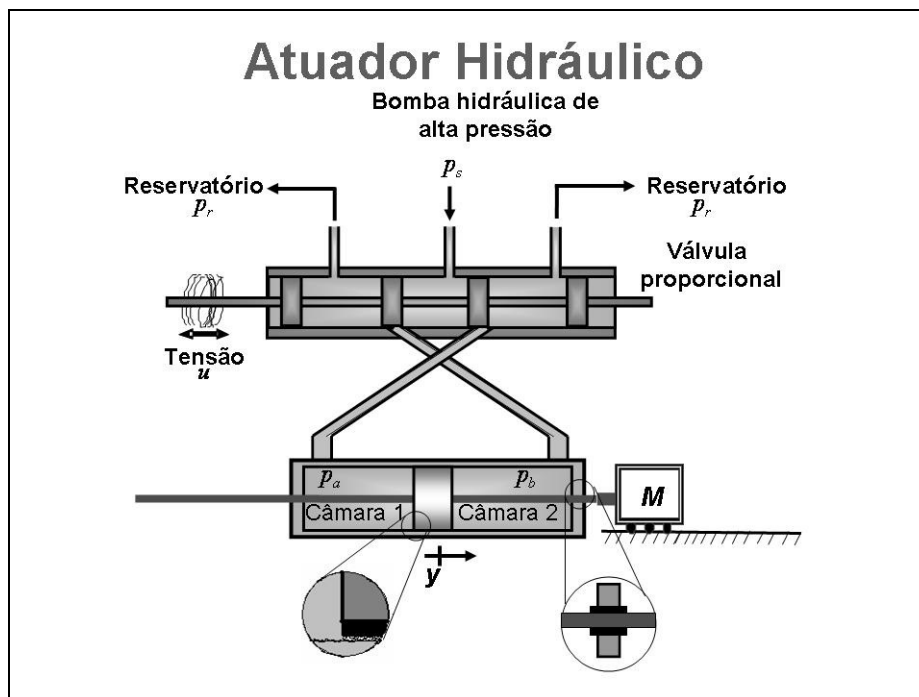


Figura 1 - Desenho Esquemático de um atuador hidráulico

Durante o seu funcionamento, o fluido é fornecido à válvula por uma unidade de potência e condicionamento hidráulico (UPCH) com uma pressão de suprimento  $p_s$ . Dado

um sinal elétrico de controle  $u$ , que energiza as bobinas dos solenóides proporcionais da válvula, ocorre um deslocamento do carretel da válvula. Por sua vez, o carretel, ao ser deslocado, gera orifícios de passagem, fornecendo fluido a alta pressão para uma das câmaras do cilindro provocando o movimento (e a produção de atrito) da haste do cilindro e conseqüentemente, o escoamento do fluido da outra câmara do cilindro para o reservatório que está a uma pressão  $p_r$ . Desta forma, obtém-se a variação das pressões  $p_a$  e  $p_b$  nas câmaras do cilindro, resultando numa força que movimenta a massa  $M$  num deslocamento  $y$ .

Conforme Dilda et al. (2007), a força gerada pelo atuador hidráulico é obtida através do produto da área da seção transversal do êmbolo do cilindro pela diferença de pressão. Sendo assim, é possível gerar grandes forças com atuadores de pequena dimensão, utilizando, para isso valores elevados de pressão.

### 3 – Características do atrito e efeitos de degradação causados por ele:

O trabalho de Valdiero (2005) apresenta as características do atrito de forma muito clara, definindo-o como um fenômeno não linear multifacetado que exibe diversas características não lineares. Tais características são compostas pelos bem conhecidos e clássicos atrito estático, atrito de Coulomb, atrito viscoso e de arraste, os quais compõem os modelos mais simples baseados em mapas estáticos; mas também são compostas por fenômenos dinâmicos mais complexos, conhecidos como atrito de Stribeck, atrito estático crescente, memória de atrito e deslocamento de predeslizamento.

As características dinâmicas do atrito são responsáveis por degradações no desempenho do sistema e necessitam serem previstas ou observadas para uma adequada compensação e conseqüente diminuição de seus efeitos. A seguir, são apresentadas cada uma destas características:

O atrito estático ocorre na velocidade zero, opõe-se a todo movimento com a mesma magnitude da força (ou torque) aplicada  $u(t)$  até um valor máximo de força de atrito estático  $F_s$  e pode ser descrito como:

$$F_{estático}(t) = \begin{cases} u(t), & se \ |u(t)| < F_s \\ F_s \delta(\dot{y}(t)) \text{sgn}(u(t)), & se \ |u(t)| \geq F_s \end{cases} \quad (1)$$

onde  $\dot{y}(t)$  é a velocidade e  $\delta(\dot{y}(t))$  é a função impulso, utilizada para descrever o fato de que o atrito estático ocorre apenas no repouso, é dada por:

$$\delta(\dot{y}(t)) = \begin{cases} 1, & \text{se } \dot{y}(t) = 0 \\ 0, & \text{se } \dot{y}(t) \neq 0 \end{cases} \quad (2)$$

O atrito de Coulomb é independente da área de contato, opõem-se ao movimento relativo e é proporcional à força normal de contato, podendo ser escrito como:

$$F_{Coulomb}(t) = F_c \operatorname{sgn}(\dot{y}(t)) \text{ quando } \dot{y}(t) \neq 0 \quad (3)$$

considerando que  $F_c$  é a magnitude do atrito de Coulomb, a qual é independente da magnitude da velocidade relativa  $\dot{y}(t)$ .

O atrito viscoso corresponde a uma situação de boa lubrificação e é linearmente proporcional a velocidade, conforme descrito pela equação (4):

$$F_{viscoso}(t) = B \dot{y}(t) \quad (4)$$

onde  $B$  é chamado coeficiente de amortecimento viscoso.

O atrito de arraste refere-se ao atrito causado pela resistência ao movimento de um corpo através de um fluido, sendo proporcional ao quadrado da velocidade e muitas vezes decorrente de um escoamento turbulento. Ele pode ser escrito como:

$$F_{arraste}(t) = F_D (\dot{y}(t))^2 \operatorname{sgn}(\dot{y}(t)) \quad (5)$$

sendo que  $F_D$  é o coeficiente de arraste. Para baixas velocidades, o valor do atrito de arraste torna-se pequeno e pode ser desprezado.

Atrito de Stribeck é um fenômeno não linear de atrito que ocorre nos trechos de baixa velocidade da curva atrito versus velocidade, onde a inclinação é negativa. Tem importante contribuição para o efeito de atrito conhecido com adere-desliza (*stick-slip*).

A característica chamada de memória de atrito é um atraso observado entre as mudanças na velocidade (ou carga normal) e a correspondente mudança na força de atrito.

A característica de deslocamento de predeslizamento é o deslocamento que ocorre antes do deslizamento entre as superfícies em contato, sendo algumas vezes chamado de microdeformação. Ele surge devido a deformação elástica e/ou plástica das rugosidades das superfícies de contato.

Vários artigos atuais (Acho (2007), Marton (2007), Wenjing (2007)) apresentam efeitos de degradação do desempenho do movimento causados pelas características não

lineares do atrito. Os efeitos do atrito dependem, entre outros, do estado de lubrificação, da temperatura, da direção do movimento, da velocidade do êmbolo, das pressões existentes nas câmaras do cilindro, e, geralmente alteram-se com o tempo.

Alguns destes efeitos são chamados de adere-desliza (*stick-slip*), *hunting*, perda de movimento (*standstill*) e *quadrature glitch*.

Conforme Valdiero (2005), o termo adere-desliza (*stick-slip*) refere-se a uma alternância entre o movimento de deslizamento e o repouso, sendo definido como um ciclo limite estável surgido durante o movimento e que é mais comum quando é utilizado o controle integral. O fenômeno *hunting* refere-se também a um ciclo limite associado ao controle integral, mas que ocorre quando o movimento oscila em torno de uma dada posição desejada de valor constante, acontece com controle de realimentação e não é possível em sistemas passivos. A perda de movimento (*standstill*) refere-se ao efeito de atrito que ocorre quando o sistema é detido no repouso por um intervalo de tempo ao passar pela velocidade nula. E o termo *quadrature glitch* refere-se ao erro de seguimento num movimento de múltiplos eixos.

#### 4- Modelagem Matemática de um atuador hidráulico, considerando o atrito:

Neste trabalho propõe-se inicialmente a utilização do modelo não linear de 5ª ordem (sem a dinâmica elétrica da válvula, por ser considerada muito rápida) através da combinação das equações, descritas pelas equações (6), (7), (8), (9) e (10), o qual resulta no sistema de equações diferenciais dadas por:

$$M\ddot{y} + F_{atr} = A(p_a - p_b) \quad (6)$$

$$\dot{p}_a = \beta \cdot f_1(y) \cdot (Q_a(x_v, p_a) - A \cdot \dot{y}) \quad (7)$$

$$\dot{p}_b = \beta \cdot f_2(y) \cdot (Q_b(x_v, p_b) + A \cdot \dot{y}) \quad (8)$$

$$F_{atr} = \sigma_0 z + \sigma_1 \dot{z} + \sigma_2 \dot{y} \quad (9)$$

$$\dot{z} = \dot{y} - \frac{\sigma_0}{g_{ss}(\dot{y})} |\dot{y}| z \quad (10)$$

onde  $M$  é a massa total em movimento, composta pela massa da haste do cilindro, pela carga e massa de fluido deslocado,  $y$  é o deslocamento da haste do cilindro,  $\dot{y}$  é a velocidade e  $\ddot{y}$  é a aceleração.  $F_{atr}$  é a força de atrito, sendo que a primeira parcela,  $\sigma_0 z$ , representa a força elástica, a segunda parcela,  $\sigma_1 \dot{z}$ , representa o amortecimento, onde  $\dot{z}$  é dado pela equação (10) e  $\sigma_2 \dot{y}$ , representa o atrito viscoso. As pressões são representadas por  $p_a$  e  $p_b$  e  $\beta$  é o módulo de elasticidade do fluido.

As vazões  $Q_a$  e  $Q_b$ , em (7) e (8) são dadas por:

$$Q_a(u, p_a) = k_s \cdot x_v \cdot g_1(p_a, \text{sign}(u)) \quad (11)$$

$$Q_b(u, p_b) = -k_s \cdot x_v \cdot g_2(p_b, \text{sign}(u)) \quad (12)$$

Sendo que:

$$g_1(p_a, \text{sign}(u)) = \sqrt{\Delta p_a} = \begin{cases} \sqrt{p_s - p_a} & \text{para } u \geq 0 \\ \sqrt{p_a - p_r} & \text{para } u < 0 \end{cases} \quad (13)$$

$$g_2(p_b, \text{sign}(u)) = \sqrt{\Delta p_b} = \begin{cases} \sqrt{p_b - p_r} & \text{para } u \geq 0 \\ \sqrt{p_s - p_b} & \text{para } u < 0 \end{cases} \quad (14)$$

onde  $p_s$  é a pressão de suprimento,  $p_r$  é a pressão de retorno,  $k_s$  é uma constante hidráulica que representa as características geométricas da válvula e também agrega propriedades consideradas constantes para o por exemplo, o peso específico do fluido.

O modelo de Lugre para o atrito é dado pela equação (9) onde  $z$  é o coeficiente das microdeformações e  $g_{ss}(\dot{y})$ , representa uma função do atrito em regime permanente e é dada por:

$$g_{ss}(\dot{y}) = F_c + (F_s - F_c) \cdot e^{-\left(\frac{\dot{y}}{\dot{y}_s}\right)^2} \quad (15)$$

em que  $F_c$  é a Força de atrito de Coulomb,  $F_s$  é a Força de atrito estático,  $\dot{y}$  é a velocidade e  $\dot{y}_s$  é a velocidade de Stribeck.



## 5- Resultados de simulação

Esta seção apresenta uma descrição detalhada da metodologia utilizada na implementação da simulação computacional do modelo matemático para o atrito dinâmico, os parâmetros adotados para o sistema e os resultados de simulação para o caso de aplicação de uma força de entrada variável.

A simulação numérica do modelo proposto foi implementada com o auxílio da ferramenta computacional MaTLab/Simulink 7.4.0(R2007a), utilizando-se o método ode4 (Runge-Kutta).

Os parâmetros do sistema utilizados nas simulações computacionais estão apresentados na Tabela 1. Estes dados foram obtidos a partir do trabalho de Valdiero (2005).

Parâmetros	Valor	Unidade
Massa do corpo - $m$	20,66	kg
Força de Atrito de Coulomb - $F_c$	94,5	N
Força de Atrito Estático - $F_s$	413	N
$\sigma_0$	$0,5 \times 10^6$	N/m
$\sigma_1$	66,46	Ns/m
Coeficiente de atrito viscoso - $\sigma_2$	280	Ns/m
Velocidade de Stribeck - $\dot{y}_s$	0,041	m/s

Tabela 1. Parâmetros do modelo do sistema com atrito dinâmico.

O diagrama de blocos utilizado para a simulação do modelo não linear de 5ª ordem representado pelo sistema de equações diferenciais ordinárias (6), (7), (8), (9) e (10) é apresentado na Figura 2:



A tabela abaixo apresenta para cada entrada o tempo de simulação e a posição obtida:

Entrada (volts)	Tempo de simulação (seg)	Posição obtida (m)
2	4,512	0,4
4	3,264	0,3998

Tabela 2 - Tempo de simulação e posição obtida para cada sinal de entrada

Também testou-se o modelo com os mesmos sinais de entrada, porém negativos, representando o movimento oposto do cilindro, observando-se somente a inversão dos gráficos.

Observa-se na Figura 4, que, conforme os dados da Tabela 2, em torno de 4,51 segundos atinge-se a posição desejada, com 2 volts e em 3,26 segundos quando a entrada é de 4 volts.

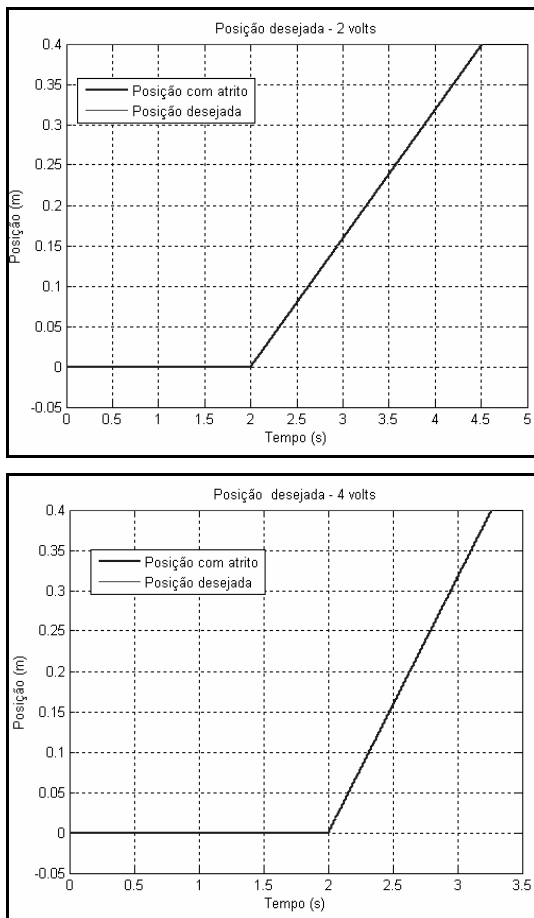


Figura 4 – Posicionamento do sistema

Tendo em vista que o comportamento dos gráficos de posição para ambos os sinais de entrada é o mesmo, alterando-se somente o tempo gasto para atingir a posição desejada, conforme dados da Tabela 2, eles não são apresentados.

A força de atrito inicia em 2 segundos, quando o sinal de entrada é dado e aumenta conforme o sinal de entrada até aproximadamente 170 N, oscilando momentaneamente e estabilizando nas proximidades de 140 N até os 4,51 segundos, que é o tempo para que a simulação de 2 v não ultrapasse o curso do cilindro.

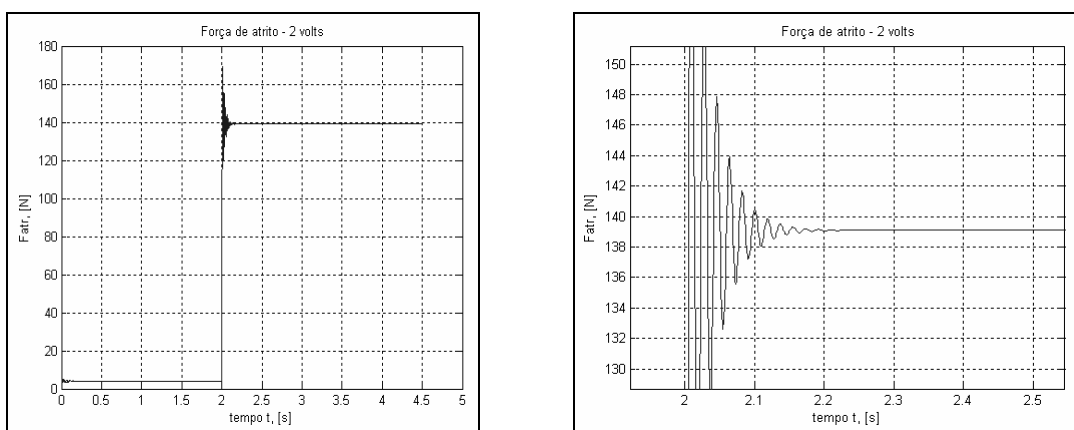


Figura 5 – Força de atrito com entrada de 2 volts

Quando é dada uma entrada de 4 v, a força de atrito sobe até aproximadamente 230 N, oscila rapidamente entre 200 N e 150 N e se mantém constante em torno de 170 N até o final da simulação.

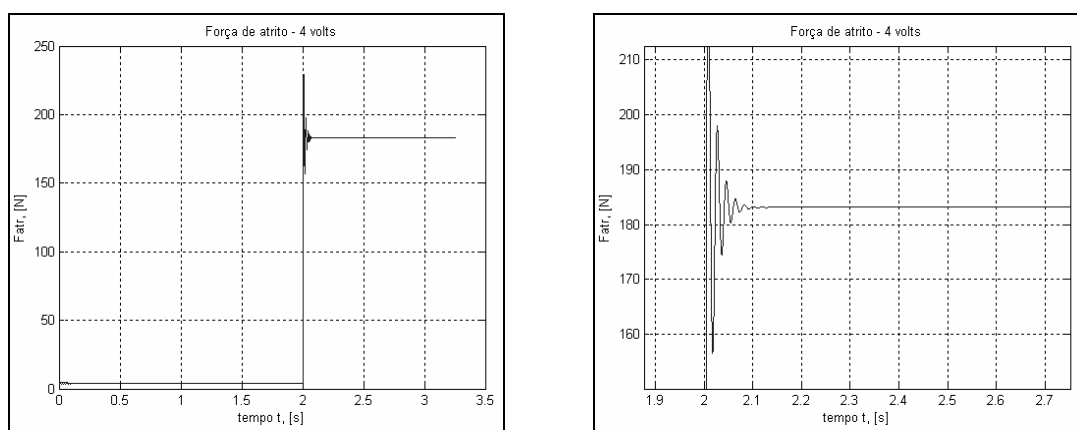


Figura 6 – Força de atrito com entrada de 4 volts

A seguir, apresentam-se os gráficos relativos à velocidade. No intervalo de tempo da simulação, para os dois sinais de entrada testados, é visível em ambos, a

oscilação que ocorre logo após a entrada do sinal. No instante em que a força de atrito estático é superada, ocorre o efeito Stribeck (passagem do atrito estático para o atrito de Coulomb) e ocorre a chamada quebra de movimento. Quando a força aplicada se torna menor que a força de atrito, a velocidade diminui, superando-se o atrito a velocidade volta a aumentar, e assim sucessivamente, ocorrendo a oscilação.

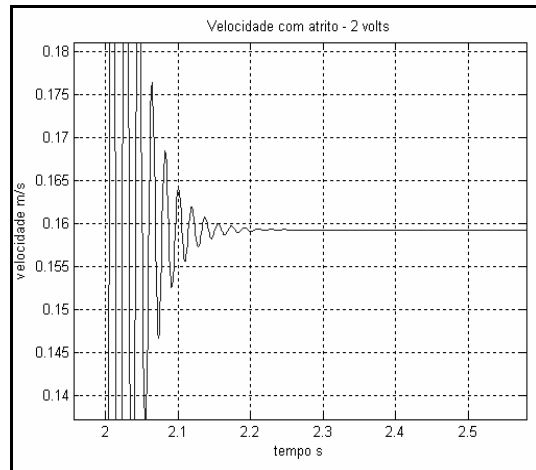
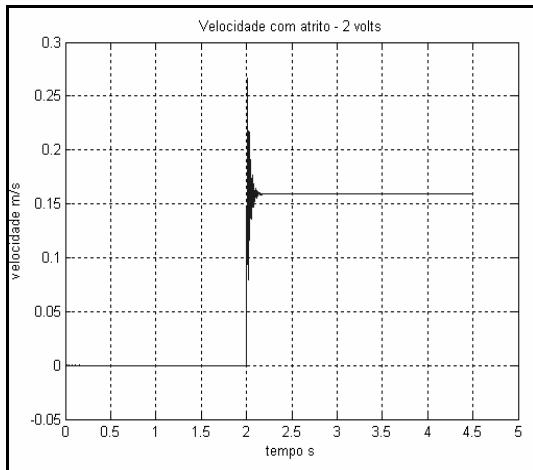


Figura 7 – Velocidade com entrada de 2 v

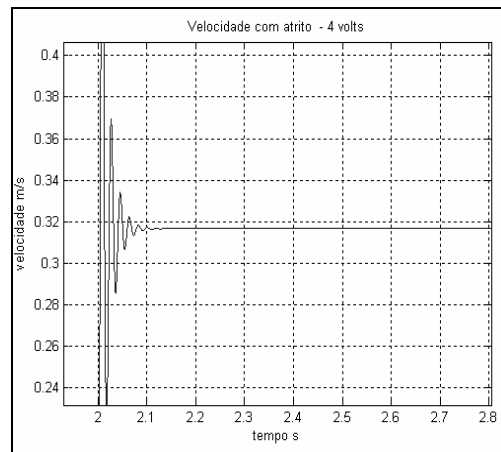
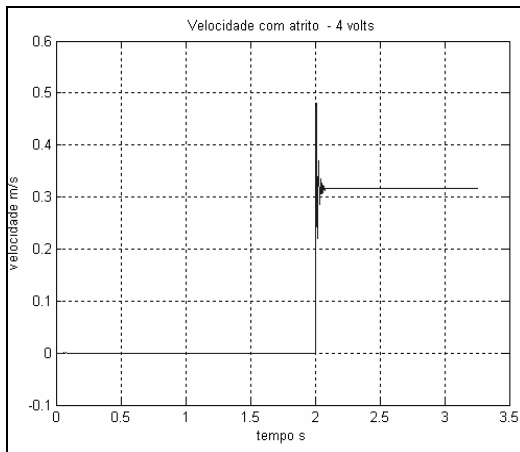


Figura 8 – Velocidade com entrada de 4 v

Aproximando-se o gráfico da velocidade, pode-se visualizar melhor a dinâmica da velocidade na presença de atrito.

## **6 – Conclusões:**

Neste trabalho foi apresentada a implementação e a execução da simulação computacional de um modelo matemático não linear de 5ª ordem para o atrito dinâmico.

Os resultados da simulação computacional do modelo em malha aberta foram obtidos para o sinal de entrada em degrau, ilustrando as características do comportamento do atrito.

Uma contribuição importante dos resultados de simulação é na verificação do projeto de sistemas de atuação, permitindo modificações e melhorias antes da construção dos protótipos experimentais.

## **7 - Agradecimentos**

Os autores agradecem à FAPERGS, à Secretaria de Ciência e Tecnologia/RS, ao CNPq e à UNIJUÍ pelo apoio.

## **Referências:**

- ACHO, L., IURIAN, C., IKHOUANE, F., RODELLAR, J.: **Robust-Adaptive Control of Mechanical Systems with Friction: Application to an Industrial Emulator**, Proceedings of the ACC, American Control Conference 2007, pp. 5970 – 5974, 2007.
- BOLLMANN, A., GUENTHER, R.: **Posicionadores hidráulicos e pneumáticos: características e técnicas de controle**, Seminário Nacional de Hidráulica e Pneumática-5, Florianópolis. Anais. Florianópolis: SENAI/CTAI, p. 57-78, 1997.
- CANUDAS-DE-WIT, C., LISCHINSKY, P.: **Adaptive friction compensation with partially known dynamic friction model**, International Journal of Adaptive Control and Signal Processing, Vol.11, pp.65-80, 1997.
- CHRISTENSEN, G. K., ZHOU, J., CONRAD, F., SORENSEN, T.: **The state of hydraulic technology and its electric competitions**. In: GARBACIK, Andrzej; STECKI, Jacek. (Ed.). Developments in fluid power control machinery and manipulators. Cracow: Fluid Power Net Publication, p. 156-159, 2000.
- CORNEJO, C., ALVAREZ-ICAZA, L.: **A nonlinear friction model for the passivity-based control of underactuated mechanical systems**, Proceedings of the 46th IEEE Conference on Decision and Control, pp. 3859 – 3864, 2007.

- DILDA, V., MIOTTO, F. E. M., VALDIERO, A. C., ANDRIGHETTO, P. L., RAFIKOV, M.: **Estudo do Comportamento Dinâmico de um Atuador Hidráulico e suas Características Não Lineares**, Congresso Nacional de Matemática Aplicada e Computacional, 2007, Florianópolis, Anais do XXX CNMAC, Florianópolis: SBMAC, p. 31, 2007.
- DUMITRIU, T.: **A Non-Linear Control Algorithm for a DC Servo Tracking System with Friction**, Proceedings of the EUROCON 2007, The International Conference on "Computer as a Tool", pp. 1718 – 1724, 2007.
- LINSINGEN, V.: **Fundamentos de sistemas hidráulicos**, 2. ed. rev. Florianópolis: Ed. da UFSC, 2003.
- MÁRTON, L., LANTOS, B.: **Modeling, Identification, and Compensation of Stick-Slip Friction**, IEEE Transactions on Industrial Electronics, Vol. 54, nº 1, pp 511 – 521, february 2007.
- MERRIT, H. E.: **Hydraulic control system**, New York: John Wiley & Sons, 1967.
- VALDIERO, A. C.: **Controle de robôs hidráulicos com compensação de atrito**, Tese de Doutorado em Engenharia Mecânica, Universidade Federal de Santa Catarina, Florianópolis, 2005.
- VALDIERO, A. C., ANDRIGHETTO, P. L., CARLOTTO, L.: **Dynamic modeling and friction parameters estimation to pneumatic actuators**, The International Symposium on Multibody Systems and Mechatronics, Uberlândia, 2005b.
- WENJING, Z.: **Parameter Identification of LuGre Friction Model in Servo System Based on Improved Particle Swarm Optimization Algorithm**, Proceedings of the 26th Chinese Control Conference, Zhangjiajie, Hunan, China, pp 135- 139, 2007.

# **AGRUPAMENTO BASEADO EM COLÔNIA DE FORMIGAS: ESTUDO COMPARATIVO DE ALGORITMOS PARA RECUPERAÇÃO DOS GRUPOS**

<sup>1</sup>Rosangela Villwock, <sup>2</sup>Maria Teresinha Arns Steiner.

<sup>1</sup>Universidade Estadual do Oeste do Paraná - UNIOESTE, rosangelamat@hotmail.com;

<sup>1,2</sup>Universidade Federal do Paraná - UFPR,  
Programa de Pós-Graduação em Métodos Numéricos em Engenharia, tere@mat.ufpr.br

## **Resumo**

O algoritmo de Agrupamento Baseado em Colônia de Formigas foi proposto inicialmente por DENEUBOURG e outros em 1991. Os padrões são espalhados numa grade e a cada formiga é atribuído um padrão. As formigas são responsáveis por carregar, transportar e descarregar os padrões na grade. Estas operações são baseadas na distribuição dos padrões. A recuperação dos grupos é feita usando-se as posições dos padrões na grade. O objetivo do presente trabalho foi testar o desempenho de algoritmos para recuperação dos grupos no Algoritmo de Agrupamento Baseado em Colônia de Formigas e, para tanto, foram abordados os exemplos IRIS e WINE. O algoritmo foi aplicado por cinco vezes e, na recuperação dos grupos, ficou definido o número de grupos. Os resultados foram satisfatórios. O método que apresentou os melhores resultados (em média), dentre os utilizados para a comparação, foi o Método Ward. No exemplo IRIS, os resultados que apresentaram menor variância, para todos os algoritmos para recuperação dos grupos, foram iguais. No exemplo WINE, o algoritmo que apresentou menor variância foi o Método Ward, porém, nos resultados que apresentaram menor variância, as Ligações Simples, Completa e Média apresentaram maior número de padrões classificados corretamente.

**Palavras-chave:** Agrupamento, Metaheurística, Algoritmos para Recuperação dos Grupos.

## **Apresentação:**

### **Agrupamento Baseado em Colônia de Formigas**

O primeiro trabalho fazendo uso dos métodos baseados em colônia de formigas foi de DORIGO (1992) *apud* ARANHA (2006). Neste trabalho, um grupo de formiga



foi usado para resolver um problema de transporte. Quando uma formiga percorre um caminho, ela deixa um sinal (feromônio) através do trajeto que usou. Os trajetos mais curtos terão os sinais mais fortes. As formigas seguintes, ao decidir que trajeto fazer, tendem a escolher trajetos com sinais mais fortes com uma probabilidade mais elevada, de modo que os trajetos mais curtos são encontrados.

O algoritmo de agrupamento Baseado em Colônia de Formigas foi proposto inicialmente por DENEUBOURG *et al.* (1991) *apud* HANDL *et al.* (2005). As formigas foram representadas como agentes simples que se movem aleatoriamente em uma grade quadrada. Os padrões foram dispersos dentro desta grade e poderiam ser carregados, transportados e descarregados pelos agentes. Estas operações são baseadas na distribuição dos padrões dentro da vizinhança local dos agentes; padrões que foram isolados ou cercados por dissimilares eram mais prováveis de serem carregados e então descarregados numa vizinhança de similares. As decisões de carregar e descarregar padrões são tomadas pelas probabilidades  $P_p$  e  $P_d$ , dadas pelas equações 1 e 2 a seguir. Os autores usaram em seu trabalho  $k_1 = 0.1$  e  $k_2 = 0.3$ .

$$P_p = \left( \frac{k_1}{k_1 + f} \right)^2 \quad (1)$$

$$P_d = \left( \frac{f}{k_2 + f} \right)^2 \quad (2)$$

Onde  $f$  é uma estimativa da fração de padrões localizados na vizinhança, que são semelhantes ao padrão atual da formiga. Neste trabalho, a estimativa  $f$  é obtida através de uma memória de curto prazo de cada formiga, onde o conteúdo da última célula da grade analisada é armazenado. Esta escolha da função de vizinhança  $f$  foi essencialmente motivada pela sua facilidade de realização por robôs simples.

LUMER & FAIETA (1994) *apud* HANDL *et al.* (2005) introduziram um número de modificações ao modelo que permitiu a manipulação de dados numéricos e melhorou a qualidade da solução e o tempo da convergência do algoritmo. A idéia era definir uma medida de similaridade ou dissimilaridade entre os objetos, já que no algoritmo proposto inicialmente, os objetos eram similares (se os objetos fossem

idênticos) ou dissimilares (se os objetos não fossem idênticos). Neste trabalho aparece pela primeira vez o mapeamento topográfico.

Neste trabalho, as decisões de carregar e descarregar padrões são baseadas nas probabilidades  $P_p$  e  $P_d$  dadas pelas equações 3 e 4 a seguir, onde  $f(o_i)$  é dada pela equação 5. Na equação 5,  $d(o_i, o_j)$  é uma função de dissimilaridade entre padrões  $o_i$  e  $o_j$  pertencentes ao intervalo  $[0, 1]$ ;  $\alpha$  é um parâmetro escalar dependente dos dados e pertencente ao intervalo  $[0, 1]$ ;  $L$  é o tamanho da vizinhança local e  $L$  é igual a  $s^2$ . Os autores usaram em seu trabalho  $k_1 = 0,1$  e  $k_2 = 0,15$  e  $\alpha = 0,5$ .

$$P_p(o_i) = \left( \frac{k_1}{k_1 + f(o_i)} \right)^2 \quad (3)$$

$$P_d(o_i) = \begin{cases} 2f(o_i) & \text{se } f(o_i) < k_2 \\ 1 & \text{se } f(o_i) \geq k_2 \end{cases} \quad (4)$$

$$f(o_i) = \max \left\{ 0, \frac{1}{s^2} \sum_{o_j \in \text{Viz}_{(xxx)}(r)} \left[ 1 - \frac{d(o_i, o_j)}{\alpha} \right] \right\} \quad (5)$$

Para RAMOS & MERELO (2002) as decisões de carregar e descarregar padrões são tomadas por uma das funções descritas no quadro 1, a seguir, baseadas nos valores de  $\chi$ ,  $\delta$  e  $\varepsilon$  dados pelas equações 6, 7 e 8, respectivamente, onde  $n$  é o número de objetos em uma vizinhança,  $d$  é a distância padronizada,  $k_1 = 0,1$  e  $k_2 = 0,3$ .

Quadro 1 – Funções de carregamento e descarregamento (RAMOS & MERELO, 2002)

Funções	Carregar	Descarregar
#1	$P_p = (1 - \chi) \varepsilon$	$P_d = \chi \delta$
#2	$P_p = (1 - \chi) \varepsilon$ (a)	$P_d = \chi \delta$ (a)
	$P_p = \varepsilon$ (b)	$P_d = \delta$ (b)
#3	$P_p = 1 - \chi$ (a)	$P_d = \chi$ (a)
	$P_p = \varepsilon$ (b)	$P_d = \delta$ (b)

$$\chi = \frac{n^2}{n^2 + 5^2} \quad (6)$$

$$\delta = \left( \frac{k_1}{k_1 + d} \right)^2 \quad (7)$$

$$\varepsilon = \left( \frac{d}{k_2 + d} \right)^2 \quad (8)$$

HANDL (2003) sugere mudanças que melhoram a separação espacial dos grupos e permitem que o algoritmo seja mais robusto. Uma delas é a restrição na função  $f$  (equação 9), a seguir, que serve para penalizar dissimilaridades elevadas.

$$f^*(i) = \begin{cases} \frac{1}{\sigma^2} \sum_j \left[ 1 - \frac{d(i,j)}{\alpha} \right] & \text{se } f^*(i) > 0 \wedge \forall j \left( 1 - \frac{d(i,j)}{\alpha} \right) > 0 \\ 0 & \text{caso contrário} \end{cases} \quad (9)$$

Segundo HANDL (2003) somente valores da função de vizinhança  $f(i) \geq 1$  tornam a operação de carregamento menos provável e, analogamente, valores da função de vizinhança  $f(i) \geq 1$  tornam a operação de descarregamento completamente determinística. Segundo a autora as decisões de carregar e descarregar padrões podem ser tomadas pelas probabilidades  $P_p$  e  $P_d$  dadas pelas equações 10 e 11, respectivamente.

$$p_p^*(i) = \begin{cases} 1 & \text{se } f^*(i) \leq 1 \\ \frac{1}{f^*(i)^2} & \text{caso contrário} \end{cases} \quad (10)$$

$$p_d^*(i) = \begin{cases} 1 & \text{se } f^*(i) \geq 1 \\ \frac{1}{f^*(i)^4} & \text{caso contrário} \end{cases} \quad (11)$$

O processo de agrupamento foi acelerado significativamente pelo uso de uma memória de curto prazo introduzida por LUMER & FAIETA (1994) *apud* HANDL *et al.* (2005). Cada agente recorda os últimos padrões carregados e as respectivas posições onde foi descarregado. Quando um padrão novo é carregado, a posição de *best matching* (melhor emparelhamento) memorizada será usada para indicar o sentido da direção aleatória do agente. O *best matching* é a posição de mínima dissimilaridade.

Na proposta de HANDL (2003), a memória de curto prazo é estendida. Os padrões armazenados na memória podem já terem sido removidos da posição registrada. A fim de determinar mais robustamente o sentido da direção, a autora permite que cada

agente “olhe adiante”. Um agente que carrega um padrão  $i$  usa sua memória para examinar todas as posições registradas, uma após a outra, usando a função de vizinhança  $f(i)$ . O *best matching* é a célula da grade para qual a função da vizinhança rende o valor o mais elevado. Entretanto, o salto para o *best matching* só é feito com alguma probabilidade, dependente da qualidade do *match*.

RAMOS & MERELO (2002) sugerem o uso do feromônio para mover a formiga. Seu algoritmo foi chamado de ACLUSTER. Os autores concluem que obtiveram resultados similares ao modelo de Lumer e Faieta e o seu algoritmo é muito mais simples. A probabilidade de mudança da posição  $k$  para a posição  $i$  é dada pela equação 12, onde  $j/k$  indica a soma sobre todos os  $j$  que estão na vizinhança de  $k$ ,  $\Delta_i$  é uma medida da diferença de orientação em relação à direção anterior, atribuindo um peso  $w_i$  para cada mudança de direção e  $W(\sigma)$  é dado pela equação 13, com  $\beta = 3,5$  e  $\gamma = 0,2$ .

$$p_{ik} = \frac{W(\sigma_i)w(\Delta_i)}{\sum_{j/k} W(\sigma_j)w(\Delta_j)} \quad (12)$$

$$W(\sigma) = \left(1 + \frac{\sigma}{1 + \gamma\sigma}\right)^\beta \quad (13)$$

Quanto ao raio de percepção, é atrativo empregar vizinhanças maiores para melhorar a qualidade do agrupamento e da distribuição na grade. Porém, além ser mais caro (porque o número das células a serem consideradas para cada ação cresce quadraticamente com o raio), inibe também a formação rápida dos grupos durante a fase de distribuição inicial. HANDL (2003) usou um raio da percepção que aumenta gradualmente com o tempo e isso acelera a dissolução de grupos pequenos preliminares.

### ***O algoritmo básico***

Na fase inicial, todos os padrões são aleatoriamente espalhados na grade. Depois, cada formiga escolhe um padrão para carregar e é colocada em uma posição na grade.

Na segunda fase, uma formiga é selecionada aleatoriamente. Esta formiga se desloca na grade executando um passo de comprimento  $L$ , numa direção determinada aleatoriamente. A formiga decide, probabilisticamente, se descarrega seu padrão na posição atual.

Se a decisão de descarregar o padrão for negativa, escolhe-se aleatoriamente outra formiga. No caso de decisão positiva, a formiga descarrega o padrão em sua posição atual na grade. Se esta célula da grade está ocupada por outro padrão, o mesmo deve ser descarregado numa célula imediatamente vizinha desta, que esteja livre, por meio de uma procura aleatória.

A formiga então procura um novo padrão para carregar. Dentre os padrões livres na grade, a formiga seleciona aleatoriamente um padrão, pula para a posição do padrão na grade, faz a avaliação da função de vizinhança e decide probabilisticamente se carrega o padrão. Este processo de escolha de um padrão livre na grade é executado até que a formiga encontre um padrão que deve ser carregado.

A segunda fase é reiniciada, escolhendo-se outra formiga até que um critério de parada seja satisfeito.

### ***A recuperação dos grupos***

Segundo HANDL (2003), o processo inicia com cada padrão formando um grupo. Depois de calcular as distâncias entre todos os grupos, devem-se fundir (ou ligar) os dois grupos com menor distância. Os tipos de ligações mais comuns são Ligações Simples, Ligações Completas, Ligações Médias e Método de Ward (JOHNSON & WICHERN, 1998). As distâncias entre grupos são definidas em termos de distância na grade. Este processo se repete até que um critério de parada seja satisfeito.

Na ligação simples (ou vizinho mais próximo), a distância entre dois grupos é a mínima das distâncias entre todos os pares de padrões  $i$  e  $j$ , com  $i$  pertencente ao primeiro grupo e  $j$  ao segundo (JAIN *et al.*, 1999). Por exemplo, se o grupo 14 é formado pelos padrões  $U$  e  $V$  e o grupo 2 é formado pelo padrão  $W$ , a distância entre os grupos 1 e 2 é calculada pela equação 14.

$$d_{(1,2)} = \min\{d_{UW}, d_{VW}\} \quad (14)$$

Na ligação completa (ou vizinho mais distante), a distância entre dois grupos é a máxima das distâncias entre todos os pares de padrões  $i$  e  $j$ , com  $i$  pertencente ao primeiro grupo e  $j$  ao segundo (JAIN *et al.*, 1999). Por exemplo, se o grupo 1 é formado pelos padrões  $U$  e  $V$  e o grupo 2 é formado pelo padrão  $W$ , a distância entre os grupos 1 e 2 é calculada pela equação 15.

$$d_{(1,2)} = \max\{d_{UW}, d_{VW}\} \quad (15)$$

Na ligação média, a distância entre dois grupos é a média das distâncias entre todos os pares de padrões, sendo que cada padrão do par é de um grupo. Se o grupo 1 é formado pelos elementos  $U$  e  $V$  e o grupo 2 é formado pelo elemento  $W$ , a distância entre os grupos 1 e 2 é calculada pela equação 16, onde  $d_{ik}$  é a distância entre o padrão  $i$  no grupo 1 e o padrão  $k$  no grupo 2,  $N_1$  é o número de padrões no grupo 1 e  $N_2$  é o número de padrões no grupo 2.

$$d_{(1,2)} = \Sigma \Sigma d_{ik} / N_1 * N_2 \quad (16)$$

O método de Ward faz a junção de dois grupos baseando-se na “perda de informação”. Considera-se como critério de “perda de informação” a soma do quadrado do erro ( $SQE$ ). Para cada grupo  $i$ , calcula-se a média (ou centróide) do grupo e a soma do quadrado do erro do grupo  $i$  ( $SQE_i$ ) que é a soma do quadrado do erro de cada padrão do grupo em relação à média. Para  $k$  grupos têm-se  $SQE_1, SQE_2, \dots, SQE_k$  e define-se  $SQE$  pela equação 17.

$$SQE = SQE_1 + SQE_2 + \dots + SQE_k \quad (17)$$

Para cada par de grupos  $m$  e  $n$ , primeiramente, calcula-se a média (ou centróide) do grupo formado (grupo  $mn$ ). Em seguida, calcula-se a soma do quadrado do erro do grupo  $mn$  ( $SQE_{mn}$ ) pela equação 18. Os grupos  $m$  e  $n$  que apresentarem o menor aumento na soma do erro quadrático ( $SQE$ ) (menor “perda de informação”) serão unidos.

$$SQE = SQE_1 + SQE_2 + \dots + SQE_k - SQE_m - SQE_n + SQE_{mn} \quad (18)$$

### ***Avaliação dos grupos***

Segundo TAN *et al.* (2005), as medidas numéricas aplicadas para julgar vários aspectos de avaliação de grupos são classificadas em três tipos: 1) os índices externos são usados para medir até que ponto rótulos de grupos correspondem a rótulos de classes externamente fornecidos; 2) os índices internos são usados para medir quão boa é a estrutura de agrupamento sem relação com informação externa e 3) os índices relativos são usados para comparar dois grupos ou agrupamentos diferentes. As medidas utilizadas neste trabalho foram: Índice Randômico (HANDL *et al.*, 2003) e Variância. O Índice Randômico é um índice externo e a Variância é um índice interno.

O Índice Randômico é dado pela equação 19, onde  $a$ ,  $b$ ,  $c$  e  $d$  são calculados para todos os possíveis pares de padrões  $i$  e  $j$  e seus respectivos grupos  $U$  (classificação correta -  $c_U(i)$  e  $c_U(j)$ ) e  $V$  (solução gerada pelo algoritmo de agrupamento -  $c_V(i)$  e  $c_V(j)$ ).  $R$  está limitado no intervalo  $[0, 1]$  e deve ser maximizado.

$$R = \frac{a + d}{a + b + c + d} \quad (19)$$

onde:

$$a = |\{i, j \mid c_U(i) = c_U(j) \wedge c_V(i) = c_V(j)\}|,$$

$$b = |\{i, j \mid c_U(i) = c_U(j) \wedge c_V(i) \neq c_V(j)\}|,$$

$$c = |\{i, j \mid c_U(i) \neq c_U(j) \wedge c_V(i) = c_V(j)\}| \text{ e}$$

$$d = |\{i, j \mid c_U(i) \neq c_U(j) \wedge c_V(i) \neq c_V(j)\}|.$$

A Variância é a soma dos quadrados dos desvios entre todos os padrões e os respectivos centróides do grupo à que pertencem, dividido pelo número total de padrões.

O objetivo do presente trabalho, conforme já comentado, é testar o desempenho de algoritmos para recuperação dos grupos no Algoritmo de Agrupamento Baseado em Colônia de Formigas.

### **Metodologia**

Neste trabalho foi aplicado o Algoritmo de Agrupamento Baseado em Colônia de Formigas para o agrupamento dos padrões. Para a recuperação do Agrupamento fornecido por este algoritmo foram usadas: Ligação Simples, Ligação Completa, Ligação Média e Ward.

### ***Exemplos Abordados***

O exemplo IRIS é composto de 150 padrões (plantas). Neste exemplo são conhecidos os grupos a que cada planta pertence. Os 150 padrões são divididos em três grupos com 50 padrões em cada grupo: Íris Setosa, Íris Versicolour e Íris Virginica. Cada padrão é constituído de quatro atributos numéricos. Os atributos são: comprimento da pétala, largura da pétala, comprimento da sépala e largura da sépala. Os dados podem ser obtidos no endereço <http://mlearn.ics.uci.edu/databases>.

O exemplo WINE é composto de 178 padrões (vinhos). Neste exemplo são conhecidos os grupos a que cada padrão pertence. Os 178 padrões são divididos em três grupos: 59 padrões pertencem ao grupo 1, 71 padrões pertencem ao grupo 2 e 48 padrões pertencem ao grupo 3. Cada padrão é constituído de 13 atributos numéricos. Os atributos são resultados de uma análise química. Os dados podem ser obtidos no endereço <http://mlearn.ics.uci.edu/databases>.

### ***A Implementação do Algoritmo de Agrupamento Baseado em Colônia de Formigas***

O algoritmo implementado utiliza como critério de parada o número de iterações. A matriz  $Q$  de entrada dos dados é de ordem  $m \times n$ , onde  $m$  é o número de padrões e  $n$  é o número de atributos. A matriz  $P$  é a matriz de dados padronizada. A padronização foi feita por dimensão e os valores resultantes pertencentes ao intervalo



[0, 1]. A medida de dissimilaridade utilizada é a Distância Euclidiana e foi utilizada a matriz de distância padronizada. Os parâmetros utilizados são descritos no quadro 2, a seguir.

Quadro 2 - Parâmetros utilizados no algoritmo de Agrupamento Baseado em Colônia de Formigas

Parâmetro	Valor
Tamanho da grade	$N_c = 10N_p$
Número de formigas	20
Número de iterações	100000
Raio de vizinhança	1
Probabilidade mínima para descarregar um padrão	$drop_{\min} = 0,5$
$\alpha$ (equação 9)	0,8

Nota:  $N_c$  é o número de células;  $N_p$  é o número de padrões

Foi utilizada vizinhança quadrada na busca dos padrões vizinhos e a direção do passo é aleatória. Definida a direção, calculou-se o tamanho máximo possível do passo. Um número randômico pertencente ao intervalo [0, 1] foi utilizado para determinar o tamanho do passo, multiplicando-se este número pelo tamanho máximo do passo.

As probabilidades de carregar ( $p_p$ ) e descarregar ( $p_d$ ) utilizadas são as descritas pelas equações 1 e 2, com o cálculo da função  $f$  dado pela equação 9. A probabilidade  $p_d$  maior que 0,5 foi escolhida para garantir a similaridade entre o dado a ser descarregado e seus vizinhos. Além disso, a probabilidade  $p_d$  da posição atual é comparada à probabilidade  $p_d$  da posição futura e o dado é descarregado somente quando a probabilidade  $p_d$  da posição futura é maior que a probabilidade  $p_d$  da posição atual. A probabilidade  $p_p$  é comparada à probabilidade  $p_d$  e o dado é carregado somente quando a probabilidade  $p_p$  é maior que a probabilidade  $p_d$ .

Quando um dado é descarregado na grade, um novo dado deve ser carregado. Esta busca é aleatória, porém, cada dado livre é avaliado somente uma vez, não sendo permitido que ele seja avaliado novamente, até que todos sejam avaliados. Caso

nenhum dado apresente probabilidade  $p_p$  maior que  $p_d$ , o dado que apresentar a maior probabilidade  $p_p$  é carregado.

Quando um dado não tem vizinhos, definiu-se a função  $f$  igual à zero. Isso faz com que a probabilidade  $p_d$  seja igual à “0”, ou seja, o dado não deve ser descarregado naquela posição e a probabilidade  $p_p$  seja igual a “1”, ou seja, o dado deve ser carregado e futuramente deixar esta posição. Na recuperação dos grupos foi definido o número de grupos.

## **Resultados**

O Algoritmo de Agrupamento Baseado em Colônia de Formigas foi aplicado aos dois exemplos (IRIS e WINE) por cinco vezes. As figuras 1 e 2, mais adiante, apresentam a distribuição dos padrões na grade para os resultados que apresentaram maior número de padrões classificados corretamente após a aplicação dos algoritmos para recuperação dos grupos, dentre os resultados com menor variância. No exemplo IRIS, para todos os algoritmos para recuperação dos grupos, os resultados foram iguais. No exemplo WINE, o algoritmo que apresentou menor variância foi o Método Ward, porém as Ligações Simples, Completa e Média apresentaram maior número de padrões classificados corretamente.

Na figura 1 observa-se claramente a separação do grupo 1 (padrões em vermelho), apesar de alguns padrões não pertencentes a este grupo estarem próximos e padrões deste grupo estarem em posições distantes na grade. Porém, a separação entre os grupos 2 e 3 (padrões em preto e azul, respectivamente) não é clara. A aplicação dos algoritmos para recuperação dos grupos podem fazer esta separação.

Na figura 2 observa-se a separação parcial do grupo 1 (padrões em vermelho), com alguns padrões não pertencentes a este grupo próximos. Os grupos 2 e 3 (padrões em preto e azul, respectivamente) também não é clara. Aqui, também, a aplicação dos algoritmos para recuperação dos grupos podem fazer esta separação.

Aplicando os algoritmos para recuperação dos grupos temos a formação de três grupos para cada uma dos exemplos (IRIS e WINE). Os quadros 3 e 4 mostram a





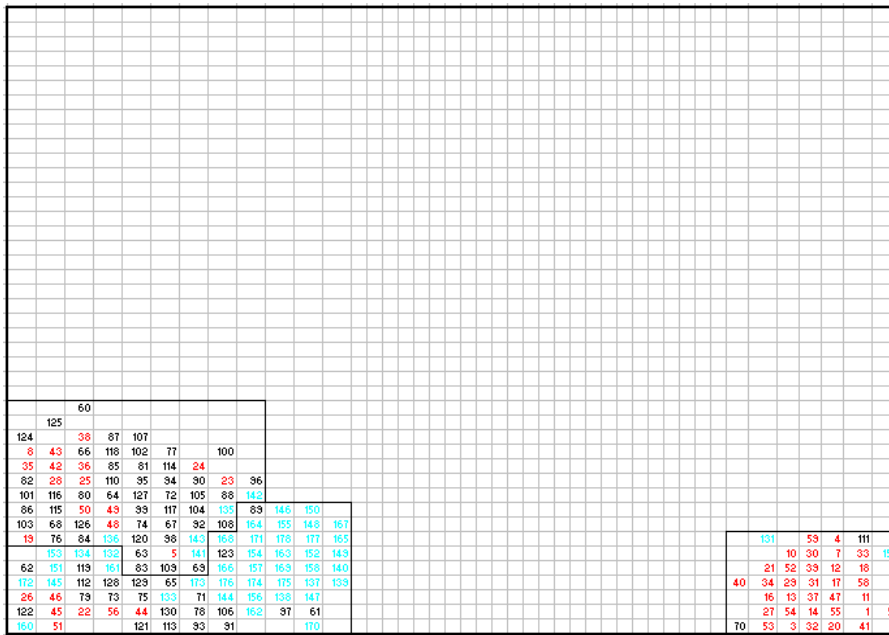


Figura 4 – Distribuição dos padrões na grade com divisão dos grupos - WINE

As figuras 3 e 4 mostram os melhores resultados para os resultados apresentados nas figuras 1 e 2, respectivamente. Para estes resultados tem-se 99 dos 150 padrões do exemplo IRIS e 125 dos 178 padrões do exemplo WINE classificados corretamente

O quadro 5, a seguir, apresenta a média e o desvio-padrão para das medidas de erro Índice Randômico (R) e Variância para os dois exemplos abordados. Os melhores resultados estão em negrito. Pode-se observar que os melhores resultados são obtidos com a aplicação do Método Ward para recuperação dos grupos.

Quadro 5 – Resultados da Aplicação dos Algoritmos de Recuperação dos Grupos

Exemplos	Medidas	Ligação Simples		Ligação Completa		Ligação Média		Ward	
		Média	Desvio-Padrão	Média	Desvio-Padrão	Média	Desvio-Padrão	Média	Desvio-Padrão
Íris	R	0,494	0,085	0,714	0,023	0,704	0,022	<b>0,718</b>	0,020
	Variância	0,126	0,027	0,075	0,010	0,076	0,012	<b>0,073</b>	0,011
Wine	R	0,405	0,038	0,608	0,025	0,593	0,017	<b>0,609</b>	0,029
	Variância	0,293	0,008	<b>0,258</b>	0,012	0,261	0,010	0,260	0,016

### Conclusões

Os resultados foram considerados satisfatórios. O método que apresentou os melhores resultados (em média) foi o Método Ward. No exemplo IRIS, os resultados

que apresentaram menor variância, para todos os algoritmos para recuperação dos grupos, foram iguais. No exemplo WINE, o algoritmo que apresentou menor variância foi o Método Ward, porém, nos resultados que apresentaram menor variância, as Ligações Simples, Completa e Média apresentaram maior número de padrões classificados corretamente.

Estes testes ainda serão aplicados a outros exemplos. Também são necessários novos testes com maior número de repetições. Outros testes estão sendo realizados porém ainda sem resultados conclusivos.

### **Agradecimento**

A primeira autora agradece à CAPES pela bolsa que vem sendo concedida para o desenvolvimento de seu Curso de Doutorado.

### **Referências Bibliográficas**

ARANHA, C. de C.. **A Survey on using Ant-Based techniques for clustering.** Institute's seminar, 2006. Disponível em <http://www.iba.k.u-tokyo.ac.jp/~caranha/research/ant-cluster-survey.pdf>. Acesso em 05/03/2007.

HANDL, J. **Ant-based methods for tasks of clustering and topographic mapping: extensions, analysis and comparison with alternative methods.** Masters thesis. University of Erlangen-Nuremberg, Germany, 2003.

HANDL, J.; KNOWLES, J.; DORIGO, M.. **Ant-based clustering: a comparative study of its relative performance with respect to k-means, average link and 1d-som.** Technical Report TR/IRIDIA/2003-24, IRIDIA, Université Libre de Bruxelles, July 2003.

HANDL, J.; KNOWLES, J.; DORIGO, M.. **Ant-Based Clustering and Topographic Mapping.** Artificial Life, V12(1), 2005.

JAIN A. K., MURTY M. N., FLYNN P. J.. **Data clustering: a review.** ACM Computing Surveys. v. 31, n. 3, set. 1999.

JOHNSON, R.A. E WICHERN, D.W. **Applied Multivariate Statistical Analysis.** 4nd. Edition. Ed. Prentice Hall, 1998.

RAMOS V. & MERELO J.J.. **Self-Organized Stigmergic Document Maps: Environment as a Mechanism for Context Learning.** Procs. of AEB'02 - 1st Spanish Conf. on Evolutionary and Bio-Inspired Algorithms, Merida Univ., Spain, Feb. 02, pp. 284-293, 2002.

TAN, P. N.; STEINBACH, M.; KUMAR, V.. **Introduction to Data Mining.** Inc. Boston, MA, USA: Addison-Wesley Longman Publishing Co., 2005.

# CÓDIGOS CORRETORES DE ERROS EM ESPAÇOS DE BLOCOS DE ROSENBLOOM-TSFASMAN

Luciano Panek

Universidade Estadual do Oeste do Paraná - Centro de Engenharias e Ciências Exatas -

UNIOESTE-CECE, Foz do Iguaçu, lucpanek@gmail.com

## Resumo:

Neste trabalho apresentamos uma nova família de métricas intercalando a métrica de blocos introduzida por Feng, Xu e Hickernell com a métrica de Rosenbloom-Tsfasman. Assim como Victor Wei, consideramos os pesos generalizados em relação a nova métrica. A partir dos pesos generalizados estendemos os resultados de Ozen e Siap: determinação de uma forma padrão de matriz geradora, limitante de Singleton generalizado, cálculo do espectro dos pesos generalizados, decodificação por máxima verossimilhança. Determinamos também o raio de empacotamento, o raio de cobertura (em função do último peso generalizado) e apresentamos uma classificação dos códigos perfeitos e quase-perfeitos. Uma proposta de algoritmo de decodificação via síndromes é apresentada.

**Palavras-chave:** códigos lineares, métrica de Rosenbloom-Tsfasman, métrica de blocos, métrica de blocos ponderada, decodificador por máxima verossimilhança.

## Introdução

Insatisfeito com o funcionamento de seu computador nos Laboratórios Bell (Bell Telephone Laboratories), Richard Hamming desenvolve em 1947 o primeiro código corretor de erros, conhecido hoje como código binário de Hamming. As idéias de Hamming foram publicadas somente em 1950 (ver [4]). Um exemplo de dimensão baixa do código

binário de Hamming apareceu em 1948 no célebre trabalho de Claude Shannon (ver [21]). Em [4] Hamming também estabelece os conceitos fundamentais da teoria dos códigos corretores de erros: métrica de Hamming, códigos equivalentes e o limitante de empacotamento de esferas. Os fundamentos teóricos sobre códigos lineares foram estabelecidos em 1956 por David Slepian (ver [22]): códigos de grupo, arranjo padrão, síndromes e algoritmo de decodificação por máxima verossimilhança via síndromes. Seguindo os passos de Hamming e Slepian, apresentamos neste trabalho os conceitos fundamentais em relação a uma nova família de métricas, as chamadas métricas de blocos de Rosenbloom-Tsfasman (descritas na próxima seção).

Iniciamos apresentando de forma sucinta os conceitos introduzidos por Hamming e Slepian em [4] e [22]. Estes conceitos são clássicos e são encontrados em qualquer bom livro de teoria dos códigos corretores de erros (por exemplo, ver [5]). Seja  $\mathbb{F}_q^N$  o espaço vetorial das  $N$ -uplas sobre o corpo finito  $\mathbb{F}_q$  com  $q$  elementos. Um vetor  $x \in \mathbb{F}_q^N$  será denotado por  $x_1x_2 \dots x_N$ , sendo  $x_i$  a  $i$ -ésima coordenada de  $x$ . Um subespaço  $k$ -dimensional  $C \subseteq \mathbb{F}_q^N$  é chamado de  $[N; k]$  **código linear**. Os vetores de um  $[N; k]$  código linear são chamados de **palavras-código**. Se  $\{g_1, g_2, \dots, g_k\}$  é uma base de  $C$  então  $G = (g_i)$  é a chamada **matriz geradora** de  $C$ :

$$C = \{x \cdot G : x \in \mathbb{F}_q^k\}.$$

O subespaço  $C^\perp = \{y \in \mathbb{F}_q^N : c \cdot y = 0 \text{ para todo } c \in C\}$ , onde  $c \cdot y = c_1 \cdot y_1 + \dots + c_N \cdot y_N$  é o **produto interno** em  $\mathbb{F}_q^N$ , é o chamado **código dual** de  $C$ . Uma matriz geradora  $H$  de  $C^\perp$  é dita uma **matriz de paridade** de  $C$ : denotando o vetor nulo  $00 \dots 0$  por  $\mathbf{0}$ ,  $c \in C$  se, e somente se,  $H \cdot c^T = \mathbf{0}$ . Para cada  $x \in \mathbb{F}_q^N$  seja  $\text{supp}_H(x) = \{i : x_i \neq 0\}$  o **suporte de Hamming** de  $x$ . O número  $w_H(x) = |\text{supp}_H(x)|$  é o **peso de Hamming** de  $x$ . A **distância de Hamming**  $d_H(x, y)$  entre  $x$  e  $y$  é o número de coordenadas distintas entre  $x$  e  $y$ , ou seja,  $d_H(x, y) = w_H(x - y)$ . A distância de Hamming é uma métrica em  $\mathbb{F}_q^N$  ( $d_H$  é positiva definida, simétrica e satisfaz a desigualdade triangular). O espaço métrico  $(\mathbb{F}_q^N, d_H)$  é conhecido como **espaço de Hamming**. Associado a um



código linear  $C$  temos a **distância mínima de Hamming** de  $C$ :

$$d_H(C) = \min \{d_H(c, c') : c \neq c' \in C\}.$$

Vale que  $d_H(C) = \min \{w_H(c) : \mathbf{0} \neq c \in C\}$ . Dado  $r$  um inteiro positivo e  $x \in \mathbb{F}_q^N$ , o conjunto de todos os vetores  $y \in \mathbb{F}_q^N$  tal que  $d_H(x, y) \leq r$  é denotado por  $B_H(x; r)$  e chamado de **bola** de centro  $x$  e raio  $r$ . O maior  $r$  tal que  $B_H(c; r) \cap B_H(c'; r) = \emptyset$  para todo  $c \neq c' \in C$  é a chamada **capacidade de correção de erros** de  $C$  (ou o **raio de empacotamento** de  $C$ ). Dado um código linear  $C$ , é bem conhecido na literatura que

$$t := \left\lfloor \frac{d_H(C) - 1}{2} \right\rfloor$$

é a capacidade de correção de erros de  $C$ . Se durante a transmissão de uma palavra código  $c$  ocorrerem no máximo  $t$  erros e recebermos a mensagem  $x$  ( $x = c + e$  com  $w_H(e) \leq t$ ), então  $c$  é a única palavra código mais próxima de  $x$ : a mensagem  $x$  é decodificada corretamente com sendo a palavra código  $c$ . O procedimento de decodificar uma mensagem  $x$  como sendo a palavra código mais próxima de  $x$  é conhecido como **decodificador por máxima verossimilhança**. Um algoritmo simples para decodificação é baseado nas síndromes dos vetores de  $\mathbb{F}_q^N$ . Dado um  $[N; k]$  código linear  $C$  e uma matriz de paridade  $H$  de  $C$ , definimos a **síndrome**  $S(x)$  de  $x \in \mathbb{F}_q^N$  como sendo o vetor  $H \cdot x^T$ . Como  $S$  é uma aplicação linear sobrejetora de  $\mathbb{F}_q^N$  em  $\mathbb{F}_q^{N-k}$  e a síndrome de uma palavra código  $c$  é o vetor nulo  $\mathbf{0}$ ,  $S(c + e) = S(e)$ . A síndrome restrita aos vetores  $e \in \mathbb{F}_q^N$  tal que  $w_H(e) \leq t$ , denotada por  $\tilde{S}$ , é uma aplicação injetora. Assim os procedimentos para o algoritmo de decodificação por síndromes são: seja  $x$  uma mensagem recebida; suponha que  $x = c + e$  com  $c \in C$  e  $e \in \mathbb{F}_q^N$  o vetor erro; calcule  $S(x)$  (lembre que  $S(x) = S(e)$ ); se  $S(x) \notin \text{Im}(\tilde{S})$ , então  $x$  não pode ser decodificada; agora se  $S(x) \in \text{Im}(\tilde{S})$ , ou seja,  $w_H(e) \leq t$ , decodifique  $x$  como sendo o vetor  $x - \tilde{S}^{-1}(S(x)) = x - e$ .

Códigos em espaços de Hamming mostram-se eficientes para a correção de erros aleatórios, correção de erros em rajada aleatórias (bursts errors) e correção de rasuras aleatórias (erasures) (por exemplo, ver [8] e [11]). Em 1997 Rosenbloom e Tsfasman ([20]) apresentaram uma nova família de métricas (as chamadas métricas de

Rosenbloom-Tsfasman) eficientes para medir erros em rajadas localizadas (neste sentido podemos dizer que os erros são do tipo rasuras em rajadas). Já em 2004 Ozen e Siap ([15]) propuseram um esquema de decodificação por máxima verossimilhança para mensagens transmitidas em canais de comunicação onde os erros ocorrem em rajadas nas primeiras coordenadas e aleatoriamente nas últimas coordenadas, intercalando a métrica de Rosenbloom-Tsfasman com a métrica de Hamming. O decodificador de Ozen e Siap é construído a partir de uma forma padrão de matriz geradora para códigos em espaços de Rosenbloom-Tsfasman ([15], Theorem 1). Os códigos considerados em [15] são lineares sobre corpos finitos. Extensões dos resultados em [15] para anéis de Galois podem ser encontrados em [14] e [13]. Em [13], [14] e [15] não são determinados a capacidade de correção de erros (ou o raio de empacotamento) dos códigos em relação a métrica de Rosenbloom-Tsfasman. Também não são propostos algoritmos de decodificação.

Neste trabalho apresentamos uma nova família de métricas intercalando a métrica de blocos introduzida por Feng, Xu e Hickernell em [3] com a métrica de Rosenbloom-Tsfasman. Assim como Wei em [23] consideramos os pesos generalizados em relação a nova métrica. A partir dos pesos generalizados estendemos os resultados de Ozen e Siap: determinação de uma forma padrão de matriz geradora, limitante de Singleton generalizado, cálculo do espectro dos pesos generalizados, decodificação por máxima verossimilhança. Determinamos também o raio de empacotamento, o raio de cobertura (em função do último peso generalizado) e apresentamos uma classificação dos códigos perfeitos e quase-perfeitos. Uma proposta de algoritmo de decodificação via síndromes é apresentada. Ficam assim estabelecidos os conceitos fundamentais da teoria dos códigos corretores de erros em espaços de blocos de Rosenbloom-Tsfasman.

A métrica de blocos de Rosenbloom-Tsfasman é um caso particular de uma família de métricas introduzidas por Alves, Panek e Firer em [1]. Em [1] são caracterizados os grupos de simetrias lineares (em particular o grupo de simetrias lineares dos espaços de blocos de Rosenbloom-Tsfasman). As métricas em [1] são extensões das métricas

introduzidas por Brualdi, Graves e Lawrence em [2] (estudadas em [6], [7], [9], [16]) e Feng, Xu e Hickernell em [3] (estudadas em [10]). A métrica de Hamming é um caso particular das métricas de Rosenbloom-Tsfasman (estudadas em [13], [14], [15], [17], [18], [19]) que são casos particulares das métricas em [2].

## Espaços de Blocos de Rosenbloom-Tsfasman

Seja  $N$  um inteiro positivo e  $\pi = (\pi_1, \pi_2, \dots, \pi_n)$  uma partição de  $N$ :

$$N = \pi_1 + \pi_2 + \dots + \pi_n$$

com  $\pi_1 \geq \pi_2 \geq \dots \geq \pi_n \geq 1$ . Para cada  $\pi_i$ ,  $1 \leq i \leq n$ , seja  $V_i$  o espaço vetorial  $\mathbb{F}_q^{\pi_i}$  sobre o corpo finito  $\mathbb{F}_q$ . Definimos o espaço vetorial  $V$  como sendo a soma direta

$$V = V_1 \oplus V_2 \oplus \dots \oplus V_n.$$

Temos que  $V$  é isomorfo a  $\mathbb{F}_q^N$ . Cada vetor  $x \in V$  pode ser escrito de forma única como

$$x = x_1 + x_2 + \dots + x_n$$

com  $x_i \in V_i$  para cada  $1 \leq i \leq n$ . Definimos o **peso de blocos de Rosenbloom-Tsfasman**  $w_\pi$  (ou o  **$\pi$ -peso**) de  $x = x_1 + x_2 + \dots + x_n \in V$  pondo

$$w_\pi(x) = \begin{cases} \max \{i : x_i \neq \mathbf{0}\} & \text{se } x \neq \mathbf{0} \\ 0 & \text{se } x = \mathbf{0} \end{cases}.$$

O  $\pi$ -peso induz uma métrica em  $V$ , chamada **métrica de blocos de Rosenbloom-Tsfasman** (ou  **$\pi$ -métrica**): se  $x, y \in V$  então

$$d_\pi(x, y) = w_\pi(x - y).$$

Nestas condições diremos que  $(V, d_\pi)$  é o  **$\pi$ -espaço de Rosenbloom-Tsfasman**. A  **$\pi$ -distância mínima** de um código linear  $C$  é o número

$$d_\pi = d_\pi(C) = \min \{d_\pi(c, c') : c \neq c' \in C\}.$$

Associado ao peso de blocos de Rosenbloom-Tsfasman temos o peso generalizado de blocos de Rosenbloom-Tsfasman: se  $D \subseteq V$  é um subespaço vetorial, então

$$\|D\| = \max \{w_\pi(x) : x \in D\}$$

é o **peso generalizado de blocos de Rosenbloom-Tsfasman** (ou  $\pi$ -**peso generalizado**) de  $D$  e

$$d_r = d_r(C) = \min \{\|D\| : D \subseteq C, \dim(D) = r\}$$

é o  $r$ -**ésimo peso mínimo de blocos de Rosenbloom-Tsfasman** (ou  $r$ -**ésimo  $\pi$ -peso generalizado**) de um código linear  $C \subseteq V$ . Um código linear  $C \subseteq V$  de dimensão  $k$  com **hierarquia** de  $\pi$ -pesos generalizados  $(d_1, d_2, \dots, d_k)$  é dito um  $[N; k; d_1, \dots, d_k]$  **código linear**. Como  $d_\pi(C) = \min \{w_\pi(c) : \mathbf{0} \neq c \in C\}$  e  $\|D\| = w_\pi(c)$  para todo  $\mathbf{0} \neq c \in D$  sempre que  $D$  é 1-dimensional, então  $d_1(C) = d_\pi(C)$ .

A métrica de blocos de Rosenbloom-Tsfasman é um caso particular da  $(P, \pi)$ -métrica introduzida por Alves, Panek e Firer em [1], que corresponde a  $P$  totalmente ordenado. Quando a partição  $\pi$  de  $N$  satisfaz  $\pi_i = 1$  para todo  $1 \leq i \leq n$  a  $\pi$ -métrica coincide com a métrica de Rosenbloom-Tsfasman proposta em [20].

## Matriz Geradora

Começaremos mostrando a monotonicidade dos  $r$ -ésimos  $\pi$ -pesos generalizados.

**Teorema 1** *Seja  $C$  um  $[N; k; d_1, \dots, d_k]$  código linear. Então*

$$0 \leq d_1 \leq d_2 \leq \dots \leq d_k \leq n.$$

**Demonstração** Observe inicialmente que  $d_{r-1} \leq d_r$ . De fato, sejam  $D_{r-1}$  e  $D_r$  subcódigos de  $C$  de dimensão  $r-1$  e  $r$  respectivamente tal que  $\|D_{r-1}\| = d_{r-1}$  e  $\|D_r\| = d_r$ . Se  $\|D_{r-1}\| > \|D_r\|$ , então para todo subcódigo  $D'_{r-1} \subset D_r$  de dimensão  $r-1$  vale  $\|D'_{r-1}\| \leq \|D_r\| < \|D_{r-1}\| = d_r$ . Mas isto contradiz a minimalidade do número  $d_r$ .  $\square$

**Observação 1** Os números da hierarquia de  $\pi$ -pesos generalizados não são necessariamente distintos. Seja  $\pi = (2, 1, 1)$  uma partição de  $N = 4$  e considere o código binário  $C = \{0000, 1000, 0100, 1100\}$ . Neste caso

$$d_1(C) = d_2(C) = \min \{w_\pi(1000), w_\pi(0100), w_\pi(1100)\} = 1.$$

Seja  $(V, d_\pi)$  um  $\pi$ -espaço de Rosenbloom-Tsfasman. Para cada  $m$ -upla  $(\pi_{t_1}, \dots, \pi_{t_m})$ , definimos a aplicação projeção

$$p_{(t_1, \dots, t_m)} : V \rightarrow V$$

em relação aos blocos  $V_{t_1}, \dots, V_{t_m}$  da seguinte forma: dado  $u = u_1 + \dots + u_n \in V$ ,

$$p_{(t_1, \dots, t_m)}(u) = u_{t_1} + \dots + u_{t_m}.$$

**Teorema 2** Seja  $C$  um  $[N; k; d_1, \dots, d_k]$  código linear. Então  $C$  admite uma matriz geradora na forma

$$\begin{pmatrix} G_{s_m \pi_1} & \cdots & G_{s_m \pi_{t_1}} & G_{s_m \pi_{t_1+1}} & \cdots & G_{s_m \pi_{t_{m-1}}} & G_{s_m \pi_{t_{m-1}+1}} & \cdots & G_{s_m \pi_{t_m}} & \mathbf{0} \\ G_{s_{m-1} \pi_1} & \cdots & G_{s_{m-1} \pi_{t_1}} & G_{s_{m-1} \pi_{t_1+1}} & \cdots & G_{s_{m-1} \pi_{t_{m-1}}} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots \\ G_{s_1 \pi_1} & \cdots & G_{s_1 \pi_{t_1}} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} & \cdots & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

onde  $G_{st}$  são submatrizes de ordem  $s \times t$ , com  $s \leq t$ , e para cada  $1 \leq i \leq m$  o posto de  $G_{s_i \pi_{t_i}}$  é igual a  $s_i$ . Nestas condições  $s_1 + s_2 + \dots + s_m = k$ .

**Demonstração** Suponha que

$$\begin{aligned} d_1 = \dots = d_{s_1} < d_{s_1+1} = \dots = d_{s_1+s_2} < \dots \\ \dots < d_{s_1+\dots+s_{m-1}+1} = \dots = d_{s_1+\dots+s_{m-1}+s_m} = d_k. \end{aligned}$$

Nestas condições  $p_{d_{s_1+\dots+s_{j-1}+1}}(C) = \dots = p_{d_{s_1+\dots+s_{j-1}+s_j}}(C)$  para cada  $1 \leq j \leq m$ . Seja

$$D_j = p_{d_{s_1+\dots+s_{j-1}+1}}(C)$$

e  $\beta_j = \{v_{1d_j}, v_{2d_j}, \dots, v_{s_j d_j}\}$  uma base de  $D_j$ ,  $1 \leq j \leq m$ . Para cada  $v_{id_j}$  em  $\beta_j$  escolha  $c_{ij} \in C$  tal que

$$c_{ij} = v_{i1} + v_{i2} + \dots + v_{id_j}$$

em relação a soma direta  $V = V_1 \oplus V_2 \oplus \dots \oplus V_n$ , ou seja,  $v_{il} \in V_l$  para cada  $1 \leq l \leq d_j$ . Afirmamos que  $\widehat{\beta}_j = \{c_{1j}, c_{2j}, \dots, c_{s_j j}\}$  é linearmente independente. De fato, suponha que

$$\alpha_1 c_{1j} + \dots + \alpha_{s_j} c_{s_j j} = \mathbf{0}.$$

Isto implica que

$$\left( \sum_{i=1}^{s_j} \alpha_i v_{id_j} \right) + \left( \sum_{i=1}^{s_j} \sum_{t=1}^{d_j-1} \alpha_i v_{it} \right) = \mathbf{0}.$$

Como  $\sum_{i=1}^{s_j} \alpha_i v_{id_j} \in D_j$  e  $\{v_{1d_j}, v_{2d_j}, \dots, v_{s_j d_j}\}$  é uma base de  $D_j$ , não resta outra possibilidade além de  $\alpha_1 = \alpha_2 = \dots = \alpha_{s_j} = 0$ . Daí que  $\widehat{\beta}_j$  é linearmente independente.

Segue da monotonicidade  $d_{s_1} < d_{s_1+s_2} < \dots < d_{s_1+\dots+s_m} = d_k$  que

$$\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \dots \cup \widehat{\beta}_m$$

é um conjunto linearmente independente de  $k$  vetores de  $C$ , portanto uma base de  $C$ . A matriz  $G$  cujas linhas são os vetores da base  $\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \dots \cup \widehat{\beta}_m$  satisfaz as condições do teorema.  $\square$

**Definição 1** A forma matricial no teorema 2 será dita do **tipo**  $((s_1, t_1), \dots, (s_m, t_m))$ .

**Lema 1** Seja  $G$  uma matriz do tipo  $((s_1, t_1), \dots, (s_m, t_m))$ . Então  $G$  é a matriz geradora de um  $[N; \sum_{i=1}^m s_i]$  código linear.

**Demonstração** Sejam  $c_{1j}, c_{2j}, \dots, c_{s_j j}$  as linhas  $s_1 + \dots + s_{j-1} + 1, \dots, s_1 + \dots + s_{j-1} + s_j$  de  $G$  e  $\widehat{\beta}_j = \{c_{1j}, c_{2j}, \dots, c_{s_j j}\}$ . Como  $\text{posto}(G_{s_i \pi_{t_i}}) \neq 0$  concluímos que  $\widehat{\beta}_j$  é linearmente independente (como na demonstração do teorema 2). Da forma escada de  $G$  segue que  $\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \dots \cup \widehat{\beta}_m$  é linearmente independente. Portanto  $G$  gera um  $[N; \sum_{i=1}^m s_i]$  código

linear. □

**Lema 2** *Seja  $C$  um código gerado por uma matriz  $G$  do tipo  $((s_1, t_1), \dots, (s_m, t_m))$ . Se  $D$  é um subcódigo de  $C$ , então  $\|D\| = t_j$  para algum  $1 \leq j \leq m$ .*

**Demonstração** Sejam  $c_{1j}, c_{2j}, \dots, c_{s_j j}$  as linhas  $s_1 + \dots + s_{j-1} + 1, \dots, s_1 + \dots + s_{j-1} + s_j$  de  $G$ ,  $\widehat{\beta}_j = \{c_{1j}, c_{2j}, \dots, c_{s_j j}\}$  e  $C$  o  $[N; \sum_{i=1}^m s_i]$  código linear gerado por  $G$ . Como  $\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \dots \cup \widehat{\beta}_m$  é uma base de  $C$ , dado  $c \in C$  existem  $\alpha_{1j}, \dots, \alpha_{s_j j} \in \mathbb{F}_q$  tal que

$$c = \left( \sum_{i=1}^{s_1} \alpha_{i1} c_{i1} \right) + \dots + \left( \sum_{i=1}^{s_m} \alpha_{im} c_{im} \right).$$

Como

$$w_\pi \left( \sum_{i=1}^{s_j} \alpha_{ij} c_{ij} \right) = 0$$

se, e somente se,  $\alpha_{1j} = \dots = \alpha_{s_j j} = 0$  e

$$w_\pi \left( \sum_{i=1}^{s_j} \alpha_{ij} c_{ij} \right) = t_j$$

se algum  $\alpha_{ij} \neq 0$ , concluímos que  $w_\pi(c) = t_j$  para algum  $1 \leq j \leq m$  já que

$$w_\pi(c) = \max \{t_j : \alpha_{ij} \neq 0 \text{ para algum } 1 \leq i \leq s_j\}.$$

Assim, se  $D$  é um subespaço linear de  $C$ , então  $\|D\| = t_j$  para algum  $1 \leq j \leq m$ . □

**Teorema 3** *Seja  $G$  uma matriz do tipo  $((s_1, t_1), \dots, (s_m, t_m))$ . Então  $G$  é uma matriz geradora de um  $[N; \sum_{i=1}^m s_i]$  código linear  $C$  com hierarquia de  $\pi$ -pesos generalizados*

$$d_1 = \dots = d_{s_1} = t_1$$

$$d_{s_1+1} = \dots = d_{s_1+s_2} = t_2$$

⋮

$$d_{s_1+\dots+s_{m-1}+1} = \dots = d_{s_1+\dots+s_{m-1}+s_m} = t_m.$$

**Demonstração** Sejam  $c_{1j}, c_{2j}, \dots, c_{s_j j}$  as linhas  $s_1 + \dots + s_{j-1} + 1, \dots, s_1 + \dots + s_{j-1} + s_j$  de  $G$  e  $\widehat{\beta}_j = \{c_{1j}, c_{2j}, \dots, c_{s_j j}\}$ . Seja  $C$  o código gerado por  $\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \dots \cup \widehat{\beta}_m$ . Mostraremos que

$$d_{s_1 + \dots + s_{j-1} + 1}(C) = \dots = d_{s_1 + \dots + s_{j-1} + s_j}(C) = t_j.$$

para todo  $1 \leq j \leq m$ . Seja  $D$  um subespaço linear de  $C$  de dimensão  $s_1 + \dots + s_{j-1} + i$  com  $1 \leq i \leq s_j$ . Se  $\|D\| < t_j$  então  $w_\pi(c) < t_j$  para todo  $c \in D$ . Daí  $c$  é uma combinação linear dos vetores de  $\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \dots \cup \widehat{\beta}_{j-1}$ , o que implica em

$$D \subseteq \text{span}(\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \dots \cup \widehat{\beta}_{j-1}),$$

o que não é possível pois

$$\dim(\text{span}(\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \dots \cup \widehat{\beta}_{j-1})) = s_1 + s_2 + \dots + s_{j-1}.$$

Isto mostra que  $\|D\| \geq t_j$ . Assim  $d_{s_1 + \dots + s_{j-1} + i}(C) \geq t_j$ . Note agora que

$$D = \text{span}(\widehat{\beta}_1 \cup \widehat{\beta}_2 \cup \dots \cup \widehat{\beta}_{j-1} \cup \{c_{1j}, c_{2j}, \dots, c_{s_j j}\}) \subseteq C$$

é tal que  $\dim(D) = s_1 + \dots + s_{j-1} + i$  e  $\|D\| = t_j$ . Portanto  $d_{s_1 + \dots + s_{j-1} + i}(C) = t_j$  para cada  $1 \leq i \leq s_j$ .  $\square$

**Definição 2** Os códigos com matriz geradora do tipo  $((s_1, t_1), \dots, (s_m, t_m))$  serão chamados de **códigos do tipo**  $((s_1, t_1), \dots, (s_m, t_m))$ .

**Corolário 1** Seja  $C$  um código do tipo  $((s_1, t_1), \dots, (s_m, t_m))$ . Então  $C$  é um  $[N; \sum_{i=1}^m s_i]$  código linear com hierarquia de  $\pi$ -pesos generalizados

$$\begin{aligned} d_1 &= \dots = d_{s_1} = t_1 \\ d_{s_1+1} &= \dots = d_{s_1+s_2} = t_2 \\ &\vdots \\ d_{s_1+\dots+s_{m-1}+1} &= \dots = d_{s_1+\dots+s_{m-1}+s_m} = t_m. \end{aligned}$$



**Corolário 2** *Seja  $C$  um código do tipo  $((s_1, t_1), \dots, (s_m, t_m))$  e  $D_1, \dots, D_m$  subcódigos de  $C$  tal que  $\|D_j\| = t_j$  e  $\dim(D_j) = s_1 + \dots + s_j$  para todo  $1 \leq j \leq m$  (note que  $D_m = C$ ). Então*

$$D_1 \subset D_2 \subset \dots \subset D_m.$$

**Demonstração** *Seja  $G$  uma matriz geradora de  $C$  do tipo  $((s_1, t_1), \dots, (s_m, t_m))$ . Se  $\|D_j\| = t_j$  e  $\dim(D_j) = s_1 + \dots + s_j$ , então  $D_j$  é gerado pelas linhas  $1, 2, \dots, s_1 + \dots + s_j$  de  $G$ . O resultado segue agora da forma escada de  $G$ .  $\square$*

**Corolário 3** *Para um código do tipo  $((s_1, t_1), \dots, (s_m, t_m))$ ,*

$$\begin{aligned} d_1 &= \dots = d_{s_1} \leq n - m + 1 \\ d_{s_1+1} &= \dots = d_{s_1+s_2} \leq n - m + 2 \\ &\vdots \\ d_{s_1+\dots+s_{m-1}+1} &= \dots = d_{s_1+\dots+s_{m-1}+s_m} \leq n. \end{aligned}$$

**Demonstração** *Basta considerar a matriz geradora na forma  $((s_1, t_1), \dots, (s_m, t_m))$ .  $\square$*

**Corolário 4** (*Limitante de Singleton*) *Para um  $[N; k; d_1, \dots, d_k]$  código linear  $C$ ,*

$$d_\pi(C) \leq n - m + 1.$$

**Demonstração** *Basta observar que  $d_\pi(C) = d_1(C)$ .  $\square$*

Os seguintes fatos são conhecidos na literatura. Seja  $\begin{bmatrix} N \\ k \end{bmatrix}$  o número total de subespaços  $U \subseteq \mathbb{F}_q^N$  de dimensão  $k$ . Então

$$\begin{bmatrix} N \\ k \end{bmatrix} = \frac{(q^N - 1) \cdot \dots \cdot (q^N - q^{k-1})}{(q^k - 1) \cdot \dots \cdot (q^k - q^{k-1})}.$$

O número total de subespaços  $V \subseteq \mathbb{F}_q^N$  de dimensão  $t$  com  $U \subseteq V \subseteq \mathbb{F}_q^N$  é igual a

$$\begin{bmatrix} N - k \\ t - k \end{bmatrix}.$$

**Teorema 4** *Seja*

$$A_j^{(r)} = |\{D : D \subseteq C, \dim(D) = r, \|D\| = j\}|$$

o espectro de  $\pi$ -pesos generalizados de um código  $C$  do tipo  $((s_1, t_1), \dots, (s_m, t_m))$  e  $\sigma_i = s_1 + s_2 + \dots + s_i$ . Então

$$A_j^{(r)} = \begin{cases} (q^{s_i} - 1) q^{\sigma_{i-1}} \begin{bmatrix} \sigma_i - 1 \\ r - 1 \end{bmatrix} & \text{se } j = t_i \text{ e } 1 \leq r \leq \sigma_i \\ 0 & \text{caso contrário} \end{cases}.$$

**Demonstração** Do lema 2 segue que  $A_j^{(r)} = 0$  se  $j \neq t_i$  para todo  $1 \leq i \leq m$ .

Suponha que  $j = t_i$  para algum  $1 \leq i \leq m$ . Seja  $G$  uma matrix geradora de  $C$  do tipo  $((s_1, t_1), \dots, (s_m, t_m))$ . Existem  $q^{s_i} - 1$  possíveis combinações lineares não nulas geradas pelas linhas  $s_1 + \dots + s_{i-1} + 1, \dots, s_1 + \dots + s_{i-1} + s_i$  de  $G$  (vetores com  $\pi$ -peso igual a  $t_i$ ) e  $q^{s_1+s_2+\dots+s_{i-1}}$  possíveis combinações lineares (nulas ou não nulas) das linhas  $1, 2, \dots, s_1 + s_2 + \dots + s_{i-1}$  (vetores com  $\pi$ -peso menor ou igual a  $t_{i-1}$ ). Sendo assim existem

$$(q^{s_i} - 1) q^{s_1+s_2+\dots+s_{i-1}}$$

possíveis escolhas de vetores em  $C$  com  $\pi$ -peso igual a  $t_i$ . O número total de supespaços de dimensão  $r$  com  $\pi$ -peso generalizado igual a  $t_i$  será igual ao número total de subespaços de dimensão  $r$  contidos em  $p_{(1,2,\dots,t_i)}(C)$  que contêm algum vetor com coordenadas não nulas no  $t_i$ -ésimo bloco, ou seja, será igual a

$$(q^{s_i} - 1) q^{s_1+s_2+\dots+s_{i-1}} \begin{bmatrix} \sigma_i - 1 \\ r - 1 \end{bmatrix}.$$

□

## Decodificação por Máxima Verossimilhança

Seja  $(V, d_\pi)$  um  $\pi$ -espaço de Rosenbloom-Tasfasman. Dados  $x \in V$  e  $r$  um inteiro positivo, o conjunto de todos os vetores  $y \in V$  tal que  $d_\pi(x, y) \leq r$  é a chamada **bola** de centro  $x$  e raio  $r$ . A bola de centro  $x$  e raio  $r$  será denotado por  $B_\pi(x; r)$ .

**Lema 3** *Se  $y \in B_\pi(x; r)$  e  $w_\pi(x) > r$ , então  $w_\pi(y) = w_\pi(x)$ .*

**Demonstração** Se  $w_\pi(y) > w_\pi(x)$ , então  $w_\pi(x - y) = w_\pi(y) > r$  e daí  $y \notin B_\pi(x; r)$ . Agora se  $w_\pi(y) < w_\pi(x)$ , então  $w_\pi(x - y) = w_\pi(x) > r$  e novamente  $y \notin B_\pi(x; r)$ . A única possibilidade é  $w_\pi(y) = w_\pi(x)$ .  $\square$

O **raio de empacotamento**  $s$  de um código linear  $C$  de  $V$  é o maior inteiro positivo tal que as bolas centradas nas palavras códigos de  $C$  são duas a duas disjuntas:

$$s := \max \{r : B_\pi(c; r) \cap B_\pi(c'; r) = \emptyset \text{ para todo } c \neq c' \in C\}.$$

**Teorema 5** *O raio de empacotamento de um  $[N; k; d_1, \dots, d_k]$  código linear  $C$  é igual a*

$$s = d_\pi(C) - 1.$$

**Demonstração** Seja  $r := d_\pi(C) - 1$ . Provaremos que

$$B_\pi(\mathbf{0}; r) \cap B_\pi(c; r) = \emptyset$$

para todo  $\mathbf{0} \neq c \in C$ . Se  $c' \in B_\pi(\mathbf{0}; r) \cap B_\pi(c; r)$ , como  $w_\pi(c) \geq r + 1$  e  $c' \in B_\pi(c; r)$ , segue do último lema que  $w_\pi(c) = w_\pi(c')$ . Logo  $w_\pi(c') = w_\pi(c) \geq r + 1$  e  $c' \notin B_\pi(\mathbf{0}; r)$ , o que é uma contradição. Conseqüentemente o raio de empacotamento de  $C$  é  $r$ .  $\square$

O processo que associa cada mensagem  $x$  a uma palavra código  $c$  mais próxima de  $x$  (segundo a  $\pi$ -métrica) é chamado de **decodificação por máxima verossimilhança**. Segue do teorema 5 que se uma mensagem  $x$  contém no máximo  $s$  erros nas  $s$  primeiras coordenadas, então existe um única palavra código  $c$  mais próxima de  $x$ .

**Observação 2** A métrica de blocos de Rosenbloom-Tsfasman é uma **ultra-métrica**:

$$d_\pi(x, y) \leq \max \{d_\pi(x, z), d_\pi(z, y)\}$$

para todo  $z \in V$ . Em espaços ultra-métricos  $(X, d)$  todo triângulo é isósceles e conseqüentemente (ver [12]), dados  $x \neq y \in X$ ,

$$B_d(x; r) \cap B_d(y; r) = \emptyset \quad (1)$$

se, e somente se,  $r = d(x, y) - \varepsilon$  com  $\varepsilon > 0$ . Como a  $\pi$ -métrica é discreta, dado um código  $C \subseteq V$ , da não-arquimedianidade de  $(V, d_\pi)$  concluímos que  $r = d_\pi(C) - 1$  é máximo com a propriedade (1) se  $d_\pi(x, y) = d_\pi(C)$ . Daí que  $d_\pi(C) - 1$  é o raio de empacotamento de  $C$ .

A  $\pi$ -distância mínima (e conseqüentemente o raio de empacotamento) de um código linear  $C$  de  $V$  pode ser facilmente calculada se conhecemos uma matriz de paridade de  $C$ .

**Teorema 6** Seja  $(V, d_\pi)$  um  $\pi$ -espaço de Rosenbloom-Tsfasman. Seja  $C \subseteq V$  um  $[N; k]$  código linear e  $H = (H_1, H_2, \dots, H_n)$  uma matriz de paridade de  $C$  com  $H_i$  submatriz de ordem  $(N - k) \times \pi_i$  para cada  $1 \leq i \leq n$ . A distância mínima de  $C$  é igual a  $l + 1$  se, e somente se,

$$l = \max \{r : \text{posto}((H_1, H_2, \dots, H_r)) = r\}.$$

**Demonstração** Basta observar que cada vetor  $c = c_1 + c_2 + \dots + c_n \in C$  fornece uma dependência linear dos  $w_\pi(c)$  primeiros blocos de  $H = (H_1, H_2, \dots, H_n)$ :

$$\mathbf{0} = H \cdot c^T = \sum_{i=1}^n c_i \cdot H_i^T = \sum_{i=1}^{w_\pi(c)} c_i \cdot H_i^T.$$

□

**Corolário 5** *Seja  $C \subseteq V$  um código linear e  $H = (H_1, H_2, \dots, H_n)$  uma matriz de paridade de  $C$  com  $H_i$  submatriz de ordem  $(N - k) \times \pi_i$  para cada  $1 \leq i \leq n$ . O raio de empacotamento de  $C$  é igual a  $l$  se, e somente se,*

$$l = \max \{r : \text{posto}((H_1, H_2, \dots, H_r)) = r\}.$$

**Demonstração** O inteiro  $l$  é o raio de empacotamento de  $C$  se, e somente se,  $l = d_\pi(C) - 1$  se, e somente se,  $l + 1 = d_\pi(C)$ . Segue do teorema 6 que  $l + 1 = d_\pi(C)$  se, e somente se,  $l$  é o maior inteiro tal que  $\text{posto}((H_1, H_2, \dots, H_s)) = l$ .  $\square$

Diremos que um código  $C$  é  $\pi$ -**perfeito** se a união das bolas de raio  $s = d_\pi(C) - 1$  centradas nos vetores de  $C$  é igual a  $V$ . Um código  $C$  que atinge o limitante de Singleton, ou seja,  $d_\pi(C) = n - m + 1$ , é dito um **código MDS** (Maximum Separable Code).

**Teorema 7** *Seja  $(V, d_\pi)$  o  $\pi$ -espaço de Rosenbloom-Tsfasman tal que  $\pi_i = 1$  para todo  $1 \leq i \leq n$  e seja  $C$  um  $[N; k]$  código linear em  $V$ . Então  $C$  é MDS se, e somente se,  $C$  é  $\pi$ -perfeito.*

**Demonstração** Se  $C$  é MDS, então  $s = d_\pi(C) - 1 = N - k$ , o que implica que  $C$  é  $\pi$ -perfeito pois

$$B_\pi(c; s) = q^{N-k}$$

para todo  $c \in C$ . Agora se  $C$  é  $\pi$ -perfeito, então  $s = N - k$ . Por outro lado temos  $s = d_\pi(C) - 1$ . Logo  $d_\pi(C) = N - k + 1$  e portanto  $C$  é MDS.  $\square$

**Teorema 8** *(ver [1], Proposition 3.1) Seja  $(V, d_\pi)$  um  $\pi$ -espaço de Rosenbloom-Tsfasman e  $C \subseteq V$  um código linear. Então  $C$  is  $\pi$ -perfeito com raio de empacotamento  $r$  se, e somente se, existe uma transformação linear*

$$L : V_{r+1} \oplus \dots \oplus V_n \rightarrow V_1 \oplus \dots \oplus V_r$$

tal que

$$C = \{(L(v), v) : v \in V_{r+1} \oplus \dots \oplus V_n\}.$$

Como consequência dos teoremas 7 e 8 temos:

**Corolário 6** *Seja  $\pi$  a partição de  $N$  tal que  $\pi_i = 1$  para todo  $1 \leq i \leq n$  e  $C \subseteq V$  um  $[N; k]$  código linear. Então  $C$  é MDS se, e somente se, existe uma transformação linear*

$$L : \mathbb{F}_q^k \rightarrow \mathbb{F}_q^{N-k}$$

tal que

$$C = \{(L(v), v) : v \in \mathbb{F}_q^k\}.$$

Seja  $C$  um  $[N; k; d_1, \dots, d_k]$  código linear e  $s = d_\pi(C) - 1$ . Definimos o **raio de cobertura**  $\rho = \rho(C)$  como sendo o menor inteiro positivo  $l$  tal que  $V$  é a união das bolas de raio  $l$  centradas nos vetores de  $C$ . É claro que  $l \leq \rho(C)$  e  $l = \rho(C)$  se, e somente se,  $C$  é  $\pi$ -perfeito. Se o código não é  $\pi$ -perfeito, seu raio de cobertura é estritamente maior do que seu raio de empacotamento.

**Teorema 9** *Seja  $C$  um  $[N; k; d_1, \dots, d_k]$  código linear,  $p_{(r, r+1, \dots, n)}(C)$  a projeção de  $C$  em relação ao espaço  $V_r \oplus \dots \oplus V_n$  e*

$$l = \begin{cases} n + 1 & \text{se } p_n(C) \neq V_n \\ \min \{r : p_{(r, r+1, \dots, n)}(C) = V_r \oplus \dots \oplus V_n\} & \text{caso contrário} \end{cases}.$$

Então

$$\rho(C) = l - 1.$$

**Demonstração** Suponha inicialmente que  $p_n(C) \neq V_n$ . Isto implica que ou  $p_n(C) = \{\mathbf{0}\}$  ou  $\{\mathbf{0}\} \subset p_n(C) \subset V_n$ . Em ambos os casos  $w_\pi(x - c) = n$  para todo  $x \in V_n - p_n(C)$  e para todo  $c \in C$ . Isto mostra que não existe  $c \in C$  tal que  $x \in B_\pi(c; n - 1)$ . Concluimos assim que  $\rho(C) = n = (n + 1) - 1$ .

Suponha agora que  $p_{(r,r+1,\dots,n)}(C) = V_r \oplus \dots \oplus V_n$  para algum  $1 \leq r \leq n$ . Seja

$$l = \min \{r : p_{(r,r+1,\dots,n)}(C) = V_r \oplus \dots \oplus V_n\}.$$

Agora seja  $x = x_l + x_{n-l} \in V$  com  $x_l \in V_1 \oplus \dots \oplus V_{l-1}$  e  $x_{n-l} \in V_l \oplus \dots \oplus V_n$ . Como  $p_{(l,l+1,\dots,n)}(C) = V_l \oplus \dots \oplus V_n$  existe  $c \in C$  tal que

$$c = c_l + x_{n-l}.$$

Assim  $w_\pi(c - x) \leq l - 1$  e daí que  $x \in B_\pi(c; l - 1)$ . Isto implica que

$$V \subseteq \bigcup_{c \in C} B_\pi(c; l - 1).$$

Como ou  $p_{l-1}(C) = \{\mathbf{0}\}$  ou  $\{\mathbf{0}\} \subset p_{l-1}(C) \subset V_{l-1}$ , existe  $\mathbf{0} \neq y \in V_{l-1}$  tal que  $y \notin p_{l-1}(C)$ . Isto implica que  $w_\pi(c - y) \geq l - 1$  para todo  $c \in C$ , ou seja,

$$y \notin \bigcup_{c \in C} B_\pi(c; l - 2).$$

Portanto  $\rho(C) = l - 1$ . □

**Corolário 7** *Seja  $C$  um código linear do tipo  $((s_1, t_1), \dots, (s_m, t_m))$ . Então*

$$\rho(C) = \begin{cases} n & \text{se } t_m = n \text{ e } s_m < \pi_{t_m} \\ n & \text{se } t_m < n \\ n - l & \text{se } t_{m-i} = n - i \text{ e } s_{m-i} = \pi_{t_{m-i}} \text{ para todo } 0 \leq i \leq l - 1 \text{ e } s_{m-l} < \pi_{t_{m-l}} \end{cases}.$$

**Corolário 8** *Seja  $C$  um  $[N; k; d_1, \dots, d_k]$  código linear e  $l$  definido como no teorema 9.*

*Então*

$$\rho(C) = \begin{cases} l - 1 & \text{se } d_k = n \\ n & \text{se } d_k < n \end{cases}.$$

Seja  $(V, d_\pi)$  um  $\pi$ -espaço de Rosenbloom-Tsfasman. Se  $C$  é um código linear em  $V$  com raio de empacotamento  $s$  e raio de cobertura  $s + 1$ ,  $C$  é dito **quase-perfeito**. Apresentaremos agora uma caracterização dos códigos quase-perfeitos em espaços de blocos

de Rosenbloom-Tsfasman. Para os espaços de Hamming o problema da classificação dos códigos quase-perfeitos permanece aberto.

**Teorema 10** *Seja  $C$  um  $[N; k; d_1, \dots, d_k]$  código linear. Então  $C$  é quase-perfeito se, e somente se,*

$$p_{(d_1+1, \dots, n)}(C) = V_{d_1+1} \oplus V_{d_1+2} \oplus \dots \oplus V_n$$

e  $p_{d_1}(C) \neq V_{d_1}$ .

**Demonstração** Se  $p_{(d_1+1, \dots, n)}(C) = V_{d_1+1} \oplus V_{d_1+2} \oplus \dots \oplus V_n$  e  $p_{d_1}(C) \neq V_{d_1}$ , então  $\rho(C) = (d_1 + 1) - 1 = d_1 = t + 1$ . Logo  $C$  é quase-perfeito.

Agora seja  $x = x_{d_1+1} + \dots + x_n$  em  $V$  tal que  $x_j \in V_j$  para cada  $d_1 + 1 \leq j \leq n$ . Se  $C$  é quase-perfeito, existe  $c \in C$  tal que  $x \in B_\pi(c; s + 1)$ . Como

$$B_\pi(c; s + 1) = c + B_\pi(\mathbf{0}; s + 1),$$

$x = c + y$  com  $y \in B_\pi(\mathbf{0}; s + 1)$ . Agora como

$$B_\pi(\mathbf{0}; s + 1) = V_1 \oplus V_2 \oplus \dots \oplus V_{d_1},$$

concluimos que  $c = x - y$  é uma palavra código com  $p_{(d_1+1, \dots, d_n)}(c) = x$ . Isto implica que  $p_{(d_1+1, \dots, n)}(C) = V_{d_1+1} \oplus V_{d_1+2} \oplus \dots \oplus V_n$ . Não podemos ter  $p_{d_1}(C) = V_{d_1}$  (pois neste caso  $\rho(C) = s$ ). Assim  $p_{d_1}(C) \subset V_{d_1}$ .  $\square$

## Algoritmo de Decodificação via Síndromes

Se  $c$  é a palavra código transmitida e  $y = c + e$  é a mensagem recebida, então diremos que  $e$  é o **vetor erro** (ou simplesmente o **erro**). Fixada uma métrica, o procedimento que determina a palavra código mais próxima de uma mensagem recebida é conhecido como **decodificador por máxima verossimilhança**. A capacidade do



decodificador por máxima verossimilhança é determinada pelo raio de empacotamento do código.

Os decodificadores por máxima verossimilhança em espaços de Hamming são eficientes na correção de erros aleatórios. Já em espaços de blocos de Rosenbloom-Tsfasman os decodificadores por máxima verossimilhança são limitados a correções de erros em rajada nos primeiros blocos.

**Teorema 11** *Seja  $C$  um código linear,  $s = d_\pi(C) - 1$  e  $t = \lfloor \frac{d_H(C)-1}{2} \rfloor$ :*

(i) *Se na transmissão de uma palavra código  $c$  ocorrem até  $t$  erros aleatórios e recebemos a mensagem  $y$ , então  $c$  é a única palavra código tal que*

$$d_H(c, y) = \min \{d_H(c', y) : c' \in C\};$$

(ii) *Se na transmissão de uma palavra código  $c$  ocorrem até  $\pi_1 + \dots + \pi_s$  erros nas primeiras  $\pi_1 + \dots + \pi_s$  coordenadas e  $y$  é a mensagem recebida, então  $c$  é a única palavra código tal que*

$$d_\pi(c, y) = \min \{d_\pi(c', y) : c' \in C\}.$$

**Demonstração** O item (ii) é consequência do teorema 5. O item (i) é conhecido na literatura clássica (ver [5]). □

**Exemplo 1** *Seja  $C = \{00000, 11111\}$  um código binário. Em relação a métrica de Hamming,  $d_H(C) = 5$  e  $t = 2$ . Isto implica que  $C$  corrige até dois erros aleatórios. Os erros do tipo 11110, 01110, 10110, 11010 são corrigidos de forma incorreta. Agora suponha que  $\pi_i = 1$  para todo  $i$ . Em relação a métrica de blocos de Rosenbloom-Tsfasman,  $d_\rho(C) = 5$ ,  $s = 4$  e qualquer erro do tipo 11110, 01110, 10110, 11010 é corrigido corretamente. A desvantagem neste caso é que os erros do tipo 10001, 01001, 00101, 00011, 00001 são corrigidos incorretamente.*

O exemplo acima ilustra duas situações: se num determinado canal de comunicação a ocorrência de erros é aleatória, então a performance do código a ser implementado devem ser avaliados segundo a métrica de Hamming; agora se a ocorrência de erros não é aleatória e ocorre em rajadas (restrita as primeiras coordenadas), então a performance do código a ser implementado deve ser avaliado segundo a métrica de Rosenbloom-Tsfasman.

Convencidos de que os códigos em espaços de Rosenbloom-Tsfasman são eficientes para corrigir longos erros em rajada nas primeiras coordenadas, passamos a descrever um algoritmo para o decodificador por máxima verossimilhança em relação a métrica de blocos de Rosenbloom-Tsfasman.

Os próximos fatos são bem conhecidos na literatura e não dependem da estrutura métrica adotada em  $V$  (ver [5]). Seja  $C$  um código linear e  $H$  uma matriz de verificação de paridade de  $C$ . Em relação a síndrome  $S : V \rightarrow \mathbb{F}_q^{N-k}$ , dada por  $S(u) = H \cdot u^T$ , valem as seguintes propriedades:  $S(c) = S(c + e)$  para todo  $c \in C$  e para todo  $e \in V$ ;  $S : V \rightarrow \mathbb{F}_q^{N-k}$  é sobrejetora; se  $C_v = \{u \in V : S(u) = v\}$ , então  $C_u \cap C_v = \emptyset$  se  $u \neq v$ ;  $C_0 = C$ ;  $V = \bigcup_{v \in \mathbb{F}_q^{N-k}} C_v$ ;  $C_v = \{u + c : c \in C\}$ ;  $|C_v|$  não depende de  $v$ .

Os vetores com  $\pi$ -peso mínimo em  $C_v$  serão chamados de  $\pi$ -líderes de  $C_v$ .

**Teorema 12** *Seja  $(V, d_\pi)$  um  $\pi$ -espaço de Rosenbloom-Tsfasman e  $C$  um código linear em  $V$ . Se  $s = d_\pi(C) - 1$  e  $u \in V$  é um  $\pi$ -líder de  $C_v$  tal que  $w_\pi(u) \leq s$ , então  $u$  é o único  $\pi$ -líder de  $C_v$ .*

**Demonstração** Suponha que  $u' \neq u$  é também  $\pi$ -líder de  $C_v$ . Temos então que  $u' - u \neq \mathbf{0}$ ,  $u' - u \in C$  e

$$\begin{aligned} w_\pi(u' - u) &= \max \{i : u'_i - u_i \neq 0\} \\ &\leq \max \{w_\pi(u'), w_\pi(u)\} \\ &< r, \end{aligned}$$

uma contradição com a minimalidade de  $r$ . □

Sendo assim, se  $\hat{c}$  é a palavra recebida, sendo  $v = S(\hat{c})$  e  $u$  um  $\pi$ -líder de  $C_v$ , decodificamos  $\hat{c}$  como sendo o vetor

$$c = \hat{c} - u.$$

Se o erro é ocorre nos primeiros  $s$  blocos, então o teorema acima assegura que o algoritmo está retornando de fato a palavra transmitida.

**Exemplo 2** *Suponha que  $\pi_i = 1$  para todo  $i$ . Seja  $C = \{00000, 10100, 01010, 11110\}$  um código binário. Temos que  $C$  é um  $[5; 2]$  código com  $d_\pi(C) = 3$ . Como*

$$G = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \end{pmatrix} = (I_2 | A)$$

é uma matriz geradora de  $C$ , então

$$H = (A^T | I_3) = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

é uma matriz de verificação de paridade de  $C$ . Vamos listar agora os  $\pi$ -líderes e suas respectivas síndromes:

	Líder				Síndrome
$C$	00000	10100	01010	11110	000
	10000	00100	11010	01110	100
	01000	11100	00010	10110	010
	11000	01100	10010	00110	110
	00001	10101	01011	11111	001
	00011	10111	01001	11101	011
	10011	00111	11001	01101	111
	10001	00101	11011	01111	101

Suponha agora que em nosso canal de comunicação os erros mais comuns são do tipo 10000, 01000, 11000. Se recebemos a mensagem  $\hat{c} = 10010$ , como  $S(\hat{c}) = 110$ , decodificamos a mesma como sendo a palavra-código

$$c = \hat{c} - 11000 = 01010.$$

Por conta das características do canal de comunicação  $c = 01010$  é de fato a palavra-código enviada (o que equivale a dizer que em  $C_{110}$  o  $\pi$ -líder é único).

O que acontece se recebemos a mensagem  $\hat{c} = 10011$ ? Como  $s(\hat{c}) = 111$  e como 10011 é o  $\pi$ -líder de  $C_{111}$ , decodificamos então  $\hat{c}$  como sendo a palavra-código

$$c = \hat{c} - 10011 = 00000.$$

E se escolhêssemos 00111 como  $\pi$ -líder de  $C_{111}$ ? Neste caso decodificaríamos  $\hat{c} = 10011$  como sendo a palavra-código

$$c = \hat{c} - 00111 = 10100.$$

Esta situação desconfortável é gerada pelo fato de que em  $C_{111}$  existem muitos  $\pi$ -líderes (na verdade todos os elementos de  $C_{111}$  são  $\pi$ -líderes). Isto acontecerá com qualquer vetor situado nas últimas quatro linhas da tabela.

## Decodificador de Ozen-Siap

Nesta seção descreveremos um método de decodificação para informações transmitidas em canais de comunicação onde os erros ocorrem em rajadas nos primeiros blocos e aleatoriamente nos últimos blocos. Nosso método é uma extensão natural do método de decodificação proposto por Ozen e Siap em [15].

Como já sabemos, a métrica de blocos de Rosenbloom-Tsfasman é eficiente para a correção de erros em rajadas nos primeiros blocos. Para a correção de erros em blocos

aleatórios utilizaremos a métrica de blocos, introduzida recentemente por Feng, Xu e Hickernell em [3].

Seja  $\pi = (\pi_1, \pi_2, \dots, \pi_n)$  uma partição de um inteiro positivo  $N$  e  $V = V_1 \oplus V_2 \oplus \dots \oplus V_n$  a decomposição de  $\mathbb{F}_q^N$  como soma direta dos espaços  $V_i = \mathbb{F}_q^{\pi_i}$ ,  $1 \leq i \leq n$  (como na definição da métrica de blocos de Rosenbloom-Tsfasman). Dado  $x = x_1 + x_2 + \dots + x_n$  em  $V$ ,  $x_i \in V_i$  para cada  $1 \leq i \leq n$ , o **peso de blocos**  $w_\Pi$  de  $x$  é o número de blocos  $x_i$  não nulos de  $x$ :

$$w_\Pi(x) = |\{i : x_i \neq \mathbf{0}\}|.$$

Associado ao peso de blocos temos a **distância de blocos**  $d_\Pi$ : dados  $x, y \in V$ ,

$$d_\Pi(x, y) = w_\Pi(x - y).$$

A distância de blocos é um métrica em  $V$ . Assim como em espaços de Hamming, a capacidade de correção de erros de um código  $C$  em relação a métrica de blocos é dada em função da distância mínima (ver [3]): se

$$d_\Pi(C) = \min \{d_\Pi(c, c') : c \neq c' \in C\}$$

é a **distância mínima de blocos** de  $C$ , então

$$\left\lfloor \frac{d_\Pi(C) - 1}{2} \right\rfloor$$

é a sua **capacidade de correção de erros** (o raio de empacotamento de  $C$  no espaço métrico  $(V, d_\Pi)$ ). A métrica de blocos mostra-se eficiente para a correção de erros em blocos aleatórios (simultâneos erros em rajadas localizados).

Estamos prontos para descrever o decodificador de Ozen-Siap, que intercala a métrica de blocos de Rosenbloom-Tsfasman com a métrica de blocos.

**Teorema 13** *Seja  $C$  um  $((s_1, t_1), \dots, (s_m, t_m))$  código linear,  $C_{t_1} = p_{(1,2,\dots,t_1)}(C)$  e  $r = d_\pi(C_{t_1})$ . Se  $\dim(C_{t_1}) = \sum s_i$ , então  $C$  corrige até  $r - 1$  erros nos  $r - 1$  primeiros blocos*

e até  $n - t_1$  erros nos últimos  $n - t_1$  blocos. Se  $\dim(C_{t_1}) < \sum s_i$ , então existe uma matriz geradora de  $C$  da forma

$$\begin{pmatrix} G_1 & G_2 \\ \mathbf{0} & G_3 \end{pmatrix}$$

com  $\text{posto}(G_1) = \dim(C_{t_1})$  sendo  $G_2$  e  $G_3$  matrizes triangulares inferiores (em relação a diagonal secundária). Nestas condições, se  $G_3$  gera um código com distância mínima de blocos igual a  $s$ , então  $C$  corrige até  $r - 1$  erros nos primeiros  $r - 1$  blocos e até  $\lfloor \frac{s-1}{2} \rfloor$  erros nos últimos  $n - t_1$  blocos.

**Demonstração** Seja  $G$  uma matriz geradora de  $C$  do tipo  $((s_1, t_1), \dots, (s_m, t_m))$ . Sejam  $G'_1$  e  $G'_2$  as restrições de  $G$  às primeiras  $\pi_1 + \dots + \pi_{t_1}$  colunas e às últimas  $\pi_{t_1+1} + \dots + \pi_n$  colunas respectivamente. Nestas condições temos

$$G = \begin{pmatrix} G'_1 & G'_2 \end{pmatrix}$$

com  $G'_2$  triangular inferior (em relação a diagonal secundária). Denotamos as linhas de  $G'_1$  e  $G'_2$  respectivamente por  $g_{1,i}$  e  $g_{2,i}$ . Sejam  $c = c_1 + c_2$  a palavra código transmitida e  $\hat{c} = \hat{c}_1 + \hat{c}_2$  a mensagem recebida, onde  $c_1, \hat{c}_1 \in V_1 \oplus \dots \oplus V_{t_1}$  e  $c_2, \hat{c}_2 \in V_{t_1+1} \oplus \dots \oplus V_n$ . Se  $d_\pi(c_1, \hat{c}_1) \leq d_\pi(C_{t_1})$  (ou seja, ocorreram no máximo  $r - 1$  erros nos  $r - 1$  primeiros blocos de  $c$ ) decodificamos  $\hat{c}_1$  como sendo  $c_1$ . A decodificação de  $\hat{c}_2$  depende do posto de  $G'_1$ .

Se  $\text{posto}(G'_1) = \sum s_i$ , dado  $c = c_1 + c_2$  em  $C$ , se  $c_1 = \sum \alpha_i g_{1,i}$  então  $c_2 = \sum \alpha_i g_{2,i}$ . Desta forma decodificamos  $\hat{c}_2$  como sendo o vetor  $\sum \alpha_i g_{2,i}$ . A mensagem recebida  $\hat{c}$  é então decodificada corretamente se  $d_\pi(c_1, \hat{c}_1) \leq d_\pi(C_{t_1})$  e  $\text{posto}(G'_1) = \sum s_i$ .

Suponha agora que  $\text{posto}(G'_1) < \sum s_i$ , ou seja,  $\dim(C_{t_1}) < \sum s_i$ . Aplicando operações elementares em  $G$  obtemos a matriz geradora na forma

$$\begin{pmatrix} G_1 & G_2 \\ \mathbf{0} & G_3 \end{pmatrix}$$

com  $\text{posto}(G_1) = \dim(C_{t_1})$  sendo  $G_2, G_3$  triangulares inferiores (em relação a diagonal secundária). Sejam  $g_{1,i}, g_{2,i}, g_{3,i}$  as linhas de  $G_1, G_2$  e  $G_3$  respectivamente. Temos que  $G_1$  é uma matriz geradora do código  $C_{t_1} = p_{(1,2,\dots,t_1)}(C)$  com  $d_\pi(C_{t_1}) = r$ . Segue do teorema 11 que  $C$  corrige até  $r - 1$  erros nos  $r - 1$  primeiros blocos (já que  $d_\pi(C_{t_1}) \leq d_\pi(C)$ ). Suponha que  $c = c_1 + c_2$  é a palavra código transmitida e  $\hat{c} = \hat{c}_1 + \hat{c}_2$  é a mensagem recebida, onde  $c_1, \hat{c}_1 \in V_1 \oplus \dots \oplus V_{t_1}$  e  $c_2, \hat{c}_2 \in V_{t_1+1} \oplus \dots \oplus V_n$ . Se  $d_\pi(c_1, \hat{c}_1) \leq d_\pi(C_{t_1})$  decodificamos  $\hat{c}_1$  corretamente como sendo  $c_1$ . Nesta etapa temos  $\hat{c} = c_1 + \hat{c}_2$ . Como

$$c_1 = \sum \alpha_i g_{1,i},$$

$$c_2 = \sum \alpha_i g_{i,2} + \sum \beta_i g_{i,3}.$$

Daí  $\hat{c}_2 = \sum \alpha_i g_{i,2} + (\sum \beta_i g_{i,3} + e)$ , onde  $e \in V_{t_1+1} \oplus \dots \oplus V_n$  é o vetor erro. Assim  $\hat{c}_2 - \sum \alpha_i g_{i,2} = \sum \beta_i g_{i,3} + e$ . Se o número de blocos não nulos de  $e$  é menor ou igual a  $\lfloor \frac{s-1}{2} \rfloor$ , decodificamos  $\hat{c}_2 - \sum \alpha_i g_{i,2}$  como sendo a palavra código  $\sum \beta_i g_{i,3}$ . Daí temos  $\hat{c}_2 = c_2$ . Portanto  $\hat{c} = \hat{c}_1 + \hat{c}_2$  é decodificado corretamente:  $\hat{c} = c_1 + c_2$ .  $\square$

### Algoritmo para o Decodificador de Ozen-Siap

Finalizamos descrevendo um algoritmo simples para o decodificador de Ozen-Siap. Para o restante do texto  $C$  denotará um  $((s_1, t_1), \dots, (s_m, t_m))$  código linear,  $C_1 = p_{(1,2,\dots,t_1)}(C)$  e  $r = d_\pi(\tilde{C})$ .

**1o. Caso:**  $\dim(\tilde{C}) = \sum s_i$

Seja  $G = \begin{pmatrix} G'_1 & G'_2 \end{pmatrix}$  uma matriz geradora de  $C$  como na demonstração do teorema 13. Sejam  $g_{1,i}$  e  $g_{2,i}$  as linhas de  $G'_1$  e  $G'_2$  respectivamente. Para o algoritmo são requeridos: uma lista contendo um  $\pi$ -líder de cada classe lateral de  $C_1$ ; uma matriz de paridade de  $C_1$  (para o cálculo da síndrome  $S_1$  em relação a  $C_1$ ). Agora suponha que  $\hat{c} = \hat{c}_1 + \hat{c}_2$  é uma mensagem recebida onde  $\hat{c}_1 \in V_1 \oplus \dots \oplus V_{t_1}$  e  $\hat{c}_2 \in V_{t_1+1} \oplus \dots \oplus V_n$ .

Os procedimentos do algoritmo são:

- (1) Calcule  $S_1(\hat{c}_1)$ ;
- (2) Se  $S_1(\hat{c}_1) = \mathbf{0}$ , faça  $c_1 = \hat{c}_1$ ;
- (3) Se  $S_1(\hat{c}_1) = v \neq \mathbf{0}$ , faça  $c_1 = \hat{c}_1 - e_v$  onde  $e_v$  é um  $\pi$ -líder de  $(C_1)_v$ ;
- (4) Então, como  $c_1 = \sum \alpha_i g_{1,i}$ , faça  $c_2 = \sum \alpha_i g_{2,i}$ ;
- (5) Daí  $c = c_1 + c_2$  é a palavra transmitida e finalize o algoritmo.

**2o. Caso:**  $\dim(\tilde{C}) < \sum s_i$

Seja

$$G = \begin{pmatrix} G_1 & G_2 \\ \mathbf{0} & G_3 \end{pmatrix}$$

uma matriz geradora de  $C$  como no enunciado do teorema 13. Sejam  $g_{1,i}$ ,  $g_{2,i}$ ,  $g_{3,i}$  as linhas de  $G_1$ ,  $G_2$  e  $G_3$  respectivamente. Para o algoritmo são requeridos: uma lista contendo um  $\pi$ -líder de cada classe lateral de  $C_1$ ; uma lista contendo um  $\Pi$ -líder de cada classe lateral do código  $C_3$  gerado por  $G_3$ ; uma matriz de paridade de  $C_1$  (para o cálculo da síndrome  $S_1$  em relação a  $C_1$ ); uma matriz de paridade de  $C_3$  (para o cálculo da síndrome  $S_3$  em relação a  $C_3$ ). Agora suponha que  $\hat{c} = \hat{c}_1 + \hat{c}_2 + \hat{c}_3$  é uma mensagem recebida onde  $\hat{c}_1 \in V_1 \oplus \dots \oplus V_{t_1}$  e  $\hat{c}_2, \hat{c}_3 \in V_{t_1+1} \oplus \dots \oplus V_n$ . Os procedimentos do algoritmo são:

- (1) Calcule  $S_1(\hat{c}_1)$ ;
- (2) Se  $S_1(\hat{c}_1) = \mathbf{0}$ , faça  $c_1 = \hat{c}_1$ ;
- (3) Se  $S_1(\hat{c}_1) = v \neq \mathbf{0}$ , faça  $c_1 = \hat{c}_1 - e_v$  onde  $e_v$  é um  $\pi$ -líder de  $(C_1)_v$ ;
- (4) Então, como  $c_1 = \sum \alpha_i g_{1,i}$ , faça  $c_2 = \sum \alpha_i g_{i,2}$  e calcule  $S_3(\hat{c}_3)$ ;



- (5) Se  $S_3(\hat{c}_3) = \mathbf{0}$ , faça  $c_3 = \hat{c}_3$ ;
- (6) Se  $S_3(\hat{c}_3) = u \neq \mathbf{0}$ , faça  $c_3 = \hat{c}_3 - e_u$  onde  $e_u$  é um  $\Pi$ -líder de  $(C_3)_u$ ;
- (7) Daí  $c = c_1 + c_2 + c_3$  é a palavra transmitida e finalize o algoritmo.

## Referências

- [1] M.M.S. Alves, L. Panek and M. Firer, **Error-block codes and poset metrics**, Advances in Mathematics of Communications 2 (2008) 95-111.
- [2] R.A. Brualdi, J.S. Graves and K.M. Lawrence, **Codes with a poset metric**, Discrete Mathematics 147 (1995) 57-72.
- [3] K. Feng, L. Xu and F.J. Hickernell, **Linear error-block codes**, Finite Fields and Their Applications 12 (2006) 638-652.
- [4] R.W. Hamming, **Error Detecting and Error Correcting Codes**, Bell System Technical Journal 29 (1950) 147-160.
- [5] W.C. Huffman and V. Pless, **Fundamentals of Error-Correcting Codes**, Cambridge University Press (2003).
- [6] J.Y. Hyun and H.K. Kim, **Maximum distance separable poset codes**, Designs, Codes and Cryptography (2008), doi:10.1007/s10623-008-9204-8.
- [7] J.Y. Hyun and H.K. Kim, **The poset structures admitting the extended binary Hamming code to be a perfect code**, Discrete Mathematics 288 (2004) 37-47.
- [8] O. Keren and Simon Litsyn, **A Class of Array Codes Correcting Multiple Column Erasures**, IEEE Transactions on Information Theory 43 (1997) 1843-1851.

- [9] Y. Lee, **Projective systems and perfect codes with a poset metric**, Finite Fields and Their Applications 10 (2004) 105-112.
- [10] S. Ling and F. Özbudak, **Constructions and bounds on linear error-block codes**, Designs, Codes and Cryptography (2007), doi:10.1007/s10623-007-9119-9.
- [11] J.J. Metzner, **On Correcting Bursts (and Random Errors) in Vector Symbol  $(n, k)$  Cyclic Codes**, IEEE Transactions on Information Theory 54 (2008) 1795-1807.
- [12] L. Narici, **Functional Analysis and Valuation Theory**, Marcel Dekker, Inc., New York (1971).
- [13] M. Ozen and I. Siap, **Codes over Galois rings with respect to the Rosenbloom-Tsfasman metric**, Journal of the Franklin Institute 344 (2007) 790-799.
- [14] M. Ozen and I. Siap, **Linear Codes Over  $\mathbb{F}_q[u]/(u^s)$  with Respect to the Rosenbloom-Tsfasman Metric**, Designs, Codes and Cryptography 38 (2006) 17-29.
- [15] M. Ozen and I. Siap, **On the structure and decoding of linear codes with respect to Rosenbloom-Tsfasman metric**, Selçuk Journal of Applied Mathematics 5 (2) (2004) 25-31.
- [16] L. Panek, M. Firer, H.K. Kim and J.Y. Hyun, **Groups of linear isometries on poset structures**, Discrete Mathematics 308 (2008) 4116-4123.
- [17] L. Panek, M. Firer and M.M.S. Alves, **Symmetry groups of Rosenbloom-Tsfasman spaces**, Discrete Mathematics (2008), doi: 10.1016/j.disc.2008.01.013.
- [18] J. Quistorff, **On Rosenbloom and Tsfasman's generalization of the Hamming space**, Discrete Mathematics (2007), doi: 10.1016/j.disc.2007.01.005.

- [19] M.M. Skriganov, **On linear codes with large weights simultaneously for the Rosenbloom-Tsfasman and Hamming metrics**, Journal of Complexity (2007), doi: 10.1016/j.jco.2007.02.004.
- [20] M. Yu Rosenbloom and M. A. Tsfasman, **Codes for the  $m$ -metric**, Problems of Information Transmission 33 (1997) 45-52.
- [21] C. Shannon, **A mathematical theory of communication**, Bell System Technical Journal 27 (1948) 379-423 and 623-656.
- [22] D. Slepian, **A Class of Binary Signaling Alphabet**, Bell System Technical Journal 35 (1956) 203-234.
- [23] V.K. Wei, **Generalized Hamming Weights for Linear Codes**, IEEE Transactions on Information Theory 37 (1991) 1412-1418.

# MODELOS MATEMÁTICOS PARA DESCREVER O AUMENTO DE PESO EM GADO DE CORTE

Thiago Formehl, Maria José de Paula Castanho, Mikael Newmann.

Universidade Estadual do Centro-Oeste – UNICENTRO, [thiagoformehl@hotmail.com](mailto:thiagoformehl@hotmail.com)

Universidade Estadual do Centro-Oeste – UNICENTRO, [zeza@unicentro.br](mailto:zeza@unicentro.br)

Universidade Estadual do Centro-Oeste – UNICENTRO, [mikaelnewmann@hotmail.com](mailto:mikaelnewmann@hotmail.com)

## Resumo:

Este trabalho tem como objetivo testar os modelos clássicos para dinâmica populacional logístico e Gompertz, bem como o modelo de Von Bertalanffy, para descrever o aumento de peso em bovinos de corte, criados em regime de confinamento na fase terminal, utilizando-se quatro tratamentos. Apresentamos a descrição de cada um desses modelos e os métodos para determinação de seus parâmetros. Concluímos que os modelos logístico e de Von Bertalanffy são apropriados para descrever este crescimento.

**Palavras-chave:** Modelo Logístico, Modelo de Gompertz, Modelo de Von Bertalanffy.

## Introdução

De todo o rebanho zebuino brasileiro 80% é da raça Nelore. Nelore é uma raça rústica e se adapta com facilidade ao clima tropical brasileiro. Mas o gado europeu, melhor para a região sul do país, tem características como carne macia e precocidade. Para unir as características destas duas raças, faz-se o chamado “cruzamento industrial”, isto é, cruzamento de gado zebu com europeu (Giani, 2007).

A raça Canchim é composta de 5/8 do gado francês Charolês e 3/8 de Nelore. Esta raça vem sendo testada e aperfeiçoada, desde sua criação, na sua produtividade, adaptabilidade e eficiência reprodutiva. Como a produção no Estado do Paraná nem sempre tem obtido resultados satisfatórios, diferentes tratamentos têm sido testados na alimentação dos animais.

Nosso objetivo, neste trabalho, é descrever o aumento de peso em gado de corte, da raça Canchim, criado em regime de confinamento com tratamentos diferentes. Tendo

em vista que os modelos assintóticos-sigmóides se adaptam a estudos de crescimento de natureza biológica, estudaremos os modelos logístico, de Gompertz e de Von Bertalanffy aplicados a dados experimentais.

## **Materiais e Métodos**

Os dados foram obtidos experimentalmente no Núcleo de Produção Animal (NUPRAN) do Setor de Ciências Agrárias e Ambientais da Universidade Estadual do Centro-Oeste (UNICENTRO). O experimento foi realizado com 24 animais da raça Canchim, com idade média de 10 meses, distribuídos em 12 baias.

Foram utilizados quatro tratamentos para alimentação destes animais, a saber: Tratamento 1: ração RAB-Confinamento associado à silagem de milho com 25% de matéria seca; Tratamento 2: ração RAB-Confinamento associado à silagem de milho com 32% de matéria seca; Tratamento 3: ração RAB-E-15 associado com silagem de milho com 25% de matéria seca; e Tratamento 4: ração RAB-E-15 associado com 32% de matéria seca.

Os modelos utilizados para descrever o aumento de peso em cada tratamento foram o Logístico e o de Gompertz, tradicionais para a dinâmica de populações (Figueiredo *et al.*, 2007) e (Bassanezi, 2002), bem como o modelo de Von Bertalanffy, (Pienaar e Turnbull, 1973).

Para aplicação dos modelos efetuamos a média aritmética dos pesos dos animais submetidos a cada tratamento (Tabelas 1 e 2). Para sua avaliação consideramos o coeficiente de correlação linear ( $r$ ) e a soma dos quadrados dos resíduos ( $SQR$ ) (Levine *et al.*, 2000).

Utilizamos o programa CURVE EXPERT 1.3 para o ajuste de curvas. Os cálculos foram realizados com o auxílio do Microsoft EXCEL 2003 e os gráficos com o MATLAB 6.1.

Na próxima seção descrevemos um resumo de cada modelo e os respectivos métodos para determinação de seus parâmetros.

### **Modelo Logístico**

Em 1834, o matemático belga Pierre de Verhulst descreveu o crescimento das populações da França e da Bélgica por meio de um modelo matemático contínuo. Este

modelo, também conhecido como Logístico, foi utilizado mais tarde, em 1920, por Pearl e Reed para a população dos Estados Unidos da América.

O que Verhulst propôs era que a taxa de crescimento da população decresce linearmente, diferente do modelo de Malthus, de 1798, que considera esta mesma taxa constante. Matematicamente:

$$\frac{dP}{dt} = P(a - bP) \quad (1)$$

em que  $P(t)$  é a população no instante  $t$  e  $\lambda(P) = a - bP$  é a taxa de crescimento.

Podemos observar que as funções  $P(t) = 0$  e  $P(t) = \frac{a}{b}$  são soluções da equação

(1). Considerando  $P(t) \neq 0$ , a população atinge a estabilidade em  $P(t) = \frac{a}{b}$ , que chamamos de  $P_\infty$ , ou seja, a população limite.

Resolvendo a equação (1) considerando  $P(0) = P_0$ , temos:

$$P(t) = \frac{aP_0}{bP_0 + (a - bP_0)e^{-at}} \quad (2)$$

Substituindo  $P_\infty = \frac{a}{b}$  em (2) temos:

$$P(t) = \frac{P_\infty P_0}{P_0 + (P_\infty - P_0)e^{-at}} \quad \text{ou} \quad P(t) = \frac{P_\infty}{1 + \left(\frac{P_\infty}{P_0} - 1\right)e^{-at}} \quad (3)$$

Dessa forma, se  $t \rightarrow \infty$  então  $P(t) \rightarrow P_\infty = \frac{a}{b}$ , valor assintótico da população.

Segundo Bassanezi (2002), para encontrar a taxa de crescimento da população,  $\lambda(P) = a - bP$ , pode ser feito um ajuste linear aos pontos  $(P_i, \lambda_i)$  em que  $\lambda_i$  é dado por (4a) e  $P_i$  é a população média dos instantes  $i$  e  $i+1$  estimada pelo modelo exponencial (4b):

$$\lambda_i = \left( \frac{P_i}{P_{i-1}} \right)^{\frac{1}{i}} - 1 \quad (4a)$$

$$P_i = P_0 e^{\lambda_i t} \quad (4b)$$

No modelo logístico a população no ponto de inflexão (onde a taxa de crescimento é

máxima) é dada por  $P(t_m) = \frac{P_\infty}{2}$ , ou seja, a população é a metade da população limite.

### Modelo de Gompertz

O modelo de Gompertz, de 1825, é utilizado para descrever crescimentos celulares (Edelstein-Keshet, 1988) como em plantas, tumores e bactérias. Neste modelo a taxa de crescimento é dada por  $\lambda(P) = a - b \ln P$ .

Em linguagem matemática podemos escrever:

$$\frac{dP}{dt} = P(a - b \ln P) \quad (5)$$

Fazendo  $\frac{dP}{dt} = 0$  temos que  $P_\infty = e^{\frac{a}{b}}$ .

Considerando  $P(0) = P_0$  e  $P(t) \neq 0$  a solução para (5) é:

$$P(t) = e^{a/b} \exp\left[-\left(\frac{a}{b} - \ln P_0\right)e^{-bt}\right] \quad (6)$$

Substituindo  $P_\infty = e^{\frac{a}{b}}$  em (6) temos:

$$P(t) = P_\infty \left(\frac{P_0}{P_\infty}\right)^{e^{-bt}} \quad (7)$$

Para encontramos os parâmetros  $a$  e  $b$  da taxa de crescimento podemos fazer um ajuste linear dos valores  $(\ln P_i, \lambda_i)$ ,  $i = 1, 2, \dots, n$ , em que  $\lambda_i$  é dado por (8) e  $\ln P_i$  é o logaritmo natural da população no instante  $i$ .

$$\lambda_i = \frac{P_{i+1} - P_i}{(t_{i+1} - t_i)P_i} \quad (8)$$

O ponto de inflexão no modelo de Gompertz ocorre quando  $P(t_m) = \frac{P_\infty}{e}$ .

### Modelo de Von Bertalanffy

Von Bertalanffy derivou este modelo do estudo das relações alométricas nos organismos (Pienaar e Turnbull, 1973). Utilizado no início dos anos 70 para descrever o aumento de peso em peixes em função do tempo (Bassanezi, 2002), o modelo de Von Bertalanffy considera além da relação alométrica do peso com a área corporal, a taxa de

anabolismo (taxa de síntese de massa por unidade de superfície do animal) e a taxa de catabolismo (taxa de diminuição da massa por unidade de massa).

Podemos escrever esta equação de modo generalizado:

$$\frac{dP}{dt} = \alpha P^\gamma - \beta P \quad (9)$$

em que  $\alpha$  é a constante de anabolismo;  $\beta$  é a constante de catabolismo e  $\gamma$  é a relação alométrica.

No modelo original de Von Bertalanffy, o parâmetro alométrico é igual a 2/3, porém Oliveira (2007) obteve bons resultados com  $\gamma = \frac{3}{4}$ , que é a relação alométrica para mamíferos, para descrever o crescimento de suínos de corte. Assim, tomaremos

$\gamma = \frac{3}{4}$  neste artigo.

Quando  $t \rightarrow \infty$ ,  $\frac{dP}{dt} = 0$  e  $P(t) \rightarrow P_\infty$ . Temos então, de (9),  $P_\infty = \left(\frac{\alpha}{\beta}\right)^4$ .

Tomando  $P(0) = P_0$ , resolvendo a equação (9), e substituindo  $P_\infty = \left(\frac{\alpha}{\beta}\right)^4$ ,

obtemos:

$$P(t) = P_\infty \left\{ 1 + \left[ \left( \left( \frac{P_0}{P_\infty} \right)^{\frac{1}{4}} - 1 \right) e^{-\frac{\beta t}{4}} \right]^4 \right\} \quad (10)$$

Ao invés de consideramos  $\beta$  como uma constante, tendo em vista que  $\beta$  varia com o tempo, encontraremos uma função que descreva esta variação. Para isso, basta isolarmos  $\beta$  na equação (10),

$$\beta [P(t), t] = -\frac{4}{t} \ln \left[ \frac{\left( \frac{P(t)}{P_\infty} \right)^{\frac{1}{4}} - 1}{\left( \frac{P_0}{P_\infty} \right)^{\frac{1}{4}} - 1} \right] \quad (11)$$

e, de posse dos dados experimentais, construir um diagrama de dispersão  $(t_i, \beta_i(t))$  e ajustar uma curva que descreva o comportamento destes pontos.



## Resultados e Discussão

Para obtenção dos modelos utilizamos os dados das Tabelas 1 e 2, em que o peso é a média aritmética dos pesos dos animais submetidos a cada tratamento, depois de  $t$  dias. Na Tabela 1, não consideramos o peso no período de adaptação dos animais ao confinamento.

O primeiro modelo testado foi o logístico. Para este modelo adotamos para  $t = 0$  o primeiro peso após o período de adaptação dos animais ao tratamento.

Tabela 1: Média dos pesos em cada tratamento após a adaptação

Tempo (dias)	Tratamento 1	Tratamento 2	Tratamento 3	Tratamento 4
	Peso (kg)	Peso (kg)	Peso (kg)	Peso (kg)
0	332,333	341	334,667	335,667
21	369,333	388,333	372,333	374
42	417,333	429	410,333	410,833
63	456,833	467	447,833	442,5
84	485,5	496,5	481,5	470,5
103	517,333	527,667	508,167	493,833

Para encontrarmos a taxa de crescimento do modelo logístico utilizamos (4a) e (4b). As retas ajustadas e os respectivos coeficientes de correlação são dados por:

$$\lambda_1(P) = 0,01143214 - 0,000016599594P \quad r_1=0.83328589$$

$$\lambda_2(P) = 0,013795996 - 0,000021614082P \quad r_2=0.95781835$$

$$\lambda_3(P) = 0,010796883 - 0,000015847349P \quad r_3=0.99131347$$

$$\lambda_4(P) = 0,012720691 - 0,000021286375P \quad r_4=0.99603972$$

em que  $\lambda_j$  com  $j = 1, 2, 3$  e  $4$  corresponde à taxa de crescimento dos animais submetidos ao tratamento  $j$ .

Substituindo as equações acima em (3), temos os modelos logísticos que descrevem o aumento de peso em gado segundo cada tratamento. Estes modelos estão representados graficamente na Figura 1.

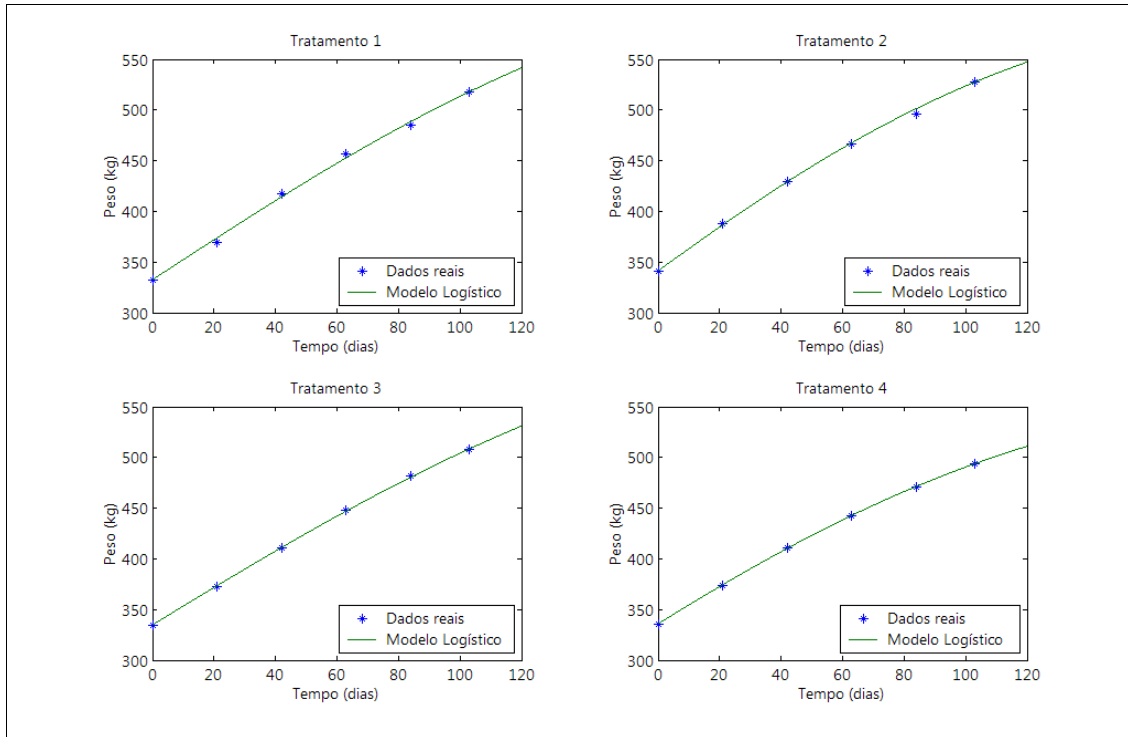
$$P_1(t) = \frac{688.7}{1,0723e^{-0,0114t} + 1}$$

$$P_2(t) = \frac{638,287}{0,8718e^{-0,0138t} + 1}$$

$$P_3(t) = \frac{681,305}{1,0358e^{-0,0108t} + 1}$$

$$P_4(t) = \frac{597,598}{0,7803e^{-0,0127t} + 1}$$

Figura 1: Modelo logístico para descrever o peso de bovinos de corte, submetidos a 4 tratamentos.



Para obtermos a taxa de crescimento do modelo de Gompertz efetuamos o ajuste linear aos pontos  $(\ln P_i, \lambda_i)$  em que  $P_i$  é o peso no instante  $i$ , constante da Tabela 1, e  $\lambda_i$  é dado por (8). Obtivemos as retas ajustadas para cada tratamento e os respectivos coeficientes de correlação:

$$\lambda_1(P) = 0,047653393 - 0,0071796808 \ln P \quad r_1 = 0.84588755$$

$$\lambda_2(P) = 0,061398098 - 0,0094297214 \ln P \quad r_2 = 0.97215670$$

$$\lambda_3(P) = 0,044396098 - 0,006690154 \ln P \quad r_3 = 0.98753496$$

$$\lambda_4(P) = 0,056100903 - 0,087013888 \ln P \quad r_4 = 0.99680635$$

em que  $\lambda_j$  com  $j = 1, 2, 3$  e  $4$  corresponde à taxa de crescimento dos animais submetidos ao tratamento  $j$ .

Substituindo na equação (7), temos os modelos de Gompertz para cada tratamento. Estes modelos estão representados graficamente na Figura 2.

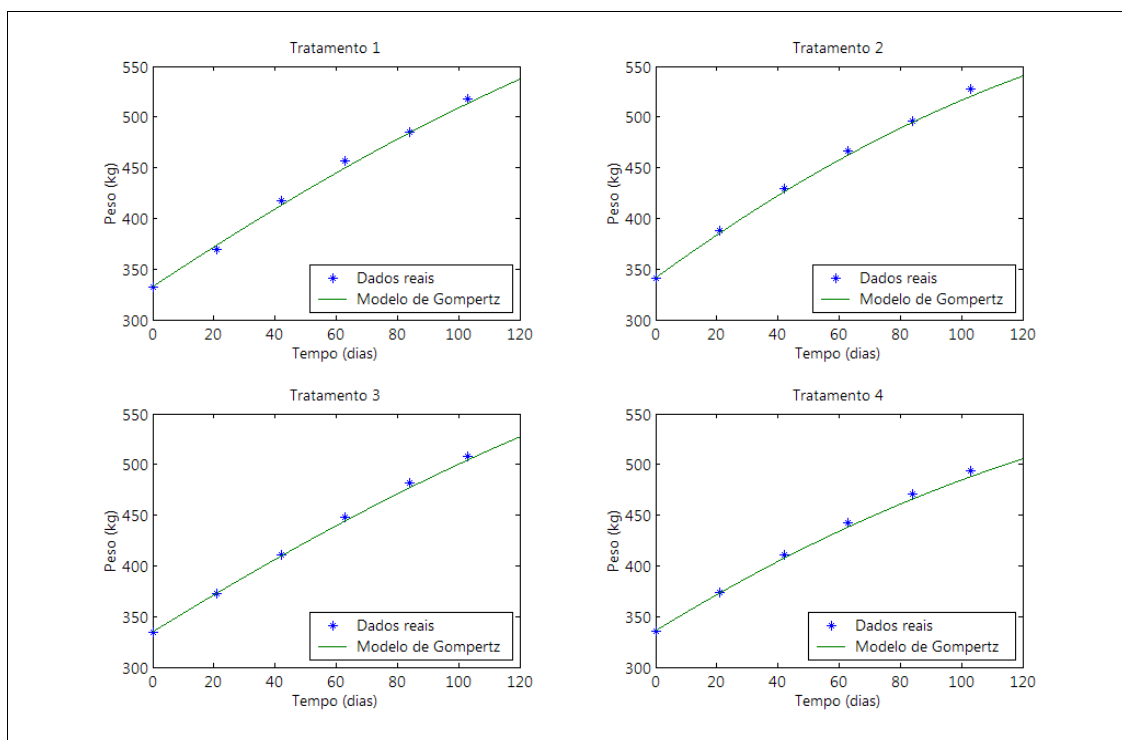
$$P_1(t) = 763(0,4356)e^{-0,0072t}$$

$$P_2(t) = 672,583(0,5070)e^{-0,0094t}$$

$$P_3(t) = 762,059(0,4392)e^{-0,0067t}$$

$$P_4(t) = 631,026(0,5319)e^{-0,0087t}$$

Figura 2: Modelo de Gompertz para descrever o peso de bovinos de corte, submetidos a 4 tratamentos.



Para aplicação do modelo de Von Bertalanffy optamos por tomar os valores do peso limite encontrados para cada tratamento nos modelos logístico e de Gompertz e, ainda, incluímos o período de adaptação ao confinamento, conforme dados contidos na Tabela 2.

Apresentamos, primeiramente, os resultados encontrados para o Modelo de Von Bertalanffy com peso limite igual ao do modelo logístico. Para encontrarmos  $\beta(t)$ , ou seja, o parâmetro que descreve o catabolismo em relação ao tempo, ajustamos uma curva logística aos pontos  $(t_i, \beta_i)$  em que  $t_i$  é o tempo em dias e  $\beta_i$  é dado por (11).

Tabela 2: Média dos pesos de cada tratamento com o período de adaptação

Tempo (dias)	Tratamento 1	Tratamento 2	Tratamento 3	Tratamento 4
	Peso (kg)	Peso (kg)	Peso (kg)	Peso (kg)
0	320,5	320,5	321	321,5
15	332,333	341	334,667	335,667
36	369,333	388,333	372,333	374
57	417,333	429	410,333	410,833
78	456,833	467	447,833	442,5
99	485,5	496,5	481,5	470,5
118	517,333	527,667	508,167	493,833

Os valores de  $\beta_j(t)$ ,  $j = 1, 2, 3$  e  $4$ , referente aos 4 tratamentos, e os respectivos coeficientes de correlação são dados por:

$$\beta_1(t) = \left( \frac{0,0308}{3,9381e^{-0,0596t} + 1} \right) \quad r_1 = 0.99765140$$

$$\beta_2(t) = \left( \frac{0,0397}{1,5724e^{-0,0546t} + 1} \right) \quad r_2 = 0.98254623$$

$$\beta_3(t) = \left( \frac{0,0307}{1,2522e^{-0,0311t} + 1} \right) \quad r_3 = 0.99641865$$

$$\beta_4(t) = \left( \frac{0,0371}{2,4347e^{-0,0565t} + 1} \right) \quad r_4 = 0.99641865$$

Substituindo na equação (10), temos os Modelos de Von Bertalanffy com peso limite igual ao utilizado no modelo logístico, para cada tratamento. Estes modelos estão representados graficamente na Figura 3.

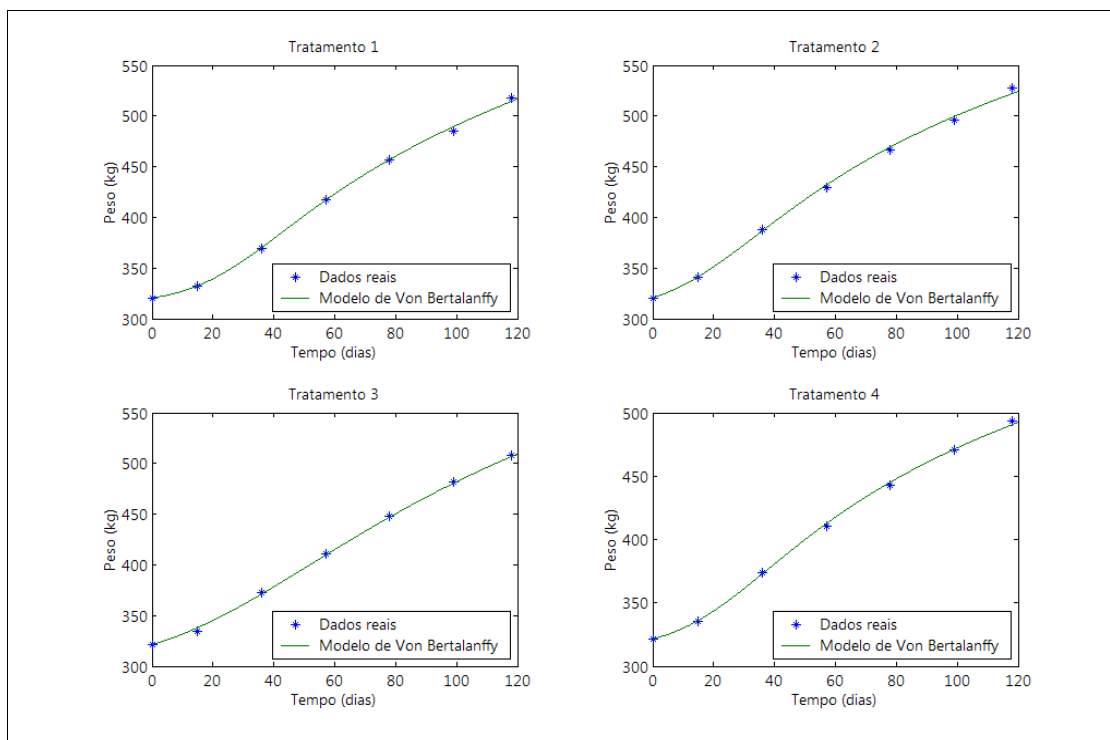
$$P_1(t) = 688,7 \left\{ 1 - 0,1740 \exp \left[ - \frac{1}{4} \left( \frac{0,0308}{3,9381e^{-0,0596t} + 1} \right) t \right] \right\}^4$$

$$P_2(t) = 638,287 \left\{ 1 - 0,1582 \exp \left[ - \frac{1}{4} \left( \frac{0,0397}{1,5724e^{-0,0546t} + 1} \right) t \right] \right\}^4$$

$$P_3(t) = 681,305 \left\{ 1 - 0,1715 \exp \left[ - \frac{1}{4} \left( \frac{0,0307}{1,2522e^{-0,0311t} + 1} \right) t \right] \right\}^4$$

$$P_4(t) = 597,598 \left\{ 1 - 0,1436 \exp \left[ - \frac{1}{4} \left( \frac{0,0371}{2,4347e^{-0,0565t} + 1} \right) t \right] \right\}^4$$

Figura 3: Modelo de Von Bertalanffy com peso limite do modelo logístico para descrever o peso de bovinos de corte, submetidos a 4 tratamentos.



Agora, utilizando o peso limite encontrado no modelo de Gompertz e os valores de  $\beta_j(t)$ ,  $j = 1, 2, 3$  e  $4$ , referente aos 4 tratamentos, (com seus respectivos coeficientes de correlação) dados por:

$$\beta_1(t) = \left( \frac{0,0250}{3,9444e^{-0,0655t} + 1} \right) \quad r_1 = 0,99792157$$

$$\beta_2(t) = \left( \frac{0,0345}{1,7383e^{-0,07t} + 1} \right) \quad r_2 = 0,98906446$$

$$\beta_3(t) = \left( \frac{0,0236}{2,3696e^{-0,0597t} + 1} \right) \quad r_3 = 0,99711051$$

$$\beta_4(t) = \left( \frac{0,0318}{2,5232e^{-0,0665t} + 1} \right) \quad r_4 = 0,99692053$$

temos os modelos de Von Bertalanffy representados graficamente na Figura 4.

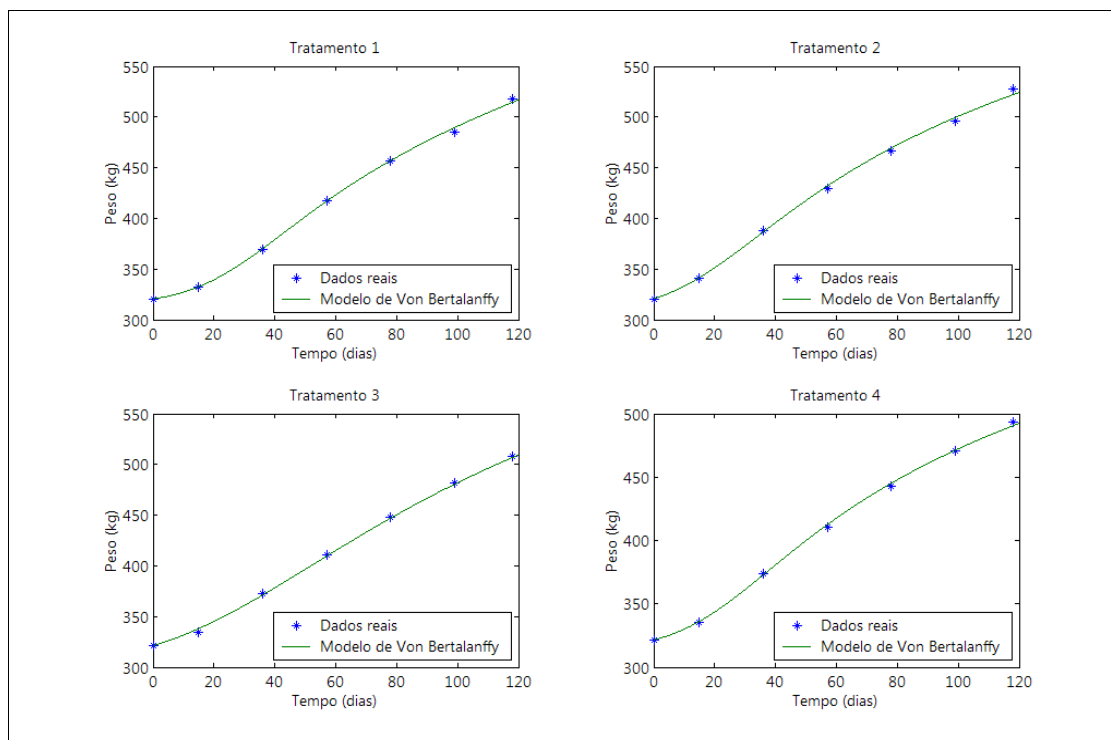
$$P_1(t) = 763 \left\{ 1 - 0,1949 \exp \left[ - \frac{1}{4} \left( \frac{0,0250}{3,9444e^{-0,0655t} + 1} \right) \right] t \right\}^4$$

$$P_2(t) = 672,583 \left\{ 1 - 0,1691 \exp \left[ - \frac{1}{4} \left( \frac{0,0345}{1,7383e^{-0,07t} + 1} \right) \right] t \right\}^4$$

$$P_3(t) = 762,059 \left\{ 1 - 0,1944 \exp \left[ - \frac{1}{4} \left( \frac{0,0236}{2,3696e^{-0,0597t} + 1} \right) \right] t \right\}^4$$

$$P_4(t) = 631,026 \left\{ 1 - 0,1551 \exp \left[ - \frac{1}{4} \left( \frac{0,0318}{2,5232e^{-0,0665t} + 1} \right) \right] t \right\}^4$$

Figura 4: Modelo de Von Bertalanffy com peso limite do modelo de Gompertz para descrever o peso de bovinos de corte, submetidos a 4 tratamentos.



Para compararmos os modelos, encontramos as somas dos quadrados dos resíduos, isto é, a soma dos quadrados das diferenças entre os valores reais e os obtidos por cada modelo para cada tratamento. Estes dados estão descritos na Tabela 3.

Tabela 3: Soma dos quadrados dos resíduos para cada modelo e tratamento

SQR	Von Bertalanffy L	Von Bertalanffy G	Logístico	Gompertz
Tratamento 1	24,715	18,512	55,341	108,744
Tratamento 2	57,367	38,241	26,435	103,879
Tratamento 3	18,597	9,722	3,176	53,556
Tratamento 4	21,825399	11,915833	1,24796	94,138095

## **Conclusão**

Comparando os modelos logístico e de Gompertz observamos que, apesar dos ajustes das taxas de crescimento no modelo de Gompertz apresentarem um melhor coeficiente de correlação na maioria dos casos, o modelo logístico apresentou uma distância menor dos valores reais em todos os tratamentos. Já nos modelos de Von Bertalanffy, o modelo com peso limite igual ao utilizado no modelo de Gompertz apresentou melhores coeficientes de correlação e menor distância dos valores reais em todos os tratamentos.

Assim, podemos concluir que os modelos logístico e de Von Bertalanffy (com peso limite igual ao do modelo de Gompertz) são apropriados para descrever o aumento de peso em gado de corte da raça Canchim com qualquer um dos tratamentos testados. A vantagem do modelo de Von Bertalanffy é ter se ajustado a todos os dados incluindo a fase de adaptação, porém, o modelo logístico apresentou menor soma de quadrados dos resíduos.

Como trabalhos futuros pretendemos reavaliar as taxas de crescimento em cada modelo e verificar se existe uma única curva que descreva os 4 tratamentos de forma satisfatória, ou seja, se há ou não diferença significativa entre os modelos que descrevem os tratamentos.

## **Referências Bibliográficas:**

- BASSANEZI, R. C.: **Ensino-Aprendizagem com Modelagem Matemática**, Contexto, São Paulo, (2002).
- EDELSTEIN-KESHET, L. **Mathematical Models in Biology**. McGraw-Hill, Inc. New York, (1988).
- FIGUEIREDO, D. G. e NEVES, A. F.: **Equações Diferenciais Aplicadas**, IMPA, Rio de Janeiro, 3rd ed, (2007).
- GIANI, G.: *Canchim – o precoce brasileiro*, Revista Tecnologias de Produção HFF&Citrus, ano IV, ed 22, (2007), 22-29.
- LEVINE, D. M; BERENSON, M. L.; STEPHAN, D.: **Estatística: Teoria e Aplicações**, LTC, Rio de Janeiro 1st ed, (2000).
- OLIVEIRA, L.; BRANDÃO, A. J. V.; BASSANEZI, R. C.: **Modelo de Von Bertalanffy Generalizado Aplicado ao Crescimento de Suínos de Corte**, Biomatemática, v. 17, UNICAMP, (2007), 101-109.

PIENAAR, L.V. e TURNBULL. K.J.: **The Chapman-Richards Generalization of Von Bertalanffy's Growth Model for Basal Area Growth and Yield in Even-Aged Stands**, Forest Science, vol 9(1), (1973), 2-21.



# **SIMULAÇÃO NUMÉRICA DO PROCESSO DE VAZAMENTO E SOLIDIFICAÇÃO DA LIGA Al-22%Ni EM MOLDE METÁLICO**

Luiz Cláudio Benine, Moisés Meza Pariona, Guilherme Anderson Salem, Josuel Kroppa Rogenski.

Universidade Estadual de Ponta Grossa – UEPG, [luiz.benine@yahoo.com.br](mailto:luiz.benine@yahoo.com.br);  
Universidade Estadual de Ponta Grossa – UEPG, [mmpariona@uepg.br](mailto:mmpariona@uepg.br); Universidade  
Estadual de Ponta Grossa – UEPG, [guilhermesalem@gmail.com](mailto:guilhermesalem@gmail.com); Universidade Estadual  
de Ponta Grossa – UEPG, [josuelkr@yahoo.com.br](mailto:josuelkr@yahoo.com.br).

## **Resumo:**

O presente trabalho teve por finalidade realizar a simulação numérica do vazamento e da solidificação da liga Al-22%Ni para o processo da fundição. Este estudo foi realizado através do método dos elementos finitos (MEF) mediante o software comercial ANSYS®. O vazamento foi realizado sob gravidade, também foi considerada a turbulência nesta etapa e logo após foi feito o processo de solidificação. Como resultado deste estudo foi mostrado o comportamento durante o vazamento e o campo térmico, o fluxo térmico e o gradiente térmico durante a solidificação.

**Palavras-chave:** elementos finitos, vazamento, solidificação.

## **Introdução**

O processo de fundição de ligas é muito difícil de modelar devido à complexidade dos fenômenos físicos envolvidos. Entre as etapas que se apresentam na fundição estão os processos de vazamento (mecânica de fluidos) e solidificação. Nestes processos apresentam-se o preenchimento do molde pelo metal líquido, o campo térmico, campo de pressão, a transferência de calor entre o metal fundido e o molde, a mudança de fase, o crescimento dendrítico, segregação, formação de vazios, porosidade, trincas, gotas frias, além de tensões térmicas no metal fundido e no molde. Cada um destes fenômenos é um tema de interesse para pesquisa. Se forem considerados simultaneamente alguns destes fenômenos citados a análise do processo da solidificação se tornaria um problema de extrema dificuldade. Outro fator que complica ainda mais é a necessidade da modelagem tridimensional, pois tipicamente uma indústria usa moldes

com geometrias complexas. Para viabilizar a simulação da solidificação de uma indústria de fundição é necessário simplificar os fenômenos apresentados anteriormente.

## Materiais e Métodos

Inicialmente foi construída a geometria do forno e do molde em duas dimensões como mostra a Figura 1. Logo em seguida foi gerada a malha do sistema como indicado na Figura 2.

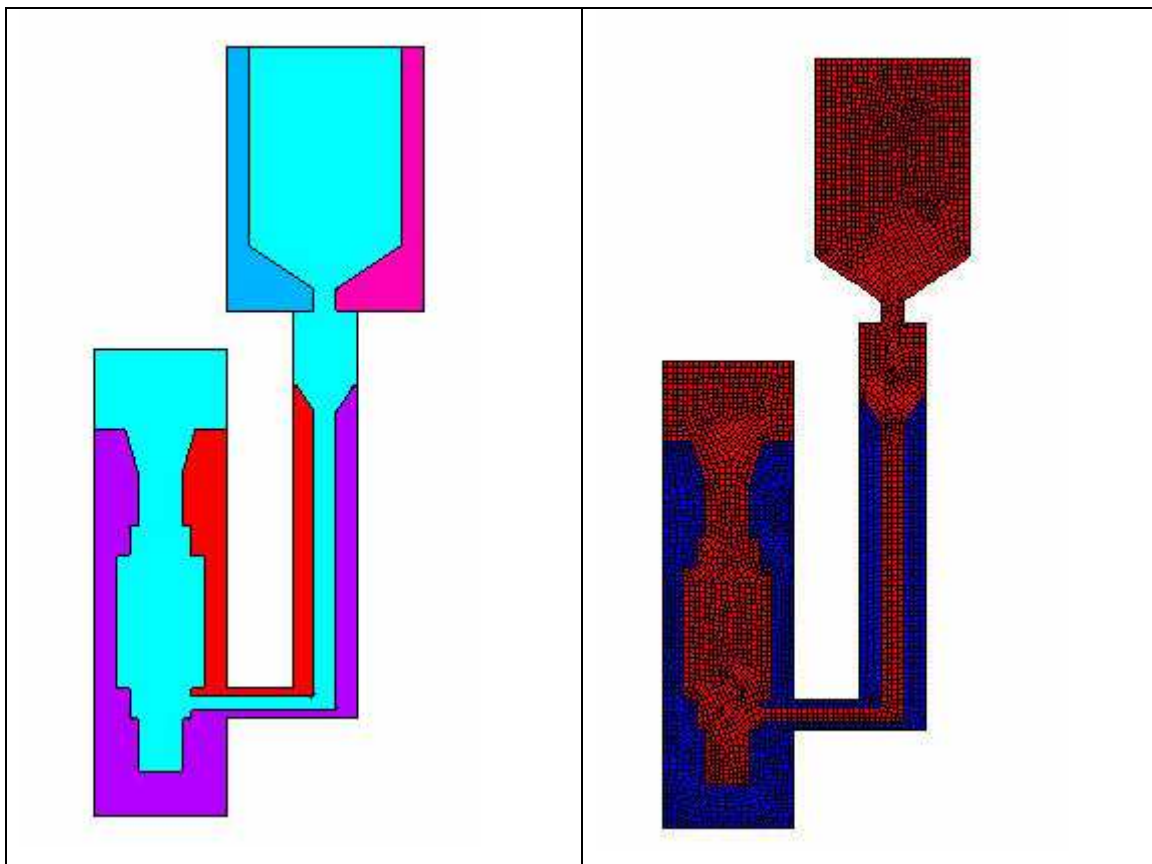


Figura 1 – Área do sistema

Figura 2 – Malha do sistema

O estudo foi feito em duas etapas. Sendo que na primeira etapa foi simulado o vazamento da liga até o preenchimento total do molde e na segunda etapa foi simulado o processo de solidificação do sistema metal-molde durante o tempo de uma hora e meia.

Para a realização do vazamento consideraram-se as seguintes condições de contorno: velocidades nas paredes internas do molde como sendo zero, gravidade a  $9,78\text{m/s}^2$ . Ainda como condição inicial a pressão ambiente foi de  $1\text{atm}$  e a temperatura

ambiente de 293K. Logo, informaram-se as propriedades termofísicas da liga, densidade ( $3297 \text{ Kg/m}^3$ ), viscosidade (0.001324) e calor-específico (963 J/Kg.K) constantes, condutividade térmica variante em função da temperatura (Tabela 1) e do molde metálico (Tabela 2).

<b>Tabela 1 – Propriedades da liga Al-22%Ni para o vazamento</b>	
Temperatura (kelvin)	Condutividade-Térmica (W/m.K)
400	196.52
450	195.985
500	194.8
548	193.1901
600	191.08
638	189.3916
710	186.1102
750	184.375
828	181.5215
868	180.4792
918	179.7289
968	179.7326
1000	179.7326

<b>Tabela 2 – Propriedades do molde metálico</b>			
Temperatura (Kelvin)	Densidade ( $\text{Kg/m}^3$ )	Calor Específico (J/Kg.K)	Condutividade térmica (W/m.K)
373	7870	519	14.6
473	7870	565	16.3
673	7870	607	19.3
873	7870	704	21.8
1073	7870	649	23.5
1273	7870	673	25.5

Na solidificação o tipo de malha e o elemento foi diferente em relação ao vazamento. Nesta etapa, como condição inicial foi considerada o campo de temperatura após finalizar o vazamento. Como condição de contorno foi aplicada a convecção na superfície externa do molde de  $70\text{W/m}^2.\text{K}$ . Foram informadas as propriedades termofísicas do metal fundido e das ligas em função da temperatura (Tabela 3).

<b>Tabela 1 – Propriedades da liga Al-22%Ni para a solidificação</b>			
Temperatura (Kelvin)	Densidade (Kg/m <sup>3</sup> )	Calor Específico (J/Kg.K)	Condutividade térmica (W/m.K)
400	3639.25	650.9	196.52
450	3629.4	605.975	195.985
500	3616.05	540.8	194.8
548	3599.2	458.48461	193.1901
600	3578.85	346.7	191.08
638	3555	249.66913	189.3916
710	3527.65	29.009	186.1102
750	3496.8	29	184.375
828	3462.45	29	181.5215
868	3424.6	29	180.4792
918	3383.25	29	179.7289
968	3338.4	29	179.7326
1000	3290.05	29	179.7326

## **Resultados e Discussão**

Através da simulação numérica em questão, após 3 segundos do vazamento, quando o molde ficou completamente cheio pelo metal líquido, foi obtido o campo de temperatura da Figura 3, este dado foi usado como condição inicial para o processo de solidificação.

Na solidificação após 1,5 h de resfriamento do metal fundido nos moldes, foi obtido como resultado o campo de temperatura (Figura 4), o campo de gradiente térmico (Figura 5) e o campo de fluxo térmico (Figura 6), que demonstram variações em todo o sistema durante o processo de solidificação. Também durante a solidificação foram obtidos as curvas de resfriamento (Figura 7).

Para realizar o processo de vazamento foi considerado este como um problema de mecânica de fluidos. Para isto foi usada a equação de Navier–Stokes, conservação de massa e equação de energia. Para o processo de solidificação foi usada a equação de transferência de calor.

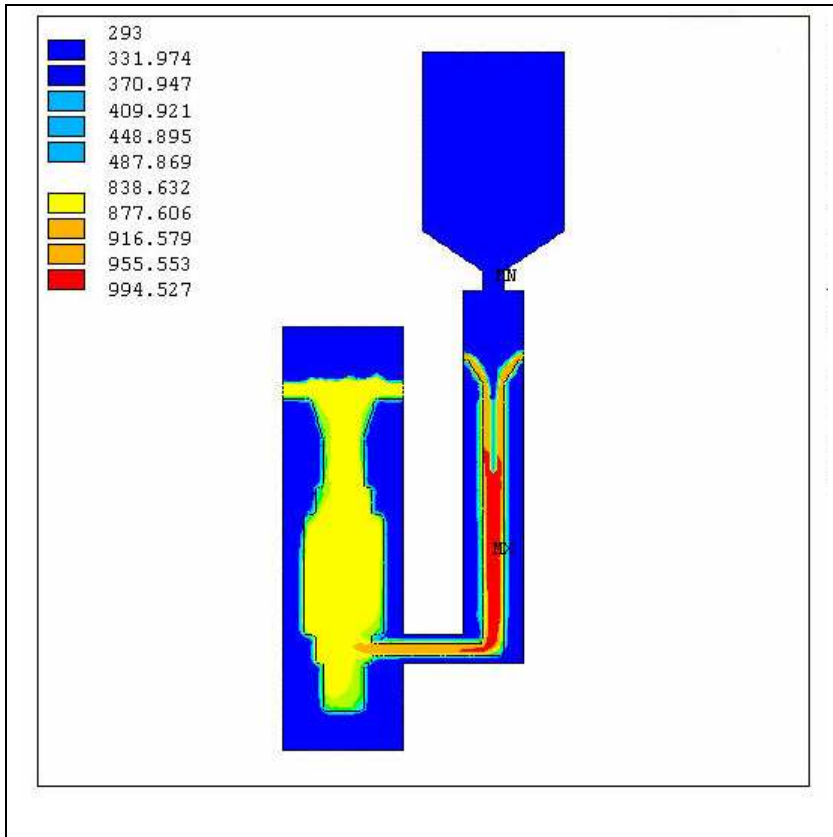


Figura 3 – Campo de temperatura após o vazamento

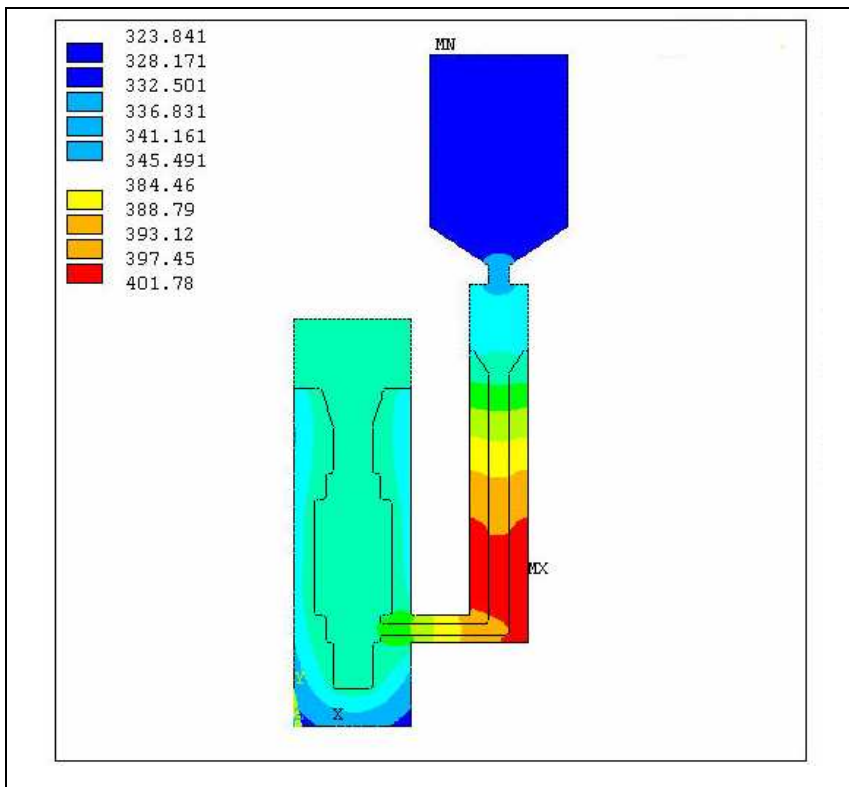


Figura 4 – Campo de temperatura após a solidificação

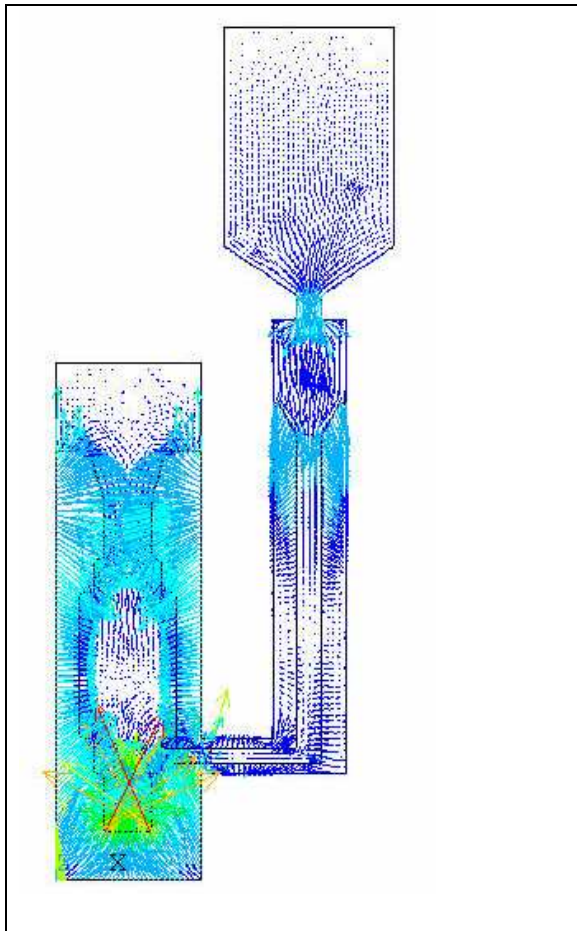


Figura 5 – Campo do gradiente térmico

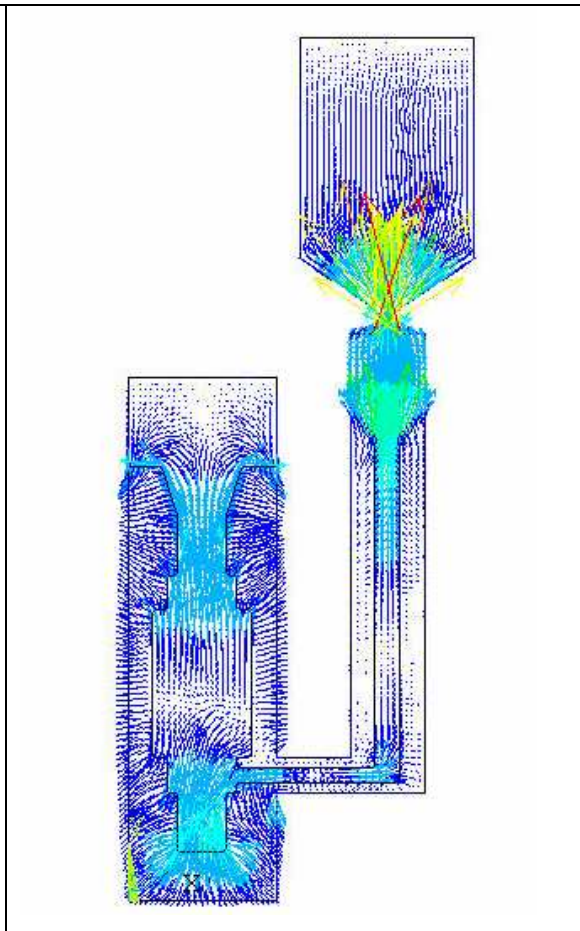


Figura 6 – Campo do fluxo de calor

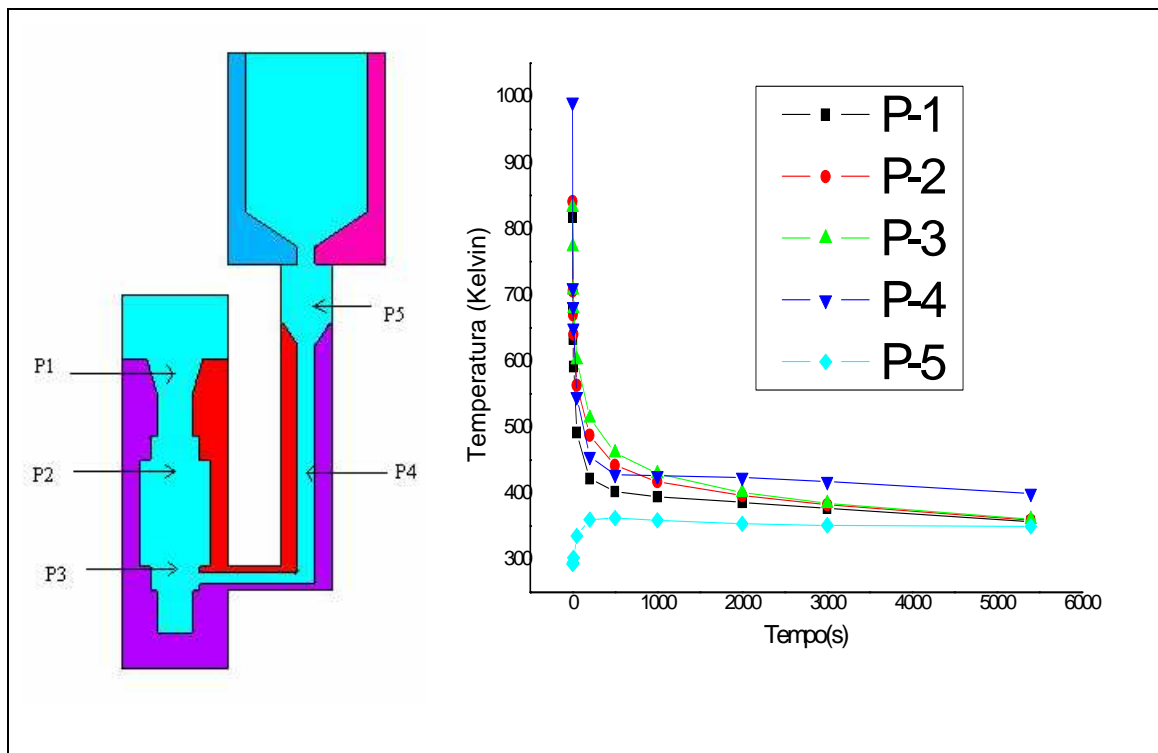


Figura 7 – Curvas de resfriamento durante a solidificação em diferentes pontos.

## **Conclusões**

Neste trabalho foi realizada a junção do processo de vazamento com o processo de solidificação. Nestes processos foram observados diferentes fenômenos que seriam difíceis de medir experimentalmente. Também, foi observado que o processo de vazamento influencia fortemente na qualidade do metal fundido. Desta forma, a simulação numérica permite otimizar e prever os processos com a finalidade de controlar a qualidade do metal fundido.

## **Agradecimentos**

Gostaria de agradecer ao meu orientador, Professor Dr. Moises Meza Pariona, aos colegas de pesquisa e à UEPG pela disponibilidade dos recursos físicos.

## **Referências Bibliográficas**

GUO Z., SAUNDERS N., MIODOWNIK A.P. and SCHILLÉ, J.-Ph., “Modeling of materials properties and behavior critical to casting simulation”. **Materials Science and Engineering**, v. A 413-414 (2005) 465-469.

Ansys handbook 10 version, Ansys Inc., Canosburg, 2005.

SHI, Z. GUO, Z. “Numerical heat transfer modeling for wire casting”, **Materials Science and Engineering**, v. A365 (2004) 311–317.

SU, X. **Computer aided optimization of an investment bi-metal casting process**, Ph.D. Thesis, University of Cincinnati, Department of Mechanical, Industrial and Nuclear Engineering, (2001).

GROZDANIC, V. “Numerical simulation of the solidification of a steel rail-wheel casting and the optimum dimension of the riser”, **Materiali in tehnologije**, v. 36, pp.39-41, (2002).

# HOMOGENEIZAÇÃO DA EQUAÇÃO DA ONDA COM DOMÍNIOS E OPERADORES VARIANDO SIMULTANEAMENTE

Joel Santos Souza; Jocemar de Quadros Chagas.

Universidade Federal de Santa Catarina - UFSC, jsouza@mtm.ufsc.br;

Universidade Estadual de Ponta Grossa - UEPG, jocemarchagas@uepg.br.

## Resumo:

Nesse trabalho estudamos a homogeneização da equação da onda para problemas de Dirichlet relaxados, definidos por meio de medidas positivas, para operadores elípticos de segunda ordem na forma de divergência, em que seus coeficientes e os seus respectivos domínios variam simultaneamente com  $\varepsilon$ .

**Palavras-chave:** Homogeneização, Condições de Dirichlet relaxadas, Domínios variáveis.

## 1 - INTRODUÇÃO

Neste trabalho, estudamos a homogeneização da equação da onda

$$u'' - Au = f,$$

onde  $A$  é um operador elíptico linear de segunda ordem com coeficientes mensuráveis limitados em  $\Omega$ . Consideramos uma seqüência de problemas de evolução com condições de Dirichlet relaxadas da forma

$$\begin{cases} u^{\varepsilon''} - \operatorname{div}(A^\varepsilon Du^\varepsilon) = f \text{ em } Q^\varepsilon = \Omega^\varepsilon \times (0, T), \quad T > 0 \\ u^\varepsilon = 0 \text{ em } \Sigma^\varepsilon = \Gamma^\varepsilon \times (0, T), \quad \Gamma^\varepsilon = \partial\Omega^\varepsilon \\ u^\varepsilon(x, 0) = u_0^\varepsilon(x) \text{ e } u^{\varepsilon'}(x, 0) = u_1^\varepsilon(x) \text{ em } \Omega^\varepsilon, \end{cases} \quad (1)$$

onde as matrizes  $A^\varepsilon$  e os domínios variáveis  $\Omega^\varepsilon$  dependem do parâmetro  $\varepsilon$  fixado. (Ou mais geralmente, consideramos uma seqüência de problemas de Dirichlet relaxados,



definidos por medidas positivas, para operadores elípticos lineares de segunda ordem sob a forma de divergência com matrizes de coeficientes também variáveis). Os conjuntos  $\Omega^\varepsilon$ , abertos, são todos contidos em um conjunto  $\Omega \subset \mathbb{R}^n$ , fixo, aberto e limitado, e as matrizes  $A^\varepsilon$ , definidas sobre  $\Omega$  com coeficientes mensuráveis, são coercivas e limitadas. O processo de homogeneização consiste em estudar o comportamento das soluções  $u^\varepsilon$  quando  $\varepsilon$  tende para zero. O presente trabalho segue na direção de [3].

## 2 - PRELIMINARES

### 2.1 - Definições e Notações

Daremos nessa seção algumas definições básicas:

Seja  $E \in \Omega$ . Definimos sua capacidade como:

$$\text{cap}(E) := \inf_{u \in H_0^1(\Omega)} \int_{\Omega} |Du|^2 dx; u \geq 1 \text{ q.s. em uma vizinhança de } E.$$

Dizemos que uma propriedade  $P(x)$  vale em *quase toda parte* (q.e.) em  $E$ , se  $P(x)$  vale em todo  $x \in E$ , exceto para um subconjunto  $N \subset E$ , com  $\text{cap}(N) = 0$ .

Dizemos que  $u : \Omega \rightarrow \mathbb{R}$  é *quase contínua* se  $\forall \varepsilon > 0, \exists E \subseteq \Omega$ , com  $\text{cap}(E) < \varepsilon$ , tal que  $u|_{\Omega \setminus E}$  é contínua.

Dizemos que  $U \subset \Omega$  é *quase aberto* se  $\forall \varepsilon > 0, \exists V \subseteq \Omega$ , com  $\text{cap}(V \Delta U) < \varepsilon$ , onde  $V$  é aberto e  $\Delta$  denota a diferença simétrica.

Observamos que toda  $u \in H^1(\Omega)$  possui uma representante quase contínua, que é unicamente definida a menos de um conjunto de capacidade nula, ou seja, se  $u \in H^1(\Omega)$ , então

$$u \cong v; v|_{\Omega \setminus E} \text{ é contínua, com } \text{cap}(E) = 0.$$

Uma *medida de Borel não negativa* sobre  $\Omega$  é uma função de conjunto aditiva contável definida sobre os subconjuntos de Borel de  $\Omega$  com valores em  $[0, +\infty]$ .

Uma *medida de Radon não negativa* sobre  $\Omega$  é uma medida de Borel não negativa

que é finita sobre todo conjunto compacto de  $\Omega$ .

$$\mu(E) = \inf\{\mu(B); B \text{ é boreliano}, E \subseteq B \subseteq \Omega\}.$$

$\mathcal{M}_0^+(\Omega)$  é o cone de todas as medidas de Borel não negativas  $\mu$  sobre  $\Omega$ , tais que:

$$\begin{cases} \text{(a) } \mu(B) = 0, \forall B \subseteq \Omega, \text{ com } \text{cap}(B) = 0, B \text{ boreliano.} \\ \text{(b) } \mu(B) = \inf\{\mu(U); U \text{ quase aberto}; B \subseteq U\}, \forall B \subseteq \Omega, \text{ boreliano.} \end{cases}$$

$\mathcal{M}_0(\Omega)$  denota o conjunto das medidas de Borel não negativas que somente satisfazem a condição (a).

Para todo conjunto quase aberto  $U \subseteq \Omega$ , definimos a medida de Borel  $\mu_U$  por

$$\mu_U(B) = \begin{cases} 0, & \text{se } \text{cap}(B \setminus U) = 0 \\ +\infty, & \text{caso contrário.} \end{cases}$$

O cone de todas as medidas de Radon sobre  $\Omega$  será denotado por  $\mathcal{M}(\Omega)$ .

O cone de todos os elementos não negativos de  $H^{-1}(\Omega)$  é denotado por  $H^{-1}(\Omega)^+$ . Como todo elemento de  $H^{-1}(\Omega)^+$  é uma medida de Radon não negativa que pertence também a  $\mathcal{M}_0^+(\Omega)$ , temos a inclusão

$$H^{-1}(\Omega)^+ \subseteq \mathcal{M}(\Omega) \cap \mathcal{M}_0^+(\Omega).$$

## 2.2 - H-Congregância

Sejam  $\alpha, \beta \in \mathbb{R}$ , com  $0 < \alpha \leq \beta < +\infty$ .

Definimos  $M_\alpha^\beta(\Omega)$  como o conjunto de todas as matrizes  $A \in L^\infty(\Omega, \mathbb{R}^{n \times n})$  tais que

$$A(x) \geq \alpha I, \quad (A(x))^{-1} \geq \beta^{-1} I, \quad \text{q.s. em } \Omega. \quad (2)$$

Em (2),  $I$  é a matriz identidade em  $\mathbb{R}^{n \times n}$ , e as desigualdades são no sentido das formas quádricas definidas por  $A(x)\xi\xi$  para  $\xi \in \mathbb{R}^n$ . (2) implica também que

$$|A(x)| \leq \beta \text{ q.s. em } \Omega$$

e que, necessariamente,  $\alpha \leq \beta$ .

**Definição 1** *Uma sequência  $(A^\epsilon)$  de matrizes em  $M_\alpha^\beta(\Omega)$  H-converge para uma matriz  $A^0$  em  $M_\alpha^\beta(\Omega)$ , se,  $\forall f \in H^{-1}(\Omega)$ , a sequência  $u^\epsilon$  de soluções dos problemas*

$$\begin{cases} u^\epsilon \in H_0^1(\Omega), \\ -\operatorname{div}(A^\epsilon Du^\epsilon) = f, \text{ em } \mathcal{D}'(\Omega) \end{cases}$$

satisfaz

$$\begin{aligned} u^\epsilon &\rightharpoonup u^0 \text{ fracamente em } H_0^1(\Omega) \\ A^\epsilon Du^\epsilon &\rightharpoonup A^0 Du^0 \text{ fracamente em } L^2(\Omega, \mathbb{R}^n), \end{aligned}$$

onde  $u^0$  é a solução do problema:

$$\begin{cases} u^0 \in H_0^1(\Omega), \\ -\operatorname{div}(A^0 Du^0) = f, \text{ em } \mathcal{D}'(\Omega). \end{cases}$$

**Observação 1** *Toda sequência de matrizes em  $M_\alpha^\beta(\Omega)$  possui uma subsequência que H-converge para uma matriz em  $M_\alpha^\beta(\Omega)$ .*

### 3 - HOMOGENEIZAÇÃO DO PROBLEMA ELÍPTICO

#### 3.1 - Resultado de Convergência

**Teorema 1** *Seja  $(A^\epsilon)$  uma sequência de matrizes em  $M_\alpha^\beta(\Omega)$  que H-converge para uma matriz  $A^0$  em  $M_\alpha^\beta(\Omega)$ , e  $u^\epsilon$  uma sequência em  $H^1(\Omega)$  tal que*

$$\begin{cases} u^\epsilon \rightharpoonup u^0 \text{ fracamente em } H^1(\Omega) \\ -\operatorname{div}(A^\epsilon Du^\epsilon) = f^\epsilon \text{ em } \mathcal{D}'(\Omega), \forall \epsilon \geq 0. \end{cases}$$

Assumindo que  $f^\epsilon = g^\epsilon + v^\epsilon$ , para todo  $\epsilon > 0$ , onde

$$\begin{cases} (g^\epsilon) \text{ é relativamente compacta em } W_{loc}^{-1,p}(\Omega), \text{ para algum } p > 1, \\ (v^\epsilon) \geq 0, \text{ em } \mathcal{D}'(\Omega). \end{cases}$$

então

$$f^\epsilon \rightharpoonup f^0 \text{ fracamente em } L^2(\Omega, \mathbb{R}^n).$$

Neste trabalho, este teorema será usado com  $(g^\epsilon)$  relativamente compacto (ou mesmo constante) em  $H^{-1}(\Omega)$ .

### 3.2 - Problema de Dirichlet Relaxado

Dados  $A \in M_\alpha^\beta(\Omega)$ ,  $\mu \in M_0^+(\Omega)$  e  $f \in H^{-1}(\Omega)$ , chamamos de problema de Dirichlet relaxado o problema de encontrar  $u$  tal que

$$\begin{cases} u \in H_0^1(\Omega) \cap L^2(\Omega, \mu), \\ \int_\Omega ADuDydx + \int_\Omega uyd\mu = \langle f, y \rangle, \forall y \in H_0^1(\Omega) \cap L^2(\Omega, \mu). \end{cases} \quad (3)$$

Por uma aplicação do lema de Lax-Milgram, o problema (3) tem uma única solução  $u$  (ver [3]), que satisfaz a estimativa

$$\alpha \int_\Omega |Du|^2 dx + \int_\Omega |u|^2 d\mu \leq \frac{1}{\alpha} \|f\|_{H^{-1}(\Omega)}^2. \quad (4)$$

### 3.3 - Reconstrução da medida $\mu$

Fixemos

$$A \in M_\alpha^\beta(\Omega), \quad \mu \in M_0^+(\Omega), \quad \lambda \in H^{-1}(\Omega)^+ \quad (5)$$

e uma solução  $w$  para o problema

$$\begin{cases} w \in H^1(\Omega) \cap L^2(\Omega, \mu), \\ \int_\Omega ADwDydx + \int_\Omega wyd\mu = \int_\Omega yd\lambda, \forall y \in H_0^1(\Omega) \cap L^2(\Omega, \mu), \end{cases} \quad (6)$$

que satisfaz

$$w \geq 0 \text{ q.e. em } \Omega. \quad (7)$$

**Observação 2** *Do lema de Lax-Milgram, existe uma única solução de (6) que pertence a  $H_0^1(\Omega)$ ; pelo princípio da comparação, esta solução satisfaz (7), de modo que o conjunto de tais funções é não vazio.*

**Proposição 1** *Assuma que (5), (6) e (7) são verdadeiras. Então  $\exists v \in H^{-1}(\Omega)^+$  tal que*

$$-\text{div}(ADw) + v = \lambda \text{ em } \mathcal{D}'(\Omega). \quad (8)$$

Por razões técnicas, a reconstrução da medida  $\mu$  de  $w$  requer a seguinte hipótese: para todo conjunto quase aberto  $U$  em  $\Omega$ , temos

$$\text{cap}(U \cap \{w = 0\}) > 0 \Rightarrow \lambda(U) > 0. \quad (9)$$

**Proposição 2** *Assuma as hipóteses (5), (6), (7) e (9). Então*

$$u \in H_0^1(\Omega) \cap L^2(\Omega, \mu) \Rightarrow u = 0 \text{ q.e. em } \{w = 0\}.$$

Além disso, para conjuntos de Borel  $B \subseteq \Omega$ , vale

$$\text{cap}(B \cap \{w = 0\}) > 0 \Rightarrow \mu(B) = +\infty.$$

A próxima proposição estabelece a reconstrução da medida  $\mu$ .

**Proposição 3** *Assuma as hipóteses (5), (6), (7) e (9), e seja  $v$  a medida de  $H^{-1}(\Omega)^+$  definida em (8). Então para todo conjunto de Borel  $B \subseteq \Omega$ , temos*

$$\mu(B) = \begin{cases} \int_B \frac{dv}{w}, & \text{se } \text{cap}(B \cap \{w = 0\}) = 0 \\ +\infty, & \text{se } \text{cap}(B \cap \{w = 0\}) > 0, \end{cases}$$

e

$$v(B \cap \{w = 0\}) = \int_B w d\mu.$$

Em particular, isso implica que  $v = w\mu$  sobre  $\{w > 0\}$ .

### 3.4 - Resultados de unicidade e densidade

Na próxima proposição, assumiremos que

$$w \in L^\infty(\Omega). \quad (10)$$

**Proposição 4** *Assuma que (5) - (7), (9) e (10) são verdadeiras. Então o conjunto  $\{w\varphi : \varphi \in C_c^\infty(\Omega)\}$  é denso em  $H_0^1(\Omega) \cap L^2(\Omega, \mu)$ .*

O seguinte resultado de unicidade é crucial para os teoremas 2 e 4. O resultado de unicidade é obtido a menos de um conjunto de capacidade nula.

**Proposição 5** *Assuma as hipóteses (5) - (7), (9) e (10). Seja  $u$  uma solução do problema*

$$\begin{cases} u \in H_0^1(\Omega) \cap L^\infty(\Omega), \\ \int_{\Omega} AD\varphi Duwdx - \int_{\Omega} ADwD\varphi udx + \int_{\Omega} u\varphi d\lambda = 0, \quad \forall \varphi \in C_0^\infty(\Omega). \end{cases}$$

Então,  $u = 0$  q.e. em  $\Omega$ .

### 3.5 - Resultado de convergência global

Para todo  $\epsilon \geq 0$ , consideramos uma matriz  $A^\epsilon$  em  $M_\alpha^\beta(\Omega)$  e uma medida  $\mu^\epsilon$  em  $M_0^+(\Omega)$ , que será fixada ao longo do resto deste trabalho. Assumimos que

$$(A^\epsilon)_{\epsilon \geq 0} \subset \mathcal{M}_\alpha^\beta(\Omega) \text{ são tais que } A^\epsilon \rightarrow A^0, \text{ H-convergência.} \quad (11)$$

Nesta seção usamos o argumento de dualidade para provar que, sob hipóteses adequadas sobre  $(\mu^\epsilon)$  (que sempre são satisfeitas para uma subsequência), as soluções  $u^\epsilon$  de problemas de Dirichlet relaxados (3) para  $A = A^\epsilon$  e  $\mu = \mu^\epsilon$  convergem para a solução  $u^0$  do problema de Dirichlet relaxado, com  $A = A^0$ , e  $\mu = \mu^0$ .

#### 3.5.1 - Definição de funções testes especiais

Para todo  $\epsilon \geq 0$ , definimos as funções  $w^\epsilon$  e  $\bar{w}^\epsilon$  como as únicas soluções para os problemas

$$\begin{cases} w^\epsilon \in H_0^1(\Omega) \cap L^2(\Omega, \mu^\epsilon), \\ \int_{\Omega} A^\epsilon Dw^\epsilon Dydx + \int_{\Omega} w^\epsilon y d\mu^\epsilon = \int_{\Omega} y dx, \quad \forall y \in H_0^1(\Omega) \cap L^2(\Omega, \mu^\epsilon), \end{cases} \quad (12)$$

e

$$\begin{cases} \bar{w}^\epsilon \in H_0^1(\Omega) \cap L^2(\Omega, \mu^\epsilon), \\ \int_{\Omega} \bar{A}^\epsilon D\bar{w}^\epsilon Dydx + \int_{\Omega} \bar{w}^\epsilon y d\mu^\epsilon = \int_{\Omega} y dx, \quad \forall y \in H_0^1(\Omega) \cap L^2(\Omega, \mu^\epsilon). \end{cases} \quad (13)$$

Além disso, pelo princípio do máximo, temos também

$$\sup \|w^\varepsilon\|_{L^\infty(\Omega)} < +\infty \quad \text{e} \quad \sup \|\bar{w}^\varepsilon\|_{L^\infty(\Omega)} < +\infty.$$

Pela proposição 1, existem duas medidas  $v^\varepsilon$  e  $\bar{v}^\varepsilon$  em  $H^{-1}(\Omega)^+$  tais que

$$-\operatorname{div}(A^\varepsilon Dw^\varepsilon) + v^\varepsilon = 1, \quad \text{e} \quad -\operatorname{div}(\bar{A}^\varepsilon D\bar{w}^\varepsilon) + \bar{v}^\varepsilon = 1, \quad \text{em } \mathcal{D}'(\Omega).$$

Finalmente, de (4) obtemos

$$\begin{aligned} \sup_{\varepsilon \geq 0} \int_{\Omega} |Dw^\varepsilon|^2 dx &< +\infty, & \sup_{\varepsilon \geq 0} \int_{\Omega} |D\bar{w}^\varepsilon|^2 dx &< +\infty, \\ \sup_{\varepsilon \geq 0} \int_{\Omega} |w^\varepsilon|^2 d\mu^\varepsilon &< +\infty, & \sup_{\varepsilon \geq 0} \int_{\Omega} |\bar{w}^\varepsilon|^2 d\mu^\varepsilon &< +\infty. \end{aligned}$$

### 3.5.2 - O principal resultado de convergência

Dadas, para toda  $\varepsilon \geq 0$ ,  $f^\varepsilon$  e  $\bar{f}^\varepsilon$  em  $H^{-1}(\Omega)$ , consideramos as soluções  $u^\varepsilon$  e  $\bar{u}^\varepsilon$  para o problema:

$$\begin{cases} u^\varepsilon \in H^1(\Omega) \cap L^2(\Omega, \mu^\varepsilon), \\ \int_{\Omega} A^\varepsilon Du^\varepsilon Dy dx + \int_{\Omega} u^\varepsilon y d\mu^\varepsilon = \langle f^\varepsilon, y \rangle, \quad \forall y \in H_0^1(\Omega) \cap L^2(\Omega, \mu^\varepsilon) \end{cases} \quad (14)$$

e para seu respectivo problema adjunto:

$$\begin{cases} \bar{u}^\varepsilon \in H_0^1(\Omega) \cap L^2(\Omega, \mu^\varepsilon), \\ \int_{\Omega} \bar{A}^\varepsilon D\bar{u}^\varepsilon Dy dx + \int_{\Omega} \bar{u}^\varepsilon y d\mu^\varepsilon = \langle \bar{f}^\varepsilon, y \rangle, \quad \forall y \in H_0^1(\Omega) \cap L^2(\Omega, \mu^\varepsilon). \end{cases} \quad (15)$$

Enunciamos agora o principal resultado de convergência:

**Teorema 2** *Admita (11), e sejam  $w^\varepsilon$  e  $\bar{w}^\varepsilon$  as soluções para (12) e (13). As seguintes condições são equivalentes:*

(a)  $w^\varepsilon \rightharpoonup w^0$  fracamente em  $H_0^1(\Omega)$ ;

(b)  $\bar{w}^\varepsilon \rightharpoonup \bar{w}^0$  fracamente em  $H_0^1(\Omega)$ ;

(c) para toda  $(f^\varepsilon)$  e  $(u^\varepsilon)$  satisfazendo (14), se  $f^\varepsilon \rightarrow f^0$  fortemente em  $H^{-1}(\Omega)$ , então  $u^\varepsilon \rightharpoonup u^0$  fracamente em  $H_0^1(\Omega)$ ;

(d) para toda  $(\bar{f}^\varepsilon)$  e  $(\bar{u}^\varepsilon)$  satisfazendo (15), se  $\bar{f}^\varepsilon \rightarrow \bar{f}^0$  fortemente em  $H^{-1}(\Omega)$ , então  $\bar{u}^\varepsilon \rightharpoonup \bar{u}^0$  fracamente em  $H_0^1(\Omega)$ .

### 3.5.3 - Um resultado de compacidade

As condições (a)-(d) do Teorema (2) não são fictícias. É o que mostra o resultado de compacidade que apresentamos a seguir:

**Teorema 3** *Admita (11). Para toda sequência  $(\mu^\varepsilon)_{\varepsilon>0}$  em  $\mathcal{M}_0^+(\Omega)$  existe uma subsequência, ainda denotada por  $(\mu^\varepsilon)$ , e uma medida  $\mu^0$  em  $\mathcal{M}_0^+(\Omega)$  tal que as condições equivalentes (a)-(d) do Teorema 2 são satisfeitas.*

### 3.5.4 - Funções testes

Introduzimos agora uma família mais geral de funções testes  $(w^\varepsilon)$ .

Para todo  $\varepsilon \geq 0$ , seja  $\lambda^\varepsilon \in H^{-1}(\Omega)$ , e seja  $w^\varepsilon$  uma solução do problema

$$\begin{cases} w^\varepsilon \in H^{-1}(\Omega) \cap L^2(\Omega, \mu^\varepsilon), \\ \int_{\Omega} A^\varepsilon Dw^\varepsilon Dydx + \int_{\Omega} w^\varepsilon y d\mu^\varepsilon = \int_{\Omega} y d\lambda^\varepsilon, \forall y \in H_0^1(\Omega) \cap L^2(\Omega, \mu^\varepsilon). \end{cases} \quad (16)$$

Assumimos que

$$\lambda^\varepsilon \in H^{-1}(\Omega)^+, \forall \varepsilon \geq 0, \quad (17)$$

$$\lambda^\varepsilon \rightarrow \lambda^0, \text{ fortemente em } H^{-1}(\Omega), \quad (18)$$

$$w^\varepsilon \geq 0 \text{ q.e. em } \Omega \forall \varepsilon \geq 0, \quad (19)$$

$$w^\varepsilon \rightharpoonup w^0 \text{ fracamente em } H^1(\Omega). \quad (20)$$

Também assumimos que, para todo conjunto quase aberto  $U$  em  $\Omega$ , temos que

$$\text{cap}(U \cap \{w^0 = 0\}) > 0 \Rightarrow \lambda^0(U) > 0, \text{ e} \quad (21)$$

$$w^0 \in L^\infty(\Omega). \quad (22)$$

**Teorema 4** *Admita que vale (11), e que  $(w^\varepsilon)_{\varepsilon \geq 0}$  e  $(\lambda^\varepsilon)_{\varepsilon \geq 0}$  satisfazem (16) - (22). Então as condições equivalentes (a)-(d) do Teorema 2 são cumpridas.*



#### 4 - PROBLEMAS DE DIRICHLET EM DOMÍNIOS VARIÁVEIS

Vamos considerar agora o caso particular do problema de Dirichlet clássico em domínios variáveis.

Seja  $(\Omega^\varepsilon)_{\varepsilon>0}$  uma sequência de conjuntos abertos, com  $\Omega^\varepsilon \subseteq \Omega$ , e seja  $\mu^0$  uma medida em  $\mathcal{M}_0^+(\Omega)$ . Para cada  $\varepsilon > 0$ , sejam  $w^\varepsilon$  e  $\bar{w}^\varepsilon$  as únicas soluções dos problemas

$$\begin{cases} w^\varepsilon \in H_0^1(\Omega^\varepsilon), \\ -\operatorname{div}(A^\varepsilon Dw^\varepsilon) = f \text{ em } \mathcal{D}'(\Omega^\varepsilon), \end{cases} \quad (23)$$

e

$$\begin{cases} \bar{w}^\varepsilon \in H_0^1(\Omega^\varepsilon), \\ -\operatorname{div}(\bar{A}^\varepsilon D\bar{w}^\varepsilon) = f \text{ em } \mathcal{D}'(\Omega^\varepsilon), \end{cases} \quad (24)$$

e sejam  $w^0$  e  $\bar{w}^0$  as soluções de (12) e (13) com  $\varepsilon = 0$ .

Dadas  $f^\varepsilon$  e  $\bar{f}^\varepsilon$  em  $H^{-1}(\Omega)$ , para  $\varepsilon > 0$ , consideremos  $u^\varepsilon$  e  $\bar{u}^\varepsilon$  soluções dos problemas:

$$\begin{cases} u^\varepsilon \in H_0^1(\Omega^\varepsilon), \\ u^{\varepsilon''} - \operatorname{div}(A^\varepsilon Du^\varepsilon) = f^\varepsilon \text{ em } \mathcal{D}'(0, T; \mathcal{D}'(\Omega^\varepsilon)), \end{cases} \quad (25)$$

e

$$\begin{cases} \bar{u}^\varepsilon \in H_0^1(\Omega^\varepsilon), \\ \bar{u}^{\varepsilon''} - \operatorname{div}(\bar{A}^\varepsilon D\bar{u}^\varepsilon) = \bar{f}^\varepsilon \text{ em } \mathcal{D}'(0, T; \mathcal{D}'(\Omega^\varepsilon)). \end{cases} \quad (26)$$

Dadas  $f^0$  e  $\bar{f}^0$  em  $H^{-1}(\Omega)$ , sejam  $u^0$  e  $\bar{u}^0$  as soluções de (14) e de (15), com  $\varepsilon = 0$ .

**Corolário 1** *Assuma (11) e sejam  $w^\varepsilon$  e  $\bar{w}^\varepsilon$  soluções de (23) e de (24) para  $\varepsilon > 0$ , e de (12) e de (13) para  $\varepsilon = 0$ . As seguintes condições são equivalentes:*

(a)  $w^\varepsilon \rightharpoonup w^0$  fracamente em  $H_0^1(\Omega)$ ;

(b)  $\bar{w}^\varepsilon \rightharpoonup \bar{w}^0$  fracamente em  $H_0^1(\Omega)$ ;

(c) para toda  $(f^\varepsilon)$  e  $(u^\varepsilon)$  satisfazendo (25) para  $\varepsilon > 0$  e (14) para  $\varepsilon = 0$ , se  $f^\varepsilon \rightarrow f^0$  fortemente em  $H^{-1}(\Omega)$ , então  $u^\varepsilon \rightharpoonup^* u^0$  fraco-estrela em  $L^\infty(0, T; H_0^1(\Omega))$ ;

(d) para toda  $(\bar{f}^\varepsilon)$  e  $(\bar{u}^\varepsilon)$  satisfazendo (26) para  $\varepsilon > 0$  e (15) para  $\varepsilon = 0$ , se  $\bar{f}^\varepsilon \rightarrow \bar{f}^0$  fortemente em  $H^{-1}(\Omega)$ , então  $\bar{u}^\varepsilon \rightharpoonup^* \bar{u}^0$  fraco-estrela em  $L^\infty(0, T; H_0^1(\Omega))$ .



solução do problema de Dirichlet

$$(D_\varepsilon) \begin{cases} u^\varepsilon \in L^\infty(0, T; H_0^1(\Omega^\varepsilon)), \\ u^{\varepsilon''} - \operatorname{div}(A^{\varepsilon'} Du^\varepsilon) = f^\varepsilon, \text{ em } \mathcal{D}'(0, T; \mathcal{D}'(\Omega^\varepsilon)), \\ u^\varepsilon = 0 \text{ em } \Sigma^\varepsilon = \Gamma^\varepsilon \times (0, T), \Gamma^\varepsilon = \partial\Omega^\varepsilon, \\ u^\varepsilon(x, 0) = u_0^\varepsilon(x), u^{\varepsilon'}(x, 0) = u_1^\varepsilon(x). \end{cases} \quad (28)$$

Queremos demonstrar o seguinte resultado:

**Teorema 5** *Suponha que (11) seja satisfeita para toda  $f^\varepsilon$  e  $u^\varepsilon$  satisfazendo (27), para  $\varepsilon = 0$ , e (28), para  $\varepsilon > 0$ ; e suponha que*

$$\begin{aligned} u_0^\varepsilon &\rightharpoonup u^0, \text{ fracamente em } H_0^1(\Omega), \\ u_1^\varepsilon &\rightharpoonup u^1, \text{ fracamente em } L^2(\Omega), \text{ e} \\ f^\varepsilon &\rightarrow f^0, \text{ fortemente em } L^\infty(0, T; H^{-1}(\Omega)), \end{aligned}$$

Então

$$u^\varepsilon \xrightarrow{*} u^0, \text{ fraco-estrela em } L^\infty(0, T; H_0^1(\Omega)).$$

Assim, resumidamente, poderíamos escrever:

$$D_\varepsilon \rightarrow DR_0, \text{ quando } \varepsilon \rightarrow 0.$$

## REFERÊNCIAS BIBLIOGRÁFICAS:

- [1] DAL MASO, G.; GARRONI, A.: **New results on the asymptotic behaviour of Dirichlet problems in perforated domains.** Math. Models Methods Appl. Sci. 4, 1994. 373-407.
- [2] DAL MASO, G.; MOSCO, U.: **Wiener's criterion and  $\Gamma$ -convergence.** Appl. Math. Optim. 15, 1987. 15-63.

- [3] DAL MASO, G.; MURAT, F.: **Comportement asymptotique et correcteurs pour des problèmes de Dirichlet linéaires avec des opérateurs et des domaines qui varient simultanément.** Ann. I. H. Poincaré - AN 21, 2004. 445-486.
- [4] CIORANESCU, D.; MURAT, F.: **Un terme étrange venu d'ailleurs. Nonlinear Partial Differential Equations and their Applications.** collège de France Seminar, vol. II e III, Research Notes in Mathematics. vol. 60 e 70, Pitman, 1982. 93-138 e 154-178.
- [5] CIORANESCU, D.; DONATO, P.; MURAT, F.; ZUAZUA, E. **Homogenization and correctors for the wave equation in domains with small holes.** Ann. Scuola Norm. Sup. Pisa. 18, 1991. 251-293.

# ESTUDO DA CURVATURA DOS AEROFÓLIOS GERADOS PELA TRANSFORMAÇÃO GENERALIZADA DE JOUKOWSKI

Rubén Panta Pazos

Universidade de Santa Cruz do Sul – UNISC, [rpazos@unisc.br](mailto:rpazos@unisc.br)

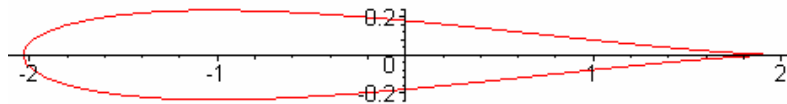
## Resumo

Neste trabalho se estudam os aerofólios gerados pela transformação generalizada de Joukowski fazendo ênfase na análise da curvatura dos aerofólios. Assim, a função curvatura representa uma espécie de assinatura do aerofólio, e podem aplicar-se transformadas discretas wavelets visando compactar a função de curvatura, para uma posterior classificação com uma base de dados. A reconstrução do aerofólio a partir da função de curvatura é obtida mediante um sistema de computação algébrica.

**Palavras-chave:** *aerofólio, curvatura, transformação generalizada de Joukowski.*

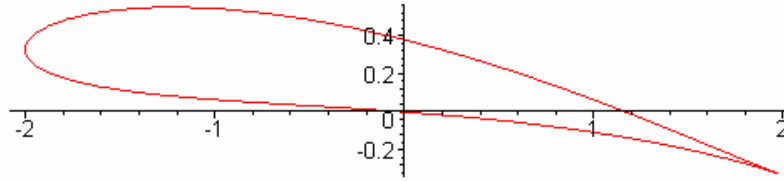
## 1. Introdução

Os métodos de variáveis complexas têm sido usados em diversas áreas da engenharia, tais como aerodinâmica, elasticidade, electromagnetismo. Especialmente a teoria da transformação conforme foi empregada por Joukowski para gerar os aerofólios que levam seu nome. Seja  $S^1$  o círculo unitário no plano complexo. Então, a imagem de  $S^1$  mediante a função  $J(z) = z + \frac{1}{z}$  é o segmento de comprimento 4 sobre o eixo real. Se a circunferência unitária for deslocada da origem sobre o eixo real, então a imagem de  $J(z)$  é o chamado aerofólio de Joukowski, ver BATCHELOR (1967), que aparece na Figura 1.



**Figura 1.** Aerofólio clássico de Joukowski obtido após aplicar  $J(z)$  sobre  $S^1 + \delta$ , com  $\delta$  real..

Quando o centro da circunferência sai do eixo real, o aerofólio curva-se para cima ou para baixo, conforme pode ver-se na Figura 2.



**Figura 2.** Aerofólio de Joukowski obtido após aplicar  $J(z)$  sobre  $S^1 + \zeta$ , com  $\zeta$  complexo.

Este trabalho está organizado da seguinte forma. Na seção seguinte é dada uma revisão sobre a Transformação Generalizada de Joukowski, visando obter clareza sobre o tema. Na seção 3 se associa o perfil de alguns aerofólios com suas curvaturas, dando ênfase na decomposição mediante funções do tipo de frações contínuas. A reconstrução da curva a partir da função de curvatura é esboçada na seção 4. Finalmente são dadas algumas conclusões e possíveis extensões.

## 2. Transformação Generalizada de Joukowski

A transformação generalizada de Joukowski se define da forma seguinte:

$$J(a, z) = z + \sum_{k=1}^n \frac{a_k}{z^k}, \quad (1)$$

onde  $a$  é um vetor  $n$ -dimensional de elementos complexos, isto é  $a \in C^n$ . Evitando os pontos singulares, (1) representa uma transformação conforme, KYTHE (1998).

Um aerofólio gerado pela transformação generalizada de Joukowski será a curva que resulta de aplicar (1) na circunferência  $S^\epsilon$ , com centro diferente da origem do plano complexo e raio  $\epsilon > 0$ .

Com o propósito de realizar uma análise em relação aos aerofólios gerados pela Transformação Generalizada de Joukowski até a segunda ordem de inversão, se procede da seguinte forma, ALMEIDA e PAZOS (2007) :

- Procuram-se pontos singulares, para evitá-los;
- Decompõe-se  $J([a_1, a_2], z)$  numa seqüência de composições de transformações conformes elementares, mediante funções do tipo frações contínuas;
- Análise da influência dos coeficientes na forma dos fluxos;
- Condições de tipo aerodinâmico. Ensaio experimentais em túnel de vento.

Destes estágios, vamos revisar apenas o correspondente à decomposição de  $J([a_1, a_2], z)$  mediante funções do tipo frações contínuas.

A fórmula da transformação generalizada de Joukowski sob consideração com inversão até 2ª ordem é uma Série de Laurent truncada,  $J([a_1, a_2], z) = z + \frac{a_1}{z} + \frac{a_2}{z^2}$  e pode ser representada por um formato de frações contínuas de  $z$ , da seguinte maneira:

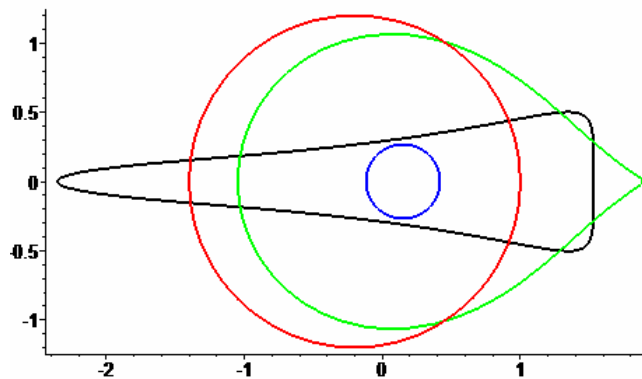
$$J([a_1, a_2], z) = z + \frac{a_1^2}{a_1 z - a_2 + \frac{a_2^2}{a_1 z + a_2}} \quad (2)$$

A equação (2) será analisada passo a passo.

Estágio 1:  $\frac{a_2^2}{a_1 z + a_2}$ : Transformação de Moebius

Estágio 2:  $a_1 z - a_2 + \frac{a_2^2}{a_1 z + a_2}$ : Soma de uma Transformação de Moebius e outra Linear.

Estágio 3:  $J([a_1, a_2], z)$ , resulta de uma inversão complexa da anterior, além de somar a identidade complexa.



**Figura 3.** Esquema demonstrando transformações ocorridas com centro deslocado da origem.

Na figura 3, podem ser visualizadas as transformações que ocorrem na circunferência unitária inicial com centro deslocado da origem (representada pela cor vermelha) até se transformar em um aerofólio mediante transformação generalizada de Joukowski. A primeira transformação é uma Transformação de Moebius Elíptica, apresenta um círculo azul. A segunda transformação é uma combinação de uma Moebius e outra linear, gerando uma figura semelhante a um aerofólio (tipo gota na posição horizontal), porém mais dilatado, visto na cor verde. E finalmente a transformação generalizada de Joukowski (na cor preta), que representa um aerofólio com coeficientes 1 e -0,47, para as inversões de primeira e segunda ordem, respectivamente.

### 3. Perfil de aerofólios e a função de curvatura

Nesta seção se dá uma breve introdução à noção de curvatura de uma curva parametrizada, e se inclui o Teorema Fundamental das Curvas no plano. Se reproduzem gráficos das funções de curvatura do aerofólio clássico de Joukowski e do aerofólio gerado pela transformação generalizada de Joukowski para  $a = [1, -0.47]$ .

Consideremos uma curva plana em coordenadas cartesianas parametrizada pelo seu comprimento de arco  $s$

$$\sigma(s) = [x(s), y(s)] \quad (3)$$

O vetor tangente unitário à curva no ponto genérico  $P$  é dado por

$$\frac{d\sigma}{ds} = [x'(s), y'(s)] \quad (4)$$

Intuitivamente é possível imaginar que a curvatura da curva mede a variação de seu vetor tangente em seu deslocamento ao longo dela, o que conduz a as seguintes definições:

Denomina-se *vetor de curvatura* em  $P$  ao vetor

$$\frac{d^2\sigma}{ds^2} = [x''(s), y''(s)] \quad (5)$$

Define-se a *curvatura*  $\kappa$  no ponto  $P$  como o módulo do vetor de curvatura em  $P$ . Com certeza, o vetor curvatura é ortogonal ao vetor tangente à curva; isto quer dizer que

$$\frac{d^2\sigma}{ds^2} = \kappa \bar{n} \quad (6)$$

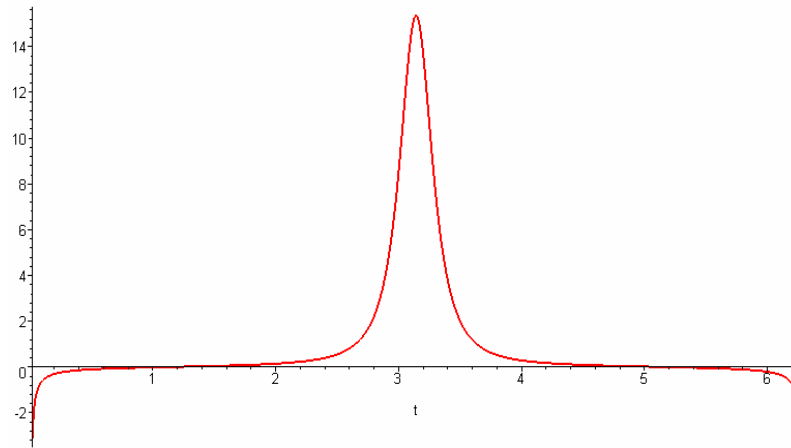
Sendo  $\bar{n}$  o vetor normal unitário em  $P$ . A equação (6) chama-se a primeira fórmula de Frenet.

**Exemplo 3.1.** Considere-se o aerofólio de Joukowski, imagem da circunferência cujo centro está deslocado em 0,1 unidades à esquerda da origem e com raio 1,1, com as seguintes equações paramétricas:

$$\sigma(t) = \begin{bmatrix} -0,1 + 1,1 \cos(t) + \frac{1,1 \cos(t) - 0,1}{(1,1 \cos(t) - 0,1)^2 + 1,21 \sin(t)} \\ 1,1 \sin(t) - \frac{1,1 \sin(t)}{(1,1 \cos(t) - 0,1)^2 + 1,21 \sin(t)} \end{bmatrix} \quad (7)$$

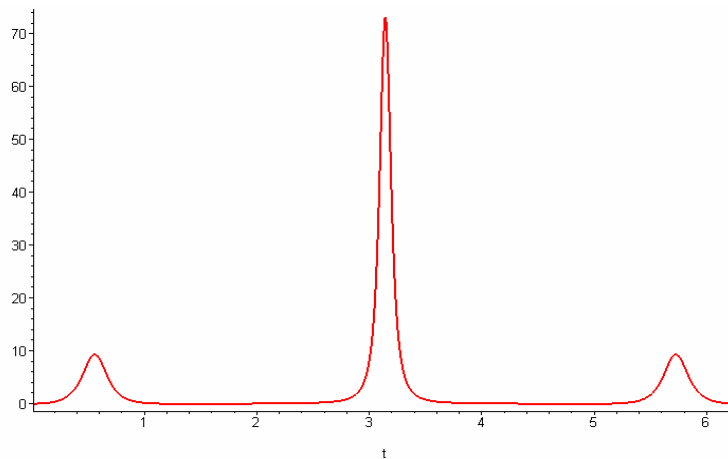
A partir de (7) se constrói a função curvatura, cujo gráfico é exibido na Figura 4.





**Figura 4.** Função de curvatura do aerofólio clássico de Joukowski.

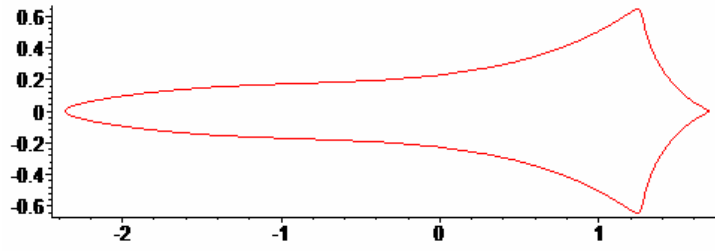
**Exemplo 3.2.** Considere-se o aerofólio gerado por  $J([1, -0.47], z)$  quando for aplicada na circunferência cujo centro está deslocado em 0,1 unidades à esquerda da origem e com raio 1,1.



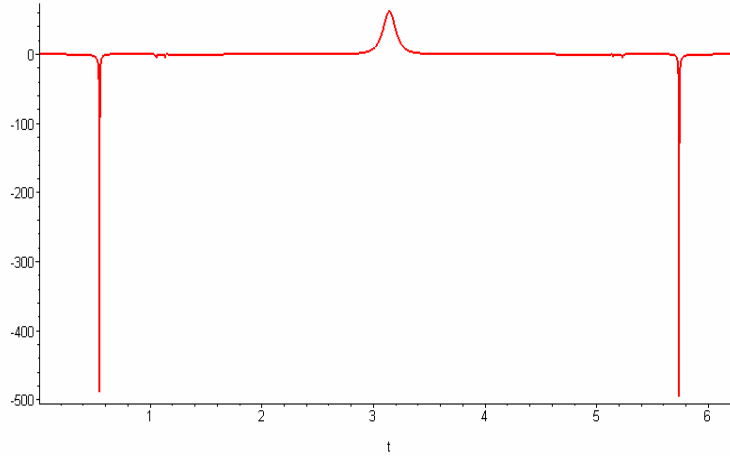
**Figura 5.** Função de curvatura do aerofólio gerado por  $J([1, -0.47], z)$ .

Assim, dada uma curva parametrizada, se define a função de curvatura. No caso dos aerofólios, por se tratar de uma curva fechada, isto representa um passo importante no reconhecimento da forma aerodinâmica dos aerofólios (bidimensional no caso) mediante uma função que depende apenas de um parâmetro. No que segue, são dados alguns exemplos não triviais.

**Exemplo 3.3.** Trata-se do aerofólio gerado por pela transformação generalizada de Joukowski de até quinta ordem de inversão,  $J([1, -0.5, 0.0, 0.1, 0.1], z)$  quando se aplica na circunferência cujo centro está deslocado 0,2 unidades à esquerda da origem e com raio 1,2. O aerofólio aparece na Figura 5, e sua função de curvatura na Figura 6.

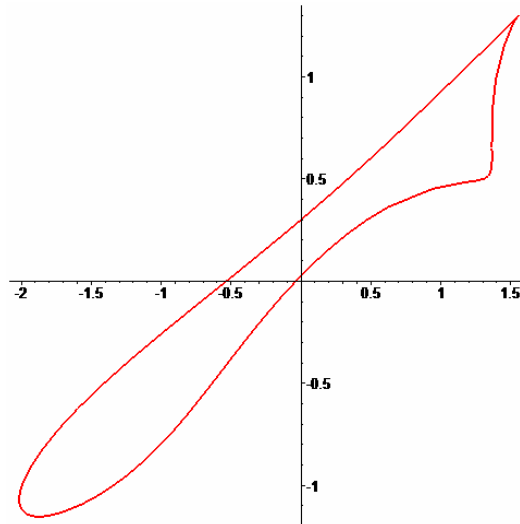


**Figura 6.** Aerofólio gerado por  $J([1, -0.5, 0.0, 0.1, 0.1], z)$  sobre  $S_{1,2}(0.2)$ .

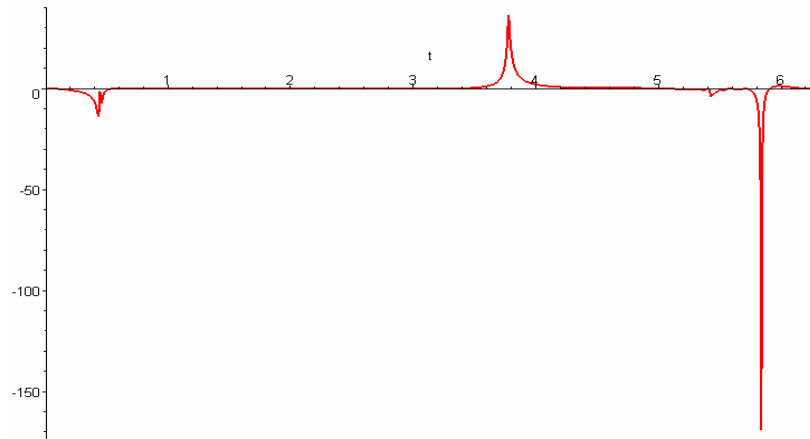


**Figura 7.** Função de curvatura do aerofólio gerado por  $J([1, -0.5, 0.0, 0.1, 0.1], z)$ .

**Exemplo 3.4.** O vetor  $a$  pode ter componentes complexos, como foi indicado na definição da transformação generalizada de Joukowski. Tome-se, como exemplo, o aerofólio gerado por  $J([0.5 + 1.05i, -0.5e^i, 0.125e^{-i}, 0.0625e^i, -0.05 + 0.03125e^{-i}], z)$  sobre  $S_{1,2}(0.2)$ .



**Figura 8.** Aerofólio gerado por  $J([0.5 + 1.05i, -0.5e^i, 0.125e^{-i}, 0.0625e^i, -0.05 + 0.03125e^{-i}], z)$  sobre  $S_{1,2}(0.2)$ .



**Figura 9.** Função de curvatura do aerofólio gerado por  $J\left([0.5 + 1.05i, -0.5e^i, 0.125e^{-i}, 0.0625e^i, -0.05 + 0.03125e^{-i}], z\right)$ .

#### 4. Reconstrução de uma curva a partir da função de curvatura

A continuação o teorema da maior importância das curvas no plano.

**Teorema 3.1** (Teorema Fundamental das Curvas) *Sejam  $\kappa(s)$  e  $\tau(s)$ , ( $s > 0$ ) duas funções analíticas dadas. Então existe uma única curva (única até sua posição em  $\mathbb{R}^3$ ) tal que  $s$  é o comprimento de arco,  $\kappa(s)$  a curvatura e  $\tau(s)$  a torção <sup>1</sup>.*

Conforme o Teorema 3.1, é possível reconstruir uma curva plana a partir de sua função de curvatura. A curva reconstruída difere da curva original por um movimento rígido.

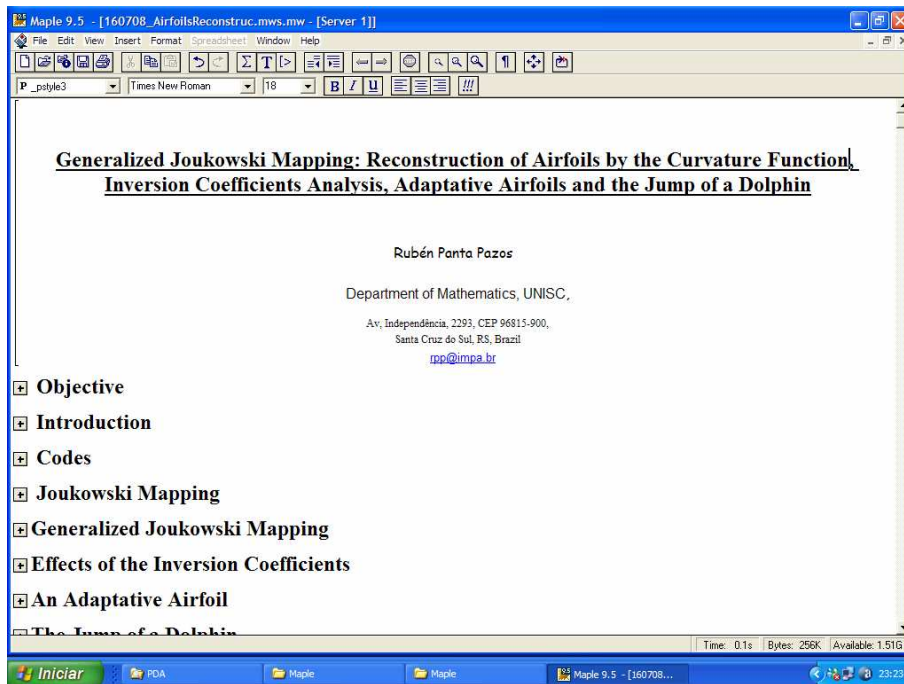
Tomado estas idéias foram realizados ensaios num sistema de computação algébrica, visando a reconstrução das curvas a partir da função de curvatura, OPREA (2007). Para funções elementares do tipo escalonado ou linear por pedaços a reconstrução resultou muito simples. Mas para funções de curvatura bem mais sofisticadas, houve de aperfeiçoar com procedimentos numéricos para resolver equações diferenciais.

A aplicação disto em áreas tão diversas como análise de formas mediante a função de curvatura, ou ainda para o fashion design e detecção de bordas, é de salientar.

---

<sup>1</sup> Para uma curva parametrizada no espaço  $\mathbb{R}^3$ ,  $\mathbf{r} = \mathbf{r}(s)$ , a curvatura é dada por  $\kappa = \frac{\|\mathbf{r}' \times \mathbf{r}''\|}{\|\mathbf{r}'\|^3}$ , e a torção

por  $\tau = \frac{[\mathbf{r}' \times \mathbf{r}''] \cdot \mathbf{r}'''}{\|\mathbf{r}' \times \mathbf{r}''\|^2}$ . Para curvas planas a torção é nula, OPREA (2007).



**Figura 10.** Reconstrução e análise de aerofólios da transformação generalizada de Joukowski mediante uma folha de trabalho de Maple.

## 5. Conclusões

Neste trabalho se analisou o enfoque de associar o perfil de aerofólios gerados pela transformação generalizada de Joukowski com a função de curvatura. O Teorema Fundamental das Curvas (Teorema 3.1) permite garantir a unicidade de esta associação, exceto um movimento rígido. Para isso, o parâmetro fundamental foi o comprimento de arco.

Do ponto de vista computacional, o emprego de um sistema de computação algébrica permite realizar simulações do tipo de aerofólio e da função de curvatura correspondente. Da mesma forma foi possível reconstruir a curva a partir de sua função de curvatura, isto é resolver um problema inverso.

A identificação dos aerofólios será possível analisando apenas a função de curvatura, que surge desta forma como a “assinatura” do aerofólio. Pode estabelecer-se ainda a “evolução” do aerofólio a partir da decomposição mediante transformações conformes elementares do tipo de frações contínuas.

As próximas tarefas compreendem a análise dos fluxos em torno a estes aerofólios e, possivelmente, o estudo das curvaturas das linhas de fluxos, e ainda o estudo do ponto de estagnação mediante procedimentos da geometria diferencial clássica.

## **Agradecimentos**

O autor deve reconhecer o auxílio fornecido pela Universidade de Santa Cruz do Sul para a realização deste trabalho; além disso, deve agradecer ao IMPA (Instituto Nacional de Matemática Pura e Aplicada, Rio de Janeiro, RJ) pelas facilidades dadas no Programa de Pós Doutorado de Verão de 2008.

## **Referências**

ALMEIDA, P. F. D. de, PAZOS, R. E. P. **Análise de aerofólios gerados pela Transformação Generalizada de Joukowski**, XXX Congresso Nacional de Matemática Aplicada e Computacional, 2007, Florianópolis. Anais do XXX Congresso Nacional de Matemática Aplicada e Computacional. Florianópolis: UFSC, (2007).

BATCHELOR, G. K., **An introduction to fluid dynamics**, Cambridge University Press, UK, (1967).

KYTHER, Prem K., **Computational Conformal Mapping**, Birkhäuser, Boston, MA, USA, (1998).

OPREA, J., **Differential Geometry and its Applications**, The Mathematical Association of America, USA, (2007).

PAZOS, R. E. P.; ALMEIDA, P. F. D. de, **Generalized Joukowski Mapping**, Maple Conference 2005, Waterloo, Canada. Maple Conference 2005 Proceedings, (2005), v. 1

# APLICAÇÃO DE TRANSFORMADAS DISCRETAS WAVELETS AO SOM DE UM CRUSTÁCEO

Patricia Fernanda Doern de Almeida Fischborn, Rubén Edgardo Panta Pazos.

Universidade de Santa Cruz do Sul – UNISC, [pattynanda@ibest.com.br](mailto:pattynanda@ibest.com.br);

Universidade de Santa Cruz do Sul – UNISC, [rpp@impa.br](mailto:rpp@impa.br), [www.rpanta.com](http://www.rpanta.com).

## Resumo:

É feita a análise do som produzido pelo crustáceo lagostim-vermelho (*Procambarus clarkii*), realizada com o uso do software *Sound Ruler* e com o método de transformadas *wavelets*. A seguir é dada uma introdução da Teoria dos *Wavelets* e suas Aplicações, dando as principais propriedades da transformada *wavelet* discreta *Haar*. São apresentadas aplicações elementares, em especial para o processamento de sinais. É incluído em forma preferencial a implementação no sistema de computação algébrica *Maple*<sup>1</sup> as *wavelets* 1D com algum renome, assim como uma aplicação em uma área importante do som.

**Palavras-chave:** *wavelets*, computação algébrica, reconhecimento de som.

## 1. Introdução

Com o crescente avanço da indústria tecnológica faz-se necessário que a cada dia sejam implementadas novas tecnologias em diversas áreas. Pensando nisso este trabalho tem o intuito de reconhecer o som emitido por um crustáceo. Sendo assim, os biólogos terão em mãos uma importante ferramenta no reconhecimento dos “seus animais”. Para algumas espécies é difícil colocar coleiras ou chips. Então trabalha-se com a possibilidade do reconhecimento de som desses animais. Pretende-se desenvolver um equipamento semelhante ao GPS, porém com o objetivo de identificar o animal pelo seu som.

Aplicando *wavelets* é possível reconhecer os sinais de som. Pretende-se criar um identificador de som que leia o sinal compacte e compare com o sinal original, tornando assim este sinal uma “impressão digital”.

---

<sup>1</sup> Produto da *Maplesoft*

As *wavelets* são determinadas funções com determinadas propriedades que as tornam adequadas a servirem de base para decomposição de outras funções.

As *wavelets* surgem frente à insuficiência da análise de Fourier em analisar sinais que não satisfazem as chamadas condições de *Dirichlet*<sup>2</sup>, por exemplo, para sinais onde existam variações abruptas.

As transformadas *wavelets* representam uma nova ferramenta matemática cujo desenvolvimento aconteceu principalmente nos últimos 25 anos do século XX. São eficientes para a análise local de sinais dependentes do tempo, não estacionárias e de variação abrupta. A primeira família de *wavelets* foi descoberta pelo matemático húngaro Alfred Haar em 1909, ao estudar novas famílias de funções ortogonais. O nome e a noção de wavelet - em sua forma atual - deve-se a equipe de estudos do Centro de Física Teórica de Marsella, composto pelo geofísico *Jean Morlet*, o matemático *Yves Meyer*, que trabalharam com a orientação do físico *Alex Grossman*. Foram criadas famílias de funções com características de ortogonalidade e com suporte compacto, que foram utilizadas para decompor um sinal transiente não estacionário e variação abrupta permitindo desta forma realizar uma análise eficiente do sinal.

Coube a *Ingrid Daubechies* um avanço muito importante na teoria de *wavelets*. Os estudos de *Daubechies* começaram a ser publicados em 1988. Em conjunto com *Stephane Mallat*, *Daubechies* desenvolveu estudos orientados ao processamento de sinais discretos.

A idéia fundamental da teoria de *wavelets* é realizar análises conforme seja a escala. O interesse crescente da teoria *wavelets* hoje em dia se deve a capacidade de representar sinais que possuem características diferentes para instantes e domínios espaciais diferentes. Assim, como pode ser visto mais adiante, podem ser implementadas de uma maneira computacional eficiente, graças ao que se denomina análise de multiresolução.

Diversas áreas da ciência e da tecnologia aplicam esta teoria com êxito, entre as quais estão a astronomia, a acústica, a engenharia nuclear, a codificação de subbandas, o processamento de sinais e imagens, a neurofisiologia, a música, as imagens de ressonância magnética, a discriminação de som, a óptica, os fractais, a turbulência, a

---

<sup>2</sup> As condições de *Dirichlet* são:

- ✓ D1.:  $f$  é uma função periódica de período  $T$  contínua em cada intervalo fundamental.
- ✓ D2.: A derivada  $f'$  também é periódica em cada período fundamental.

A variação em cada ponto de descontinuidade é finita.

predição sísmica, a visão humana, além de aplicações na matemática pura como a solução de equações diferenciais em derivadas parciais.

O trabalho está organizado da forma seguinte: na seção 1 foi dada uma breve introdução sobre o assunto, já na seção 2 será dada uma explicação básica da terminologia das *wavelets* e na seção 3 é feita à análise do som emitido pelo crustáceo lagostim-vermelho utilizando as *wavelets* mediante o software *Maple* e *Sound Ruler*.

## 2. Wavelets e multiresolução

**Definição 2.1:** Seja  $\psi : D \rightarrow \mathfrak{R}$  uma função do tipo  $\mathcal{L}_2(\mathbb{R})$ , denominada *função geradora*, de tal forma que as funções  $\psi_{a,b}$  definidas por transformações de dilatação (ou contração) e de translação a partir de  $\psi$

$$\psi_{a,b}(t) = \frac{1}{|a|^{1/p}} \psi\left(\frac{t-b}{a}\right), \text{ com } p > 0, a, b \in \mathfrak{R}, a \neq 0 \quad (1)$$

formam uma base de funções. A família de *wavelets* são as funções  $\psi_{a,b}$ , sendo com frequência de valor  $p = 2$  o mais usado.

Também se emprega a fórmula geradora

$$\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k), \text{ com } j, k \in \mathfrak{R} \quad (2)$$

Interessa saber se dada uma determinada família de funções  $\{\psi_{a,b}\}_{(a,b) \in \mathfrak{R}^* \times \mathfrak{R}}$  pode empregar-se para realizar uma decomposição de uma função dada  $f(t)$ . Nessa direção segue a seguinte definição

**Definição 2.2:** *Transformada wavelet contínua.* Seja  $\{\psi_{a,b}\}_{(a,b) \in \mathfrak{R}^* \times \mathfrak{R}}$  uma família de funções definida por (1), e dada uma função quadrada  $f(t)$  integrável, a transformada wavelet contínua de  $f \in \mathcal{L}_2(\mathbb{R})$  a decomposição de  $f$  mediante o conjunto de funções base  $\psi_{a,b}$ :

$$\psi_f(a,b) = \int_D f(t) \psi_{a,b}^*(t) dt \quad (3)$$

onde \* representa o conjugado complexo, e  $D$  o suporte compacto de  $\psi(t)$ , a função geradora.



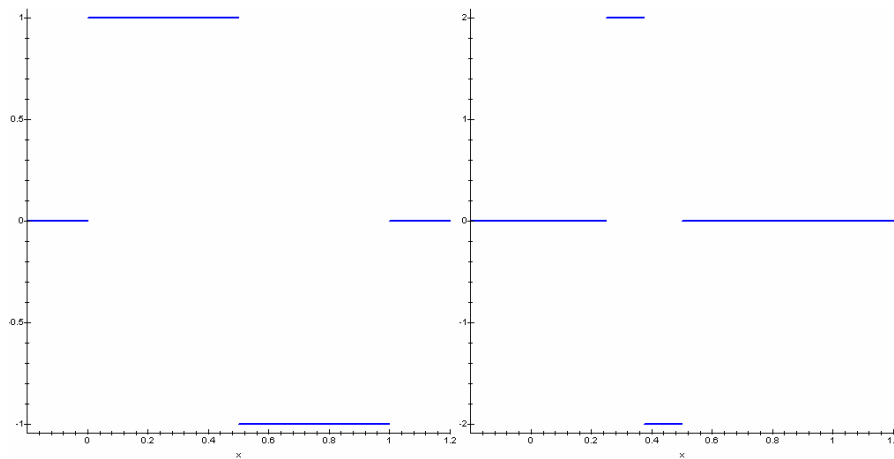
A constante  $\frac{1}{\sqrt{a}}$  em (1) se utiliza como constante de normalização. A família de *wavelets* é normalizada quando

$$\int_{\mathfrak{R}} |\psi_{a,b}(t)|^2 dt = \int_{\mathfrak{R}} |\psi(t)|^2 dt = 1, \quad (4)$$

de tal forma que todas as funções wavelet mantenham a mesma energia.

Substituindo (1) na fórmula da transformada wavelet (3) se pode escrever a transformada de  $f(t)$  como uma correlação entre o sinal e as *wavelets* do tipo  $f\left(\frac{t}{a}\right)$  da forma seguinte:

$$\psi_f(a,b) = \frac{1}{\sqrt{a}} \int_D f(t) \psi^*\left(\frac{t-b}{a}\right) dt \quad (5)$$



**Figura 1** - A função geradora Haar  $\psi$  e suas funções  $\psi_{1,0}$  e  $\psi_{1,1}$

As transformadas de Fourier das funções *wavelets* são

$$\hat{\psi}_{a,b}(\omega) = \int_{\mathfrak{R}} \frac{1}{\sqrt{a}} \psi^*\left(\frac{t-b}{a}\right) e^{-i\omega t} dt = \sqrt{a} \hat{\psi} \dots (a\omega) e^{-i\omega b}, \quad (6)$$

onde  $\hat{\psi}(\omega)$  é a transformada de Fourier da função geradora  $\psi(t)$ . No domínio da frequência, as *wavelets* estão afetadas em um fator de escala  $\frac{1}{a}$  e estão multiplicadas por um fator de fase em  $e^{-i\omega b}$  e ainda por um fator de normalização de  $\sqrt{a}$ .

**Exemplo 2.1:** *Wavelets Haar.*

Considere a seguinte função  $\psi(t) = \chi_{[0,1]}(2t) - \chi_{[0,1]}(2t - 1)$ , sendo  $\chi_{[0,1]}$  a função característica de intervalo  $[0, 1)$ .

As *wavelets Haar* definidas por  $\psi_{j,k}(t) = 2^{j/2} \psi(2^j t - k)$ , com  $j, k \in \mathbb{Z}$ , formam uma base ortogonal de  $\mathcal{L}_2(\mathbb{R})$ . Para  $k = 0$  considerasse a família  $\{\psi_{j,0} \mid j \in \mathbb{Z}\}$ ; se observa que quanto maior é  $j$ , menor será o conjunto *cozero*<sup>3</sup>  $\left[0, \frac{1}{2^j}\right]$ .

A correspondente transformada wavelet discreta *Haar* se define da seguinte forma sobre um sinal discreto  $s = [s_1, s_2, \dots, s_n]$ : se geram dos subsinais de longitude  $\frac{n}{2}$ , denominados  $a_1$  e  $d_1$  da seguinte maneira:

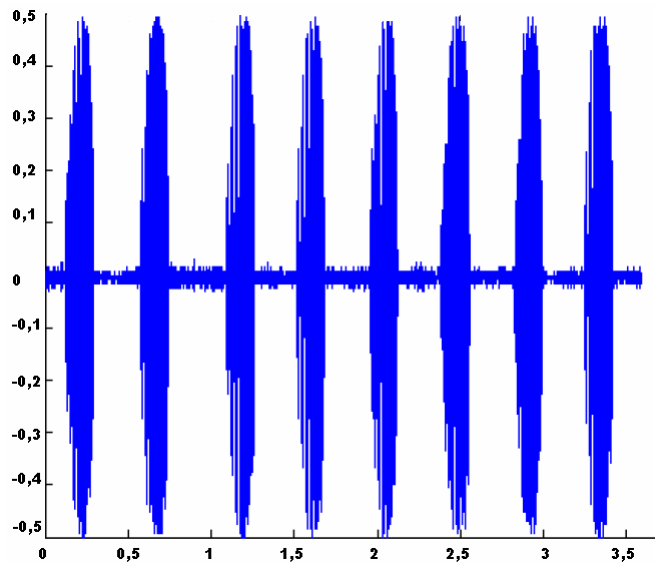
$$a_1 = \left[ \frac{s_1 + s_2}{\sqrt{2}}, \frac{s_3 + s_4}{\sqrt{2}}, \dots, \frac{s_{n-1} + s_n}{\sqrt{2}} \right] \quad (7)$$

e

$$d_1 = \left[ \frac{s_1 - s_2}{\sqrt{2}}, \frac{s_3 - s_4}{\sqrt{2}}, \dots, \frac{s_{n-1} - s_n}{\sqrt{2}} \right] \quad (8)$$

## 2. Modelo matemático das características físicas fundamentais do som

Utilizando um sistema de computação algébrica *Maple* foi possível criar o sinal do som de um crustáceo que pode ser visto na Figura 2.

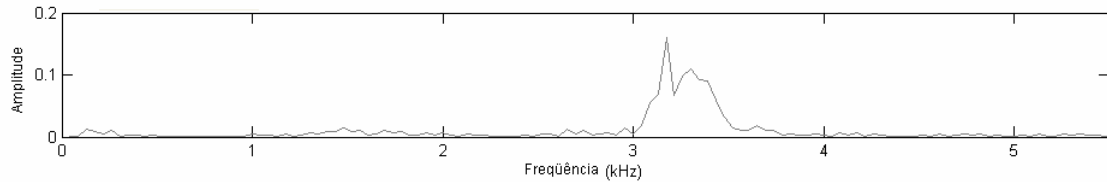


**Figura 2** – Sinal do som de um crustáceo.

Antes de aplicar a Transformada *Haar* ao sinal do som é possível, mediante o *software Sound Ruler* fazer algumas análises clássicas deste sinal, mediante o oscilograma e espectrograma.

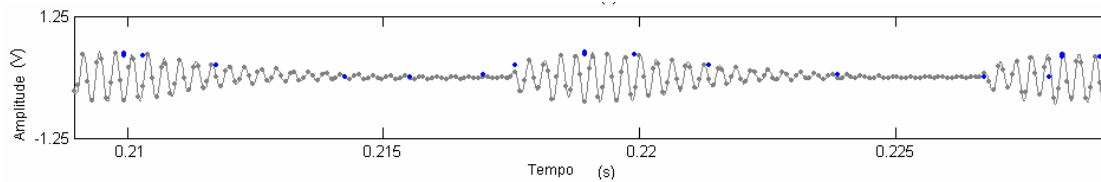
<sup>3</sup> O conjunto *cozero* de uma função é o conjunto de todos os valores  $t$  tais que o valor  $f(t)$  é diferente de zero. Simbolicamente se representa:  $\{t \in \mathbb{R} \mid f(t) \neq 0\}$ .

O *Sound Ruler* é uma ferramenta de software para medir e ilustrar sons e para ensinar acústica. Propicia o controle da análise manual e a objetividade e velocidade da análise automatizada. É especialmente útil na análise de sons simples e repetitivos. Estes sons podem ser medidos rapidamente e com grande detalhe, pulso por pulso.



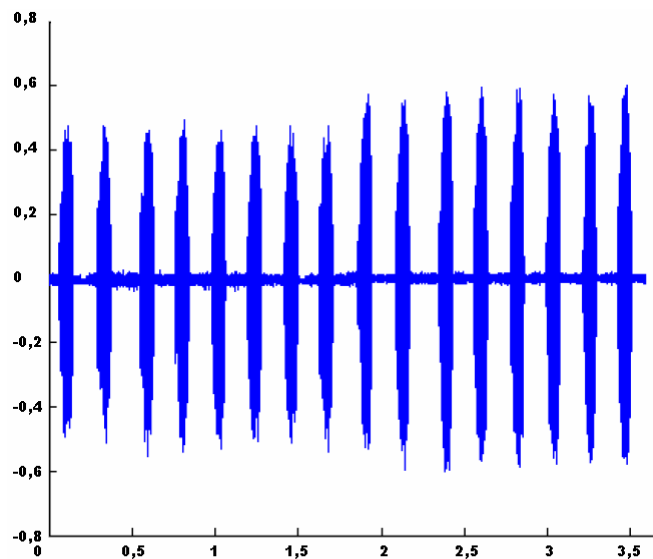
**Figura 3** – Espectrograma do som produzido pelo crustáceo.

Através do espectrograma pode-se analisar a “forma” que tem o som emitido por esse animal, começa-se desde aqui a analisar a “impressão digital” deste animal.



**Figura 4** – Oscilograma do som produzido pelo crustáceo.

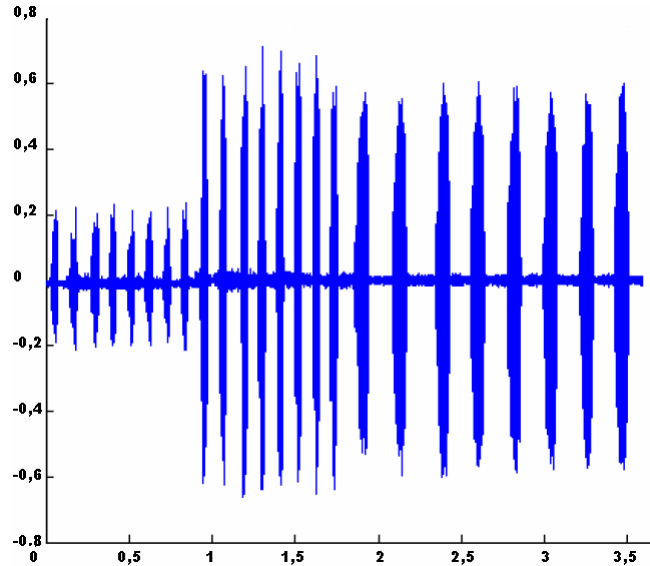
Através da aplicação da transformada discreta *Haar* de forma consecutiva, três vezes, compactou-se o sinal e o mesmo na forma compactada. Após ter o sinal do som é aplicada a Transformada *Haar* e temos como resultado o gráfico da Figura 3.



**Figura 5** - Transformada Haar aplicada ao sinal original.

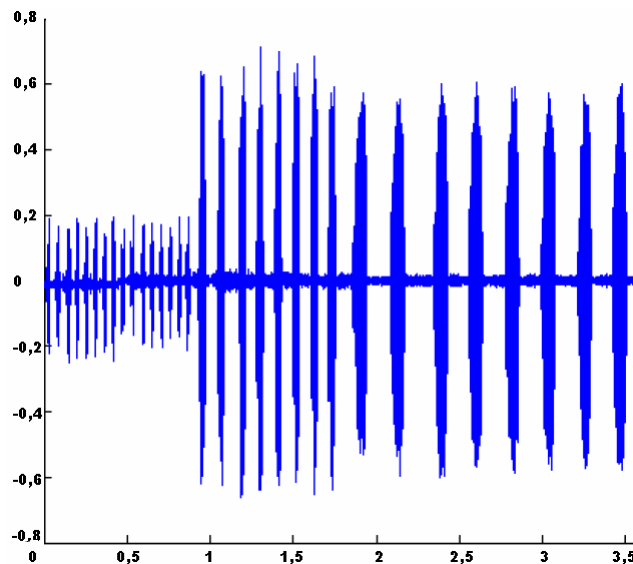
Comparando a Figura 2 com a Figura 5, pode-se ver que o sinal compactado reduziu a sua amplitude.

Vamos ver o que acontece se compactarmos mais uma vez esse sinal. O resultado aparece na Figura 6.



**Figura 6** – A transformada Haar aplicada por segunda vez ao sinal original.

Aplicando-se mais uma vez a transformada *Haar* ao som emitido pelo crustáceo obtém-se a Figura 7.



**Figura 7** – Terceira compactação mediante Transformada Haar.

Com certeza, após a terceira compactação, se tiver um banco de dados pode identificar-se o som do crustáceo.

É importante ressaltar a energia do sinal que se conserva nas três compactações, o que indica que não perderam-se as propriedades do sinal. As propriedades do sinal são

preservadas nas compactações o que garante que na reconstituição do sinal obter-se-á o sinal original.

Este trabalho esta em fase de desenvolvimento, novas *wavelets* serão aplicadas com o intuito de melhorar ainda mais a compactação do sinal.

### **Referências Bibliográficas:**

BACHMAN, G., NARICI, L. and BECKESTEIN, E.: **Fourier and wavelet analysis**, Springer – Verlag, (2000).

DAUBECHIES, I.: **Tem lectures on wavelets**, CBS – NSF Regional Conferences in Applied Mathematics, 61, SIAM, (1992).

HAAR, A.: **Zur theorie der orthogonalen funktionen systems**, Math. Ann, 69, (1910), 331 – 371.

HERNÁNDEZ, E.: **Ondículas y tecnologia**, Bol. Soc. Esp. Mat. Apl., 25, (2003), 39 – 54.

GABOR, D.: **Theory of communication**, J. Inst. Elect. Eng., 92 (III), London, (1946), 429 – 457.

GOMES, J. e VELHO, L.: **Computação Gráfica**, Instituto de Matemática Pura e Aplicada, vol. 1, Rio de Janeiro, (1998).

MALLAT, S.: **A wvelet tour of signal processing**, Academic Press, (1998).

MEYER, Y.: **Ondelettes et opérateurs, I: Ondelettes**, Hermann, Paris, (1990).

PAZOS, Rubén E. Panta. **Teoria wavelets e suas aplicações**, XXV Colóquio de Matemática, Sociedade Matemática Peruana, Lima, (2007).

POPULARIKAS, A. D. et. al.: **The transforms and applications handbook**, CRC Press, Boca Raton, FL., (1995).

STRANG, G.: **Wavelets and dilation equations: a brief introduction**, SIAM Review, 31 (4), (1989), 614 – 627.

# MODELAGEM MATEMÁTICA DOS PROCESSOS DE EVAPORAÇÃO E TRANSPIRAÇÃO NO SISTEMA SOLO - PLANTA

Angéli Cervi, Pedro Augusto Pereira Borges

Universidade Regional do Noroeste do Estado do Rio Grande do Sul – UNIJUÍ,  
agcervi@yahoo.com.br; Universidade Regional do Noroeste do Estado do Rio Grande do Sul –  
UNIJUÍ, pborges@unijui.edu.br

## **Resumo:**

Na agricultura irrigada é fundamental o conhecimento das necessidades hídricas das culturas, ou seja, o estudo das variações do teor de água do solo juntamente com a absorção da água pelas plantas é de grande interesse para pesquisas em sistemas agro-florestais e agrícolas, pois o movimento de nutrientes depende basicamente do movimento da água no solo. O presente trabalho tem por objetivo desenvolver um modelo matemático para determinar o teor de água retirada do solo pela planta, considerando os processos de irrigação e evapotranspiração. O monitoramento do teor de água em colunas de solo submetidas ao ar ambiente somente na superfície foi feito através de transdutores térmicos, obtendo-se curvas de teor de água em função do tempo e da posição na coluna. A evaporação foi calculada usando as medidas do teor de água na superfície. O problema da dinâmica da água na coluna de solo foi modelado através da equação de Richards, em uma dimensão, usando o Método das Diferenças Finitas. Com estas informações, a transpiração foi calculada resolvendo o problema inverso.

**Palavras-chaves:** teor de água, evapotranspiração, problemas inversos.

## **Introdução**

O crescente uso da água e a redução de sua disponibilidade têm se tornado preocupação cada vez maior no mundo todo. No Brasil, a agricultura irrigada consome uma grande parcela da água usada nas diversas atividades humanas, razão por que se ressalta que, na maioria das áreas irrigadas, é comum observar ausência de manejo racional da água resultando, geralmente, em aplicação excessiva, com desperdício de água e energia, além da ocorrência de problemas ambientais, ou em deficiência hídrica

para as plantas, com baixa produtividade e prejuízos econômicos ao produtor. Práticas adequadas de manejo da irrigação contribuem para aumentar a produtividade e melhorar a qualidade dos produtos agrícolas, para minimizar o uso de água e preservar os recursos hídricos (Oliveira *et al.*, 2007).

Assim, o solo, que funciona como um reservatório aberto para a atmosfera, exerce função importante para o crescimento das plantas, fornecendo água, ar e nutrientes.

O estudo sobre o solo é muito amplo, pois envolve uma série de variáveis, como por exemplo, estrutura, compactação, textura, teor de água, densidade, condutividade hidráulica, entre outros. Neste contexto a água é um elemento importante em quase todos os mecanismos de movimento de nutrientes, e o conhecimento detalhado do movimento da mesma durante o desenvolvimento de uma cultura fornece elementos importantes para aperfeiçoar a produtividade. O sistema água/solo/planta engloba uma série de processos de transferência de nutrientes que viabilizam o crescimento e reprodução das espécies vegetais.

Segundo Reichardt (1990), o solo é um sistema complexo constituído de materiais sólidos, líquidos e gasosos. As partículas sólidas formam um arranjo poroso tal que os espaços vazios são chamados poros, e têm a capacidade de armazenar líquidos e gases. A parte líquida constitui-se essencialmente de água, contendo minerais dissolvidos e materiais orgânicos solúveis. Ela ocupa parte dos espaços vazios entre as partículas sólidas. Esta água é absorvida pelas raízes das plantas ou é drenada para camadas de solo mais profundas e, por isso, precisa ser periodicamente repostada pela chuva ou pela irrigação, para garantir uma produção vegetal adequada. Por isto a importância agrícola do conhecimento deste reservatório de água para as plantas e dos princípios que governam seu funcionamento. Já a parte gasosa ocupa os espaços vazios entre as partículas sólidas não ocupados pela água.

A fase líquida e gasosa são complementares, isto é, a máxima presença de uma implica na ausência da outra. Assim, a fase líquida (água) pode estar presente nos poros do solo completa ou parcialmente, e o solo é chamado de saturado ou não saturado. Quando o espaço poroso está totalmente cheio de água, o solo é dito saturado, já quando o espaço poroso possui parcialmente água e parcialmente ar, o solo é dito não saturado (Libardi, 1999).

O movimento da água no solo pode ocorrer por influência de diferentes forças. Em solos muito saturados e com presença de pequenos canais a água pode escoar com

certa facilidade, movimentada basicamente pelas forças gravitacionais (percolação). Para solos apenas saturados a tensão superficial da água contida nesses pequenos canais poderá exceder a força gravitacional e provocar movimentos ascendentes (capilaridade). Em solos não saturados as forças gravitacionais e a tensão superficial não são significativas para o movimento da água, porque não há massa suficiente para preencher os canais. A água acaba sendo então absorvida pelas partículas de solo e também passando de partícula para partícula, devida à diferença de teor de água entre elas.

A transferência de água na forma de vapor do sistema solo-planta para a atmosfera é denominada de evapotranspiração (Reichardt, 1990), sendo que do solo para a atmosfera denominada evaporação e da planta para a atmosfera transpiração.

Assim, a água retirada do solo pelas raízes se dá pela água que a planta retém no seu metabolismo somado à água cedida para a atmosfera por transpiração.

O estudo desses processos de movimento dos nutrientes em plantas, e o conhecimento da dinâmica da água no solo são de amplo interesse em pesquisas de monitoramento de florestas, sistemas agro-florestais e agrícolas.

O presente trabalho é uma pesquisa sobre o movimento da água no solo com presença de um sistema radicular. O problema proposto é a estimação da quantidade de água retirada do solo pela planta, com base em dados experimentais próprios do monitoramento do teor de água do solo.

### **Descrição dos experimentos**

A parte experimental da pesquisa está sendo executada na casa de vegetação do campus da UNIJUÍ, e nos laboratórios de Física do Solo e de Medidas Físicas para Modelagem Matemática desta mesma universidade, na forma de dois tipos de experimentos: curva de retenção da água e monitoramento do teor de água do solo com presença de planta.

A determinação da curva característica está sendo realizada em laboratório, a mesma se relaciona às características intrínsecas do solo, que expressa o teor de água em função do potencial matricial da água no solo (Libardi, 1995).

O segundo experimento consiste em monitorar o teor de água do solo durante 30 dias. Uma planta foi instalada no centro de um cilindro metálico de raio 20 cm e altura 40 cm. Nove sensores de teor de água foram instalados conforme ilustra a Figura 1.



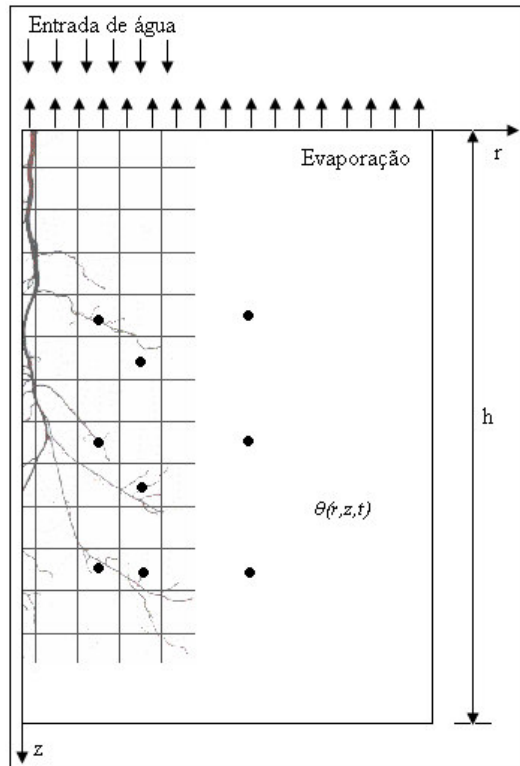


Figura 1: Sistema radicular com sensores.

Com o objetivo de testar os modelos propostos e apresentar resultados para esta fase do trabalho, foi criada uma coleção de dados artificiais da distribuição do teor de água nas direções  $r$  e  $z$  em função do tempo.

### Modelo Matemático

A quantidade de água absorvida pelas raízes e o movimento da água no solo tem sido pesquisados empírica e teoricamente (Reichardt, 1985 e 1990, Boast e Robertson, 1982; Kanemasu *et al*, 1976; Klocke *et al*, 1990; Ritchie e Johnson, 1990; e Rodríguez *et al*, 1997 entre outros) associada ao fenômeno de evapotranspiração.

O modelo proposto neste trabalho é a determinação da quantidade de água retirada do solo pelas raízes de plantas, considerando os processos de evapotranspiração e irrigação, obtido a partir da complementação da equação de Richards com um termo fonte (absorção da água pelas raízes) e condições de fronteiras adequadas. Em coordenadas cilíndricas tem sua formulação matemática para o caso 2D da seguinte forma:

$$\frac{\partial \Theta}{\partial t} = \frac{1}{r} \frac{\partial}{\partial r} \left[ K_r(\Theta) r \frac{\partial \psi}{\partial r} \right] + \frac{\partial}{\partial z} \left[ K_z(\Theta) \frac{\partial \psi}{\partial z} \right] + S(r, z, t) \quad (1)$$

com as seguintes condições de contorno:

$$\frac{\partial \Theta}{\partial z}(r, 0, t) = \Phi_i \quad \text{para } 0 < r < r_i \text{ e } t < t_i \quad (1.1)$$

$$\frac{\partial \Theta}{\partial z}(r, 0, t) = \Phi_e \quad \text{para } 0 < r < R \text{ se } t > 0 \quad (1.2)$$

$$\frac{\partial \Theta}{\partial z}(r, h, t) = 0 \quad \text{para } 0 < r < R \text{ e } t > 0 \quad (1.3)$$

$$\frac{\partial \Theta}{\partial r}(R, z, t) = 0 \quad \text{para } 0 < z < h \text{ e } t > 0 \quad (1.4)$$

$$\Theta(r, z, 0) = \Theta_0 \quad \text{para } 0 < r < R \text{ e } 0 < z < h \quad (1.5)$$

Onde:

$\Theta$  é o teor de umidade do solo (adimensional)

$r_i$  é o raio de irrigação (m)

$t_i$  é o tempo de irrigação (h)

$R$  é o raio do tubo (m)

$h$  é a altura do tubo (m)

$t$  é o tempo (s)

$\Phi_i$  é o fluxo de água de irrigação ( $m^3/h$ )

$\Phi_e$  é o fluxo de água de evaporação ( $m^3/h$ )

$S$  é a absorção de água pelas raízes ( $1/h$ )

$k_r$  e  $k_z$  são as condutividades hidráulicas nas direções r e z, respectivamente (m/s)

$z$  é a profundidade (m)

O potencial matricial do solo será obtido experimentalmente através do ajuste da curva característica do solo (curva de retenção). Um modelo empírico bastante versátil foi proposto por Van Genuchten (1980):

$$\psi_m = \frac{1}{a} \sqrt[n]{\Theta^{\frac{-1}{m}} - 1} \quad (2)$$

Onde:

$\Psi_m$  é o potencial matricial

$\Theta$  é o teor de água adimensional

$a, m, n$  são parâmetros do ajuste de curvas

Mualem (1976) e Van Genuchten (1980) desenvolveram a equação para determinar a condutividade hidráulica, usando os parâmetros da curva de retenção de água no solo:

$$k(\Theta) = k_0 \Theta^l \left[ 1 - \left( 1 - \Theta^{\frac{1}{m}} \right)^m \right]^2 \quad (3)$$

Onde:

$k_0$  é a condutividade hidráulica do solo saturado

$m$  é o parâmetro da curva característica

$l$  é um parâmetro considerado por Mualem (1976) igual a 2

A absorção de água do solo pelas raízes será calculada usando a metodologia do Problema Inverso (PI), acoplada ao Problema Direto (PD) de movimento de água no solo e crescimento de raízes.

O PD foi resolvido numericamente usando o Método das Diferenças Finitas, com valores conhecidos empiricamente para a condição de fronteira na superfície. O PI consiste em resolver o PD estimando valores para a função de absorção  $S(r, z, t)$  de água pelas raízes, visando à descrição ótima do comportamento dos dados experimentais de teor de água no interior do solo, obtidas experimentalmente. Foi usado o Método de Procura em Rede, que consiste em estimar um intervalo de valores em seqüência crescente para os parâmetros a ajustar, admitindo que os valores ótimos estão contidos no intervalo estimado, localizar um ponto mínimo e ponto máximo e pesquisar um ponto ótimo e restringir o intervalo dos parâmetros  $a$  e  $b$  e encontrar um novo ponto ótimo, até chegar ao fluxo ótimo.

### **Resultados Preliminares**

A análise dos resultados do Problema Direto visa mostrar a coerência do modelo com o que pode ser esperado experimentalmente, sobre a distribuição do teor de água do solo, considerando as variações na posição e intensidade de fluxo nos ápices em

função do espaço e do tempo. A Figura 2 apresenta a superfície de teor de água calculada com fluxo de água constante em todos os ápices e em relação ao tempo, para um intervalo de tempo de  $30h$ . A localização dos ápices foi determinada artificialmente, porque os dados experimentais ainda não estão disponíveis. A função da evaporação foi considerada linear após ajuste de curvas dos dados de teor de água da primeira camada de solo, conforme hipótese descrita em Vione (2007). A ação da evaporação nas camadas superiores de solo é constatada pelo baixo teor de água em relação às camadas inferiores, o que forma uma curva  $\Theta(z)$ . Se estas curvas forem construídas para cada instante de tempo mostrariam a variação do teor de água e o deslocamento da água na direção vertical, de baixo para cima, em função do gradiente de teor de água ocasionado pela evaporação na superfície. Observa-se também a redução do teor de água na região onde se localizam os ápices, que é o efeito da transpiração.

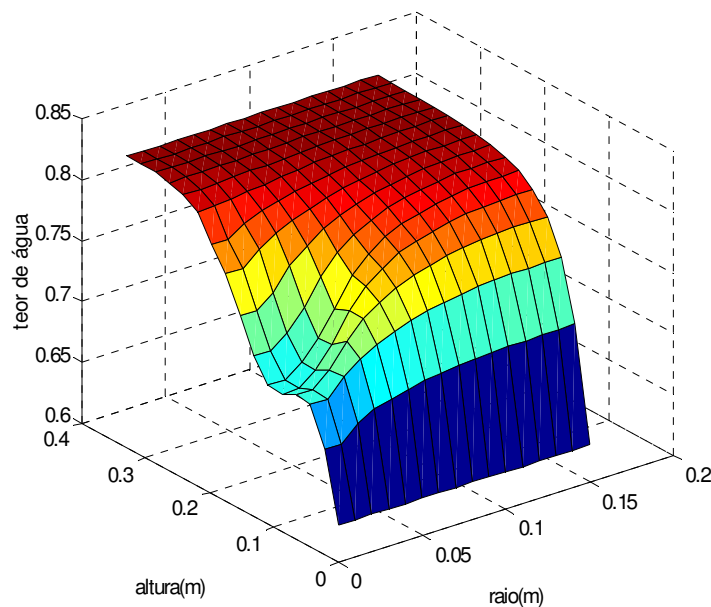


Figura 2 - Superfície do teor de água:  $t = 30 h$ .

A Figura 3 apresenta os dados experimentais fictícios e os resultados do problema direto, com transpiração constante ótima calculada pelo problema inverso. A concordância para  $0 m$  era esperada porque o modelo de evaporação foi implantado em função do teor de água desta posição. Para  $0,25 m$  existe boa correspondência devido à tendência uniforme dos dados e porque não há variação significativa do teor de água.

Para as demais profundidades observa-se que os resultados obtidos são apenas razoáveis. Porém, deve-se observar que os dados experimentais são fictícios e que, provisoriamente, têm somente função ilustrativa e de teste da convergência do algoritmo do problema inverso.

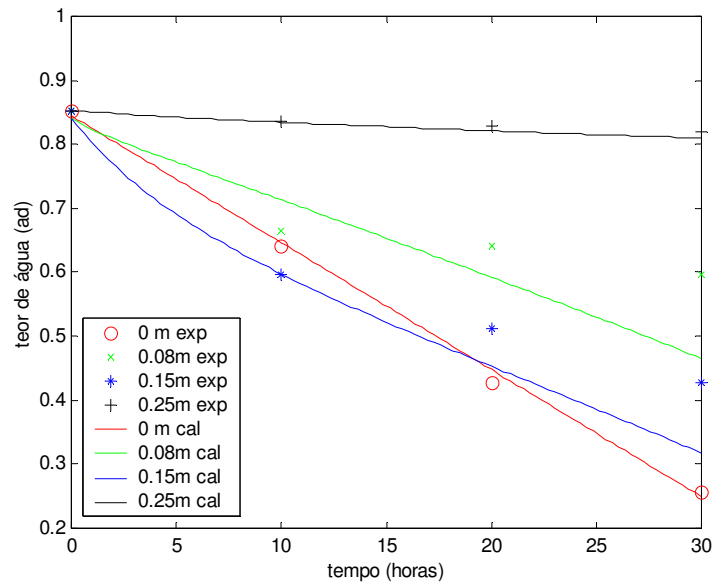


Figura 3 - Dados experimentais fictícios e calculados:  $t = 30$  h.

### Agradecimentos

Nossos agradecimentos à UNIJUI pela infra-estrutura cedida e a CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) pela bolsa de mestrado, apoios fundamentais para a realização do presente trabalho.

### Referências Bibliográficas

BOAST, C. W.; ROBERTSON, T. M. A micro-lysimeter method for determining evaporation from bare soil: description and laboratory evaluation. **Soil sci. Soc. Am. J.**, Madison, v.46, p. 689-696, 1982.

KANEMASU, E.T., Stone, L.R., POWERS, W.L. Evapotranspiration model tested for soybean and sorghum. **Agron. J., Madson**, v. 68, p. 569-572, 1976.

KLOCKE, D.L., MARTIN, D.L., TODD, R.W. et al. Evaporation measurements and predictions from soils under crop canopies. **Trans. ASAE**, St. Joseph, v. 33, n. 5, p. 1590-1596, 1990.

LIBARDI, P.L. **Dinâmica da água no solo**. Piracicaba, 2.ed, 509p., 1999.

\_\_\_\_\_ **Dinâmica da água no solo**. Piracicaba, 497p. 1995.

MUALEM, Y. A new model for predicting the hydraulic conductivity of unsaturated porous media. **Water Resour. Res.**; Washington, v.12, p. 513-522, 1976.

OLIVEIRA, Rubens A. et al. Desempenho do irrigâmetro na estimativa da evapotranspiração de referência. **Rev. bras. eng. agríc. ambient.**, Campina Grande, v.12, n.2, p.166-173, 2008. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S1415-43662008000200009&lng=pt&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S1415-43662008000200009&lng=pt&nrm=iso). Acesso em: 11 Mar 2008.

REICHARDT, K. **A água em sistemas agrícolas**. São Paulo: Manole Ltda, 1990.

\_\_\_\_\_ **Processos de transferência no sistema solo-planta-atmosfera**. 4.ed. Campinas: Fundação Cargill, 445p, 1985.

VAN GENUCHTEN, M. A closed form equation for predicting the hydraulic conductivity of unsaturated soils. **Soil Science Society of America Journal**, Madison, v. 44, n. 5, p. 892-898, Sept/Oct. 1980.

VIONE, M. T. **Modelagem Matemática da evaporação da água no solo**. Dissertação de Mestrado, Unijuí, Ijuí-RS, p. 77, 2007.

# MODELAGEM MATEMÁTICA DAS PROPRIEDADES DOS PRODUTOS DE COMBUSTÃO DE COMBUSTÍVEIS VEGETAIS, CONSIDERANDO O EQUILÍBRIO QUÍMICO

Danieli Michael<sup>1</sup>

Ângela Patricia Spilimbergo<sup>2</sup>.

Universidade Regional do Noroeste do Estado do Rio Grande do Sul - UNIJUÍ,

[danielimichael@yahoo.com.br](mailto:danielimichael@yahoo.com.br)

[patspi@unijui.edu.br](mailto:patspi@unijui.edu.br)

**Palavras-chave:** modelagem matemática, produtos de combustão, equilíbrio químico

## Resumo:

Os processos de combustão em diferentes instalações energéticas, como por exemplo, fornalhas para secagem de grãos, caldeiras, propulsores e fornos para cozimento de tijolos e telhas, são descritos muitas vezes por modelos que consideram que o meio reagente permanece no estado de equilíbrio químico. Entre esses modelos os mais conhecidos são Alemassov *et al.* (1973) e Gordon e McBride (1971). Eles levam em conta processos de dissociação em meios reagentes de alta temperatura e são descritos por um sistema de equações algébricas não lineares

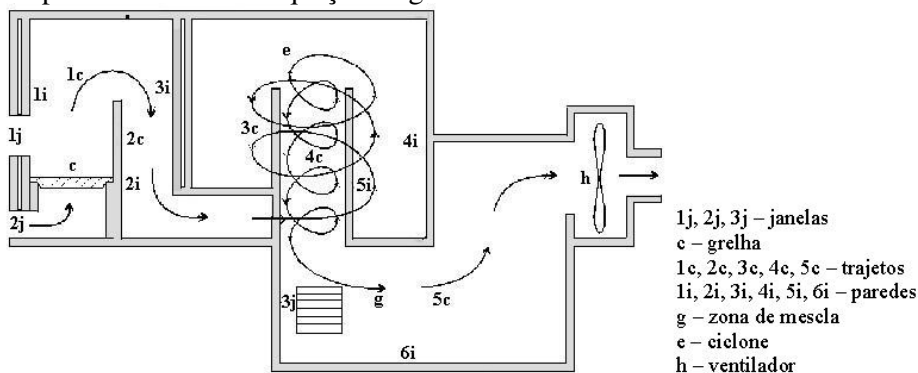


Figura 1. Esquema de uma fornalha para secagem de grãos

Neste trabalho foram determinadas as propriedades dos produtos de combustão em fornalhas (Fig. 1), que são instalações que se destinam a gerar ar quente com o objetivo de realizar a secagem de diferentes tipos de grãos, utilizando para isto combustíveis vegetais como, por exemplo, casca de arroz, palha de trigo e casca de

<sup>1</sup> Acadêmica do Curso de Matemática – Licenciatura. Bolsista de Iniciação Científica - PIBIC/UNIJUÍ

<sup>2</sup> Professora Orientadora

feijão. Para calcular essas propriedades foi utilizado o modelo Alemassov *et al.* (1973) em conjunto com os métodos das “grandes moléculas”. Essas propriedades são: composição dos produtos de combustão ( $r_q$ ), temperatura (T), calor específico ( $C_p$ ), massa molecular média ( $\mu$ ), viscosidade ( $\eta$ ), condutibilidade térmica ( $\lambda$ ), entre outras.

O modelo Alemassov *et al.* (1973) está descrito detalhadamente em Auth e Iskhacova (1996) e é constituído basicamente por três tipos de equações:

- equação da dissociação das moléculas (radicais) nos átomos:

$$\prod_i P_i^{a_{ij}} / P_j = K_j \quad j = 1, \dots, m \quad (1)$$

onde,  $P_i$  e  $P_j$  são as pressões parciais do átomo  $i$  e molécula (radical)  $j$ ,  $a_{ij}$  é a quantidade do átomo  $i$  em uma molécula (radical)  $j$  e  $K_j$  é a constante de dissociação pela pressão.

- equação da conservação da quantidade de átomos nos produtos de combustão:

$$\sum_j a_{ij} \cdot P_j + P_i = M_p \cdot b_{ip} \quad i = 1, \dots, n \quad (2)$$

onde,  $M_p$  é a constante de proporcionalidade que assegura  $P_q = n_q$  ( $P_q$  e  $n_q$  são respectivamente, a pressão parcial e a quantidade dos moles da  $q$ -ésima substância dos produtos de combustão) e  $b_{ip}$  é a quantidade do  $i$ -ésimo átomo na fórmula condicional do propelente (Spilimbergo, Castelli e Auth, 1999).

- equação de Dalton:

$$\sum_{q=1}^{m+n} P_q = P \quad q = i, j. \quad (3)$$

Se forem conhecidos os valores da pressão  $P$  e da temperatura  $T$  é possível determinar as grandezas  $P_i$ ,  $P_j$ , e  $M_p$ . Mas como regra para condições de combustão, a temperatura  $T$  é incógnita e neste caso, junto com (1)-(3), é necessário utilizar a equação da energia:  $I_p - I_{pc} = 0$ , onde,  $I_p$  e  $I_{pc}$  são as entalpias mássicas do propelente e dos produtos de combustão, respectivamente. As relações (1)-(3) fornecem um volumoso sistema de equações algébricas não lineares, e para sua resolução utiliza-se o método de Newton com algumas modificações para assegurar a convergência. A seguir são determinadas então as propriedades dos produtos de combustão, cujas principais são estão listadas a seguir.

- Frações molares:  $r_q = P_q / P$ .



- Massa molecular média:  $\mu = \sum_q r_q \cdot \mu_q$ , sendo  $\mu_q$  a massa molecular de cada molécula

(radical) ou átomo.

- Entalpia:  $I_{pc} = \sum_q (H_q \cdot r_q) / \sum_q (\mu_q \cdot r_q)$ .

- Entropia:  $S = \sum_q P_q \cdot (S_q^o - R_o \cdot \ln P_q) / (P \cdot \mu)$ , sendo  $R_o$  a constante universal de gás.

- Calores Específicos:

a) Calor específico “congelado” - corresponde ao calor específico habitual para mistura gasosa, sem que ocorram mudanças na composição dos produtos de combustão. É calculado através da expressão:  $C_{pf} = \sum (P_q \cdot C_{pq}) / P$ , onde  $C_{pq}$  é o calor específico a pressão constante para a q-ésima substância.

b) Calor específico “equilibrado”- sob altas temperaturas realiza-se o processo de dissociação, ocorrendo transformações nos produtos de combustão, para o qual é necessário energia complementar e  $C_{pe}$  leva em conta esse fenômeno, sendo calculado

por:  $C_{pe} = C_{pf} + \frac{1}{M_p \cdot T} \cdot \left[ \sum_q P_q \cdot I_q \cdot \left( \frac{\partial \ln P_q}{\partial \ln T} \right)_P \right] - \frac{1}{T} \cdot \left( \frac{\partial \ln M_p}{\partial \ln T} \right)_P$ , onde  $I_q$  é a

entalpia mássica da q-ésima substância.

O método que normalmente é utilizado para determinar o conteúdo da fase condensada nos produtos de combustão apresenta uma série de limitações, entre as quais podemos destacar regras de fases e convergência limitada. Devido a isso, foi utilizado no modelo Alemassov *et al.* (1973), o “método das grande moléculas” - GM (Auth e Iskhakova, 1996), cuja a essência proposta por Khudiacov (Alemassov *et al.*, 1973), está em considerar cada fase condensada como um conjunto de “moléculas grandes”, cada uma possuindo 1000 moléculas habituais, ocorrendo portanto, que cada substância, na fase condensada, simula um conjunto de GM comportando-se como na fase gasosa.

Utilizando o aplicativo existente foram realizados vários cálculos para determinar as propriedades dos produtos de combustão de combustíveis vegetais (casca de arroz, palha de trigo e casca de feijão). As informações necessárias sobre os combustíveis foram obtidas de Jenkins (1990). Nestes trabalhos, para cada combustível estão apresentados os dados:  $\Delta H$  (poder calorífico alto), o percentual mássicos dos átomos ( $g_i$ ) e o percentual mássicos dos resíduos ( $g_r$ ). Por exemplo:

Tabela 1. Dados para lenha-eucalipto

$\Delta H$ (MJ/kg)	$g_C$ (%C)	$g_H$ (%H)	$g_O$ (%O)	$g_N$ (%N)	$g_r$ (% resíduos)
19,23	48,18	5,92	44,2	0,39	1,12

Estes dados são suficientes para que sejam realizados os cálculos, mas para utilizá-los no aplicativo é necessário transformar a entalpia do combustível na escala que é utilizada em Alemassov et al. (1973) e Gordon e McBride (1971) e também incluir a “cinza” na fórmula condicional do combustível. Neste trabalho foi admitido que a “cinza” é constituída somente pela substância condensada  $SiO_2^*$  e que o  $O_2$  da cinza foi retirado do combustível (Spilimbergo *et al.*, 1998).

Na tabela 2 estão apresentadas as entalpias e as fórmulas condicionais para os três combustíveis pesquisados neste trabalho, indicados em Jenkins (1990) e a seguir estão apresentados alguns resultados numéricos (Fig. 2 a 5) sobre as propriedades dos produtos de combustão dos combustíveis pesquisados neste trabalho. As simulações numéricas foram realizadas para  $P=10^5$  Pa,  $\alpha_{ox} = 0.1, \dots, 2.5$  ( $\alpha_{ox}$  - coeficiente de excesso do oxidante (Spilimbergo, Castelli e Auth, 1999)), sendo o meio reagente formado pelas substâncias gasosas: H, O, N, C, Ar, Si, CO, OH, NO,  $H_2$ ,  $O_2$ ,  $H_2O$ ,  $CH_4$ ,  $N_2$  e condensadas  $SiO_2^*$  e  $C^*$ .

Tabela 2. Entalpias e Fórmulas Condicionais

Combustível	$I_c$ (kJ/kg)	C	H	O	N	Si
Palha de Feijão	-8141,22	3,58	5,59	2,29	0,07	0,09
Casca de Arroz	-5970,00	2,88	3,93	3,04	0,07	0,41
Palha de Trigo	-5339,86	3,60	5,00	2,84	0,07	0,19

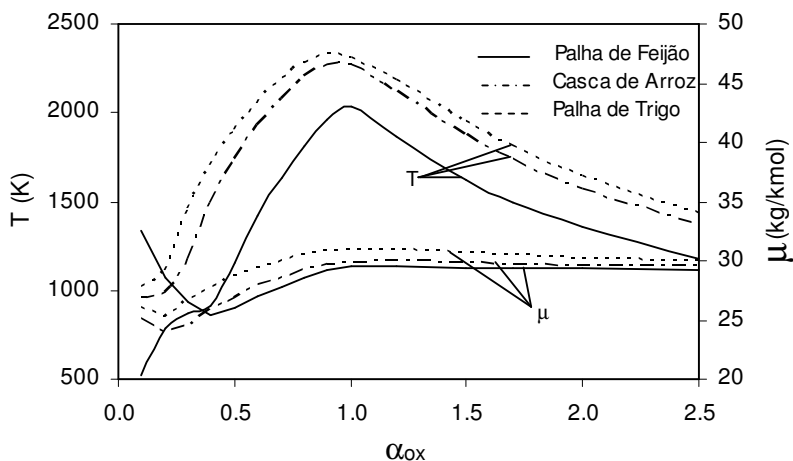


Figura 2. Variação da temperatura e massa molecular dos produtos de combustão

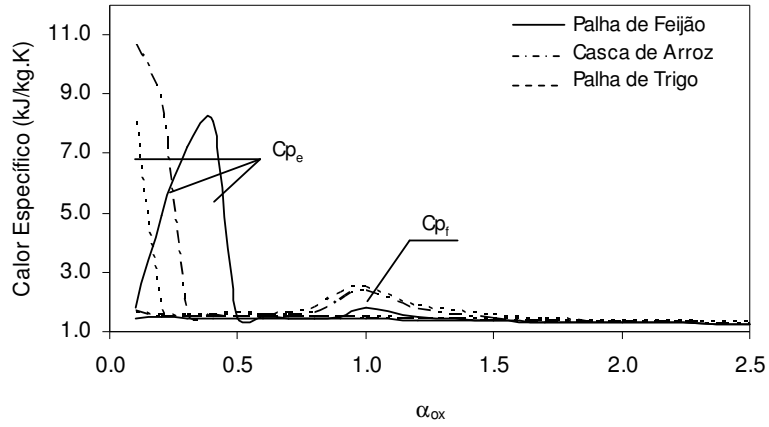


Figura 3. Alteração do calor específico “equilibrado” e “congelado”

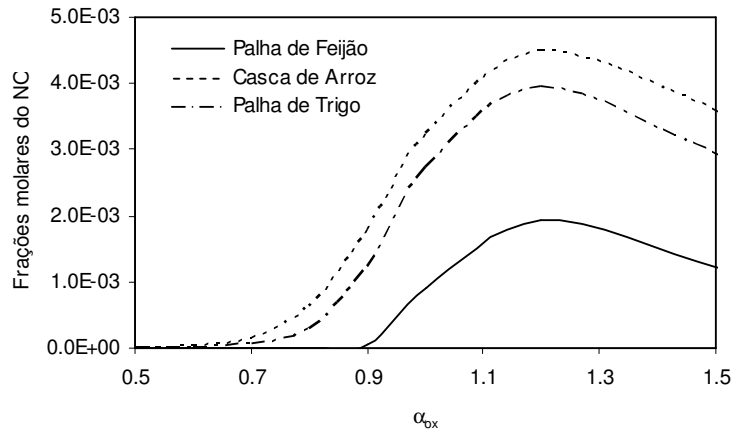


Figura 4. Alterações das frações molares do poluente NO

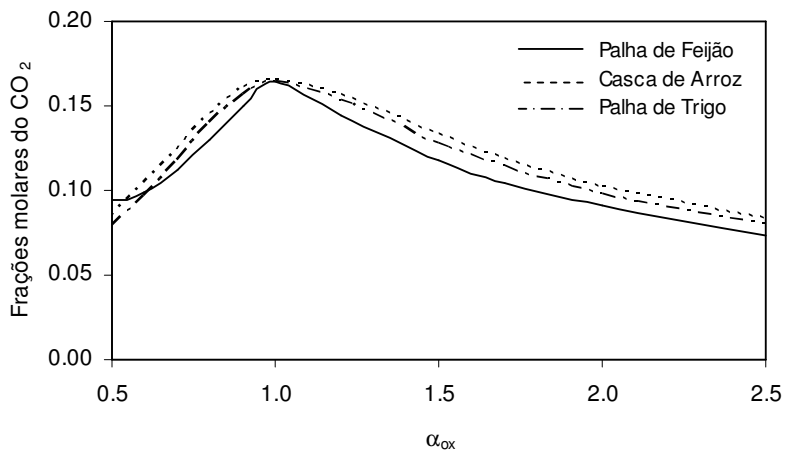


Figura 5. Alteração das frações molares do  $CO_2$

As Figuras 2 e 3 mostram as alterações da temperatura, da massa molecular e dos calores específicos “equilibrado” e “congelado” em função de  $\alpha_{ox}$  e as Fig. 4 e 5 as alterações das frações molares do CO<sub>2</sub> e do poluente NO, também em função de  $\alpha_{ox}$ .

### **Referências Bibliográficas**

ALEMASOV, V.E.; DREGALIN, A.F.e TISHIN, A.P.: **Propriedades termodinâmicas e termofísicas dos produtos de combustão**, VINITI, Moscou, (1973).

AUTH,C.J. e ISKHAKOVA, R.L.: **Pesquisa das propriedades termodinâmicas e termofísicas dos produtos de combustão de biomassa**, Proceedings of 6<sup>th</sup> Brazilian Congress of Engeneering and Thermal Sciences, Florianópolis, SC, Vol. III, (1996), 1769-1774.

GORDON, S., McBRIDE, B. J.: **NASA SP273 - Computer program for calcylation do complex chemical equilibrium compositions, rocket performancem incident and reflected schoks and chapman-jouguet**, Washington, (1971).

JENKINS, B. M.: **Fuel Properties for Biomass Materials**, Proceedings of International Symposium on Application and Management of the Energy in Agriculture, Indiana, (1990).

SPIILIMBERGO, A.P., CASTELLII, C.A. e AUTH, C.J.: **Simulação numérica das propriedades dos produtos de combustão de diferentes espécies de carvão**, Proceedings of the XX Computational Methods in Engineering (CILAMCE), São Paulo, (1999), CD-ROOM.

SPIILIMBERGO, A.P., AUTH, C.J., ISKHACOVA, R.L., ROSANELLI, M.: **Simulação numérica de processos de combustão utilizando os métodos das grandes moléculas e do meio local reagente**, Anais da Feira e Congresso de ar condicionado, refrigeração, aquecimento e ventilação do Mercosul, Porto Alegre, (1998), CD-ROOM.

# ESTUDO E CONSTRUÇÕES DE FRACTAIS

Eunice Cristina dos Santos <sup>1</sup>, Kelly Roberta Mazzutti Lübeck.

Universidade Estadual do Oeste do Paraná - UNIOESTE, ecristina41@hotmail.com; Universidade

Estadual do Oeste do Paraná - UNIOESTE, kellyrobertaml@gmail.com

## Resumo:

Por vários séculos o mundo natural era descrito como regular, a geometria de Euclides era a rainha, e através dela poderíamos descrever as árvores, as nuvens, os raios, sons e outros fenômenos naturais com seus cilindros, círculos, retas, quadrados, cones e demais formas geométricas. Porém, através de estudos percebeu-se que os fenômenos naturais não são regulares e, a partir disso, surgiram diversos problemas denominados “monstros matemáticos”, que em 1975 foram chamados de “fractais” por Benoit Mandelbrot. Sentindo a necessidade de incluir esta “nova geometria” na sala de aula e, também, proporcionar um estudo investigativo sobre a teoria dos fractais utilizamos os recursos do desenho geométrico e da geometria dinâmica do software livre iGeom, para auxiliar na investigação de tais teorias. Neste artigo mostraremos algumas propriedades de fractais clássicos como o Triângulo de Sierpinski, a Curva de Koch e o Floco de Neve, além de suas construções com o uso do software iGeom.

**Palavras-chave:** Geometria Dinâmica, Fractal, Dimensão de Hausdorff.

## Apresentação:

Os fractais foram introduzidos por Benoit Mandelbrot, em meados de 1975, como os subconjuntos de um espaço euclidiano cujas dimensões Topológicas e de Hausdorff não são iguais. Além desta definição não ser muito esclarecedora (principalmente para aqueles que estão tendo um primeiro contato com esta teoria) ela também é bastante restritiva

---

<sup>1</sup>Bolsista FPTI/PDTA.

e, como o próprio Benoit afirmou, ela provavelmente será substituída futuramente, mas por enquanto ainda é considerada como uma definição para os fractais.

A teoria fractal está diretamente relacionada com a dimensão de um fractal, que não é necessariamente inteira (sua dimensão é na maioria das vezes quebrada). Benoit Mandelbroit utilizou um exemplo para elucidar o conceito de dimensão baseado nas diferentes perspectivas de se observar um novelo de lã. Quando observamos um novelo de lã a uma grande distância vemos apenas um ponto, um ente sem dimensões pela geometria euclidiana. Ao nos aproximarmos vemos um objeto em três dimensões. Mas, se nos aproximarmos ainda mais, percebemos um elemento de dimensão um (a linha) envolto de si mesmo. Podemos perceber, dessa forma, que apesar do novelo de lã ser um emaranhado de linhas (que possui dimensão 1) ele ocupa um lugar no espaço, assim sua dimensão está entre 1 e 3.

“Mas como calcular a dimensão dos fractais?”

Para responder a contento esta questão, começamos por considerar alguns casos específicos.

Consideremos um segmento de reta de comprimento  $l$ , um quadrado e um cubo de arestas  $l$ .

Dividamos o segmento em quatro, obtemos  $4^1$  partes iguais, cada uma com comprimento:  $\left(\frac{1}{4}\right)^1 l$ .



Figura 1: Segmento  $l$ .

No quadrado, ao dividirmos cada lado em quatro partes iguais, obtemos  $4^2$  partes iguais, cada uma com área:  $\left(\frac{1}{4}\right)^2 l^2$ .

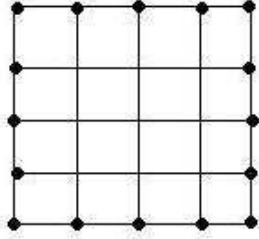


Figura 2: Quadrado de aresta  $l$ .

No cubo, procedendo de maneira análoga, obtemos  $4^3$  partes iguais, cada parte com volume igual a:  $\left(\frac{1}{4}\right)^3 l^3$ .

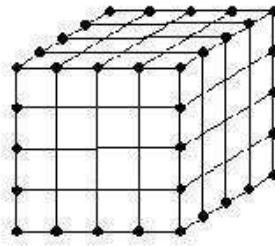


Figura 3: Cubo de aresta  $l$ .

Assim, definindo  $s$  : o fator de contração, razão pela qual dividimos o segmento;  $k$  : a quantidade de subconjuntos originados pelo fator de contração;  $d$  : a dimensão; para o segmento de dimensão 1 temos:  $s = 1/4$  e  $k = 1/s^1 = 4$ , para o quadrado de dimensão 2:  $s = 1/4$  e  $k = 1/s^2 = 4^2$  e, no cubo de dimensão 3:  $s = 1/4$  e  $k = 1/s^3 = 4^3$ .

Generalizando para qualquer dimensão  $d$  temos:  $k = 1/s^d$ , isto é,  $k = (1/s)^d$ . Então, aplicando o logaritmo em ambos os lados temos que  $d = \frac{\lg k}{\lg(1/s)}$  e, portanto,  $d = \frac{\ln k}{\ln 1/s}$ , ou seja, a **dimensão de Hausdorff** para conjuntos auto-similares é dada por:

$$d = \frac{\ln k}{\ln 1/s}.$$

A dimensão de Hausdorff e a dimensão topológica de um conjunto, não precisam necessariamente ser a mesma, mas eventualmente podem coincidir como no caso da Curva de Peano. Além disso, a dimensão de Hausdorff é sempre maior ou igual a dimensão topológica.

Outras características peculiares dos fractais são a auto-similaridade e a complexidade infinita. A primeira refere-se as semelhanças das partes com o todo, enquanto que a segunda está relacionada com as suas interações infinitas, ou seja, um fractal pode ser expresso através de um processo cíclico, sendo constituído de um número infinito de interações, onde a cada interação, sua estrutura torna-se mais complexa.

Dizemos que um subconjunto fechado e limitado do plano euclidiano  $\mathbb{R}^2$  é **auto-similar** se pode ser descrito da forma  $S = S_1 \cup S_2 \cup S_3 \cup \dots \cup S_k$  onde  $S_1, S_2, S_3, \dots, S_k$  são conjuntos não-sobrepostos, cada um dos quais é congruente à contração de  $S$  pelo mesmo fator  $s$ , com  $0 < s < 1$ .

Dessa forma, cada subconjunto é uma pequena cópia do conjunto todo. Um fato interessante (curioso) é que os fractais que são “exatamente” auto-similares não representam, na maioria das vezes, fenômenos naturais, como no caso do triângulo de Sierpinski onde cada triângulo interno a ele é uma cópia em menor escala dele mesmo. Já uma couve-flor, que representa um fractal natural, possui pequenos “ramos” que são semelhantes ao fractal, mas não cópias exatas do mesmo.

Veremos, a seguir, a construção de alguns fractais com o uso do software iGeom, e suas respectivas dimensões.

O software iGeom <sup>2</sup> é um programa livre, desenvolvido por pesquisadores da USP, que trabalha com geometria dinâmica, ou seja, é a implementação computacional

---

<sup>2</sup>Disponível na página: <http://www.matematica.br/igeom>.



da régua e do compasso, permitindo que o aluno experimente, que procure ele próprio elaborar e testar conjecturas (conforme King e Shattschneider, apud Brandão, Isotani e Moura, 2004). Este software possui as ferramentas básicas para as construções geométricas favorecendo o estudo da geometria plana além de possibilitar o estudo de outras geometrias: como a analítica e a fractal. Esta última geometria pode ser explorada devido o recurso de scripts recorrentes (roteiros que realizam determinadas construções, isto é, funções matemáticas tendo como domínio e como imagem objetos geométricos). O uso de scripts permite o estudo de contagem e convergência associados aos fractais (na forma de regras de formação) e também uma introdução ao conceito de algoritmo e programação, permitindo ao aluno ater-se nos pontos de recorrência e não em detalhes de implementação, aproximando-se de uma programação funcional.

## Triângulo de Sierpinski

Construção utilizando o software iGeom:

1. Marca-se o ponto  $A$  e o  $B$  e seleciona-se a opção “iniciar script”;
2. Traça-se o segmento  $AB$  e a partir dele, o triângulo equilátero  $ABC$ ;
3. Marca-se os pontos médios  $D, E$  e  $F$  respectivos dos segmentos  $AB, BC$  e  $CA$ ;
4. Traça-se os segmentos  $DE, EF$  e  $FD$ ;
5. A partir da opção “esconder”, esconde-se os rótulos de todos os elementos exceto os rótulos dos pontos;
6. Seleciona-se os pontos  $D, E, F$  e então aplica-se a opção “área poligonal” e da mesma forma procede-se para os pontos  $A, D, F$ ;  $D, B, E$  e  $F, E, C$ ;
7. Através da opção “área poligonal”, muda-se a cor da área poligonal  $DEF$  de cinza para branco;

8. Esconde-se os rótulos das áreas poligonais;
9. Aplica-se recorrência nos pontos na ordem seguinte: selecione  $A$  e  $D$  e clique em “recorrência”, e de maneira idêntica proceda com os pontos:  $D$  e  $B$ ,  $B$  e  $E$ ,  $E$  e  $D$ ,  $D$  e  $F$ ,  $F$  e  $A$ ,  $F$  e  $E$ ,  $E$  e  $C$  e por fim  $C$  e  $F$ ;
10. Esconde-se todos os rótulos do pontos;
11. Na opção “gravar” salva-se a construção com extensão “.scr”;
12. Limpa-sea tela e seleciona-se dois pontos quaisquer;
13. Através da opção “aplicar script”, aplica-se o script;
14. Abre-se uma janela onde se digita a profundidade desejada obtendo-se, assim, o fractal **Triângulo de Sierpinski**.

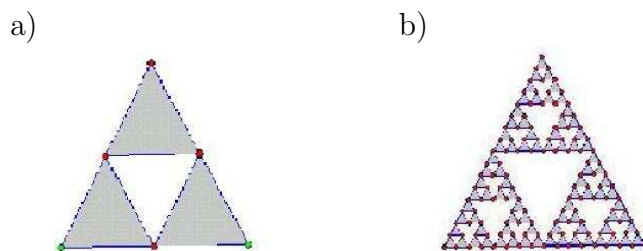


Figura 4: a) Gerador do Triângulo. b) Iteração de nível 3.

Dividimos cada lado do triângulo em duas partes quando marcamos o ponto médio, então a contração é dada por  $s = 1/2$ . Desde o gerador, cada triângulo gera mais três novos triângulos dentro de si, então podemos perceber que o número de subconjuntos originados pelo fator de contração é 3. Assim,  $d = \frac{\ln 3}{\ln 2}$  ou seja, a dimensão de Hausdorff do Triângulo de Sierpinski é aproximadamente 1,58.



4. Define-se duas circunferências, uma centrada em  $F$  e de raio  $FG$  e outra centrada em  $G$  e de raio  $GF$ , na interseção delas marca-se o ponto  $H$ ;
5. Muda-se a cor do segmento  $AB$  para branco e traçam-se os segmentos:  $AF$ ,  $FH$ ,  $HG$ ,  $GB$ ;
6. Esconde-se as circunferências;
7. Seleciona-se o ponto  $A$  e  $F$  e aplica-se recorrência procedendo de maneira análoga com os demais pontos sempre respeitando a ordem citada em 5;
8. Através da opção “gravar script”, salva-se a construção com extensão “.scr”;
9. Limpa-se a tela e a partir de dois pontos quaisquer se aplica a construção com profundidade desejada obtendo-se assim o fractal **Curva de Koch**.

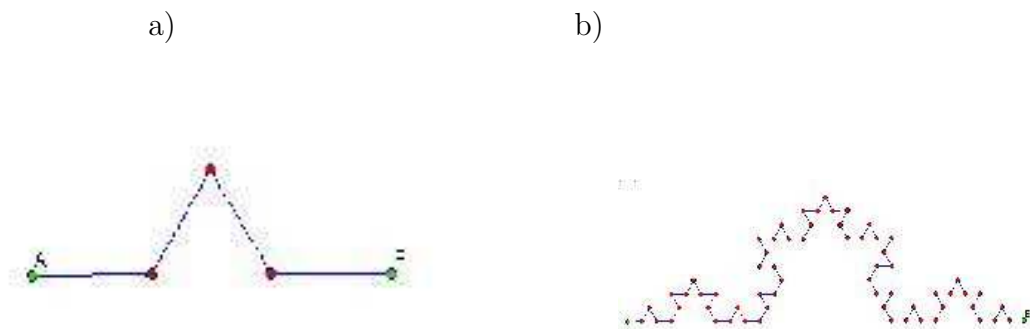


Figura 6: a) Gerador da curva de Koch. b) Interação de nível 2.

A curva tem comprimento infinito e é auto-similar, pois cada parte é igual a totalidade da figura. Como inicialmente dividimos o segmento em 3 partes, temos que  $s = 1/3$  e esse fator, gerou 4 novos segmentos de reta, e portanto  $k = 4$ . Então, a dimensão da curva é dada por:  $d = \frac{\ln(4)}{\ln 3} = 1,26$ .

## Floco de Neve

A curva de Koch quando aplicada sobre os lados de um triângulo equilátero dá origem ao fractal chamado floco de neve.

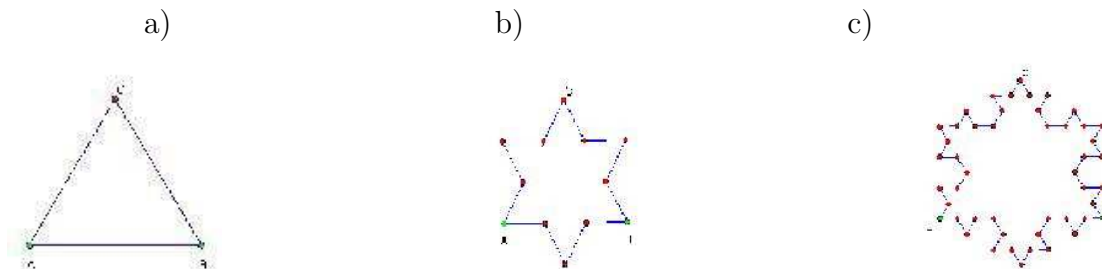


Figura 7: a) Gerador do Floco de Neve. b) Iteração de nível 1. c) Iteração de nível 2.

Construção com o uso do iGeom:

1. Em um triângulo equilátero, seleciona-se dois pontos e aplica-se o fractal “Curva de Koch” construído anteriormente, procedendo-se assim em cada par de pontos do triângulo obtendo-se dessa forma uma estrela de seis pontas;
2. Seleciona-se dois pontos de um lado da estrela e aplica-se o fractal, procedendo assim para todos os lados e quantas vezes for preciso.

Dividimos inicialmente cada lado do triângulo em três segmentos,  $s = 1/3$ , gerando assim novos quatro segmentos,  $k = 4$ . Desprezando o interior da figura e consideremos apenas o contorno, então a dimensão será:  $d = \frac{\ln 4}{\ln 3} = 1,26$ .

Procuramos, dessa forma, construir e estudar alguns fractais clássicos, que possuem relevância histórica, com o auxílio do software iGeom, para desenvolver a base para futuras inserções na teoria de fractais. Acreditamos que uma análise e construção destes objetos geométricos nos possibilitará uma inserção na elaboração de diversos novos fractais, possibilitando, dessa forma, uma visualização geométrica real do nosso mundo

natural.

**Referências Bibliográficas:**

BARBOSA, R. M., **Descobrimdo a geometria fractal para sala de aula**, Autêntica, Belo Horizonte, 2<sup>a</sup> ed, (2005).

BRANDÃO, L. O., ISOTANI, S.,MOURA, J.G.: **in Anais do 15° Simpósio de informática na Educação**, Manaus, (2004).

LAGO, H., CASTRO, A., BATANETE, A.: **Natureza - caos ou ordem?** Trabalho de fundamentos e ensino da álgebra da Universidade de Coimbra - FCTUC, Faculdade de Ciências e Tecnologia , Departamento de Matemática, Fundamento e Ensino da Álgebra, (2005).

MANDELBROT, B. B., **The Fractal Geometry of Nature**, W. H. Freeman and Company, New York, (2004).

# EXISTÊNCIA DA SIGMA-ÁLGEBRA DE LEBESGUE

Cleber Fernando Colle, Fernando Mucio Bando.

Universidade Estadual do Oeste do Paraná - UNIOESTE, clebecolle@msn.com; Universidade Estadual  
do Oeste do Paraná - UNIOESTE, bando@unioeste.br

## Resumo:

No desenvolvimento da Integral de Lebesgue, que é uma generalização da integral de Riemann, se faz necessário estender a medida em um semi-anel  $\mathcal{S}$ ,  $\mu : \mathcal{S} \rightarrow [0, +\infty)$ , para uma  $\sigma$ -álgebra que satisfaça as condições de Lebesgue. Nosso objetivo é mostrar a existência dessa  $\sigma$ -álgebra definida por  $\Lambda = \{A \subseteq \Omega \mid A \cap S \in \Lambda_0 \forall S \in \mathcal{S}\}$ , cuja a extensão  $\nu : \Lambda \rightarrow [0, +\infty]$  de modo natural deve ser  $\sigma$ -aditiva, e para todo  $S \in \mathcal{S}$ ,  $\nu(S) = \mu(S)$ .

**Palavras-chave:** Análise, Integral de Lebesgue,  $\sigma$ -álgebra de Lebesgue, Medida.

## Apresentação:

Dada uma função  $f : \Omega \rightarrow [-\infty, +\infty]$  e dado  $t \in [-\infty, +\infty]$  denotaremos  $[f > t] = \{x \in \Omega \mid f(x) > t\}$ ,  $[f \geq t] = \{x \in \Omega \mid f(x) \geq t\}$ , etc.

A função indicadora de um conjunto  $A \subseteq \Omega$  é  $1_A : \Omega \rightarrow \mathbb{R}$ , definida por  $1_A(x) = 1$  se  $x \in A$  e  $1_A(x) = 0$  se  $x \notin A$ . Seja  $\mathcal{S}$  um semi-anel de subconjuntos de  $\Omega$ . As funções  $\mathcal{S}$ -simples são as funções do tipo

$$h = \sum_{j=1}^m r_j 1_{S_j} : \Omega \rightarrow \mathbb{R},$$

onde os  $S_j$  são conjuntos disjuntos,  $S_j \in \mathcal{S}$ ,  $r_j \in \mathbb{R}$  e  $m$  é finito. Evidentemente  $h(\Omega) = \{h(x) \mid x \in \Omega\}$  é finito, pois temos  $h(x) = r_j$  se  $x \in S_j$  onde  $1 \leq j \leq m$ ,  $h(x) = 0$  se  $x \notin \bigcup_{j=1}^m S_j$ .

A integral de  $h$  relativamente a uma medida  $\mu : \mathcal{S} \rightarrow [0, +\infty)$  é definida por

$$I(h) = \sum_{j=1}^m r_j \mu(S_j).$$

**Proposição 1** *O valor de  $I(h)$  está bem definido, isto é, se*

$$h_1 = \sum_{j=1}^m r_j 1_{S_j} = h_2 = \sum_{i=1}^n c_i 1_{T_i},$$

então  $I(h_1) = I(h_2)$ .

**Demonstração.** Inicialmente, para cada  $r_j \neq 0$  em  $h(\Omega)$ , temos  $[h = r_j] = S_j$ , assim

$$I(h) = \sum_{\substack{r \in h(\Omega) \\ r \neq 0}} r \mu([h = r]).$$

Se  $\sum_{j=1}^m S_j = \emptyset$  e  $\sum_{i=1}^n T_i = \emptyset$ , então trivialmente  $I(h_1) = I(h_2)$ . Suponha  $\sum_{j=1}^m S_j \neq \emptyset$  e  $\sum_{i=1}^n T_i \neq \emptyset$ . Da hipótese de  $h_1 = h_2$  para cada  $r_j \neq 0$  em  $h_1(\Omega)$  existe  $c_i \neq 0$  em  $h_2(\Omega)$  tal que  $r_j = c_i$  para algum  $j \in 1, \dots, m$  e  $i \in 1, \dots, n$ . Caso  $x \notin \sum_{j=1}^m S_j$  ou  $x \notin \sum_{i=1}^n T_i$  então  $h_1(x) = h_2(x) = 0$ . Se  $r_j = c_i \neq 0$  então  $[h_1 = r_j] = [h_2 = c_i]$ , logo  $\mu([h_1 = r_j]) = \mu([h_2 = c_i])$ , assim  $r_j \mu([h_1 = r_j]) = c_i \mu([h_2 = c_i])$ . Portanto

$$\sum_{\substack{r \in h_1(\Omega) \\ r \neq 0}} r \mu([h_1 = r]) = \sum_{\substack{c \in h_2(\Omega) \\ c \neq 0}} c \mu([h_2 = c]). \blacksquare$$

Sendo  $V$  um espaço vetorial qualquer, uma função  $T : V \rightarrow \mathbb{R}$  é chamado um operador linear quando para todos  $f$  e  $g \in V$  valem

$$T(f + g) = T(f) + T(g) \text{ e } T(rf) = rT(f) \forall r \in \mathbb{R}.$$

Seja  $\mathcal{H} = \{h : \Omega \rightarrow \mathbb{R} \mid h \text{ é } \mathcal{S} - \text{simples}\}$  um espaço vetorial. Dada uma medida qualquer  $\mu : \mathcal{S} \rightarrow [0, +\infty)$ , então  $I : \mathcal{H} \rightarrow \mathbb{R}$  é um operador linear, (ISNARD, 2007).

## Reticulados Vetoriais



Dados  $t_1$  e  $t_2 \in [-\infty, +\infty]$  denotaremos por  $t_1 \vee t_2$  e  $t_1 \wedge t_2$  o máximo e mínimo de  $\{t_1, t_2\}$ . Dadas funções  $f$  e  $g : \Omega \rightarrow [-\infty, +\infty]$  denotamos por  $f \vee g$  e  $f \wedge g$  as funções  $x \mapsto f(x) \vee g(x)$  e  $x \mapsto f(x) \wedge g(x)$ .

Definimos também  $f^+ = f \vee 0$  e  $f^- = (-f)^+$ , obtendo assim  $f^+ - f^- = f$  e  $f^+ + f^- = |f|$ , onde  $|f|$  denota a função  $x \mapsto |f(x)|$ .

**Definição 1** *Um reticulado vetorial de funções de  $\Omega \rightarrow \mathbb{R}$  é qualquer espaço vetorial  $V$  de funções tal que se  $f \in V$  então  $|f| \in V$ .*

O conjunto  $\mathcal{H}$  das funções  $\mathcal{S}$ -simples de  $\Omega \rightarrow \mathbb{R}$  trivialmente é um reticulado vetorial, pois se  $h = \sum_{j=1}^n r_j 1_{S_j} \in \mathcal{H}$ , então  $|h| = \sum_{j=1}^n |r_j| 1_{S_j} \in \mathcal{H}$ .

**Proposição 2** *Seja  $V$  qualquer reticulado vetorial de funções de  $\Omega \rightarrow \mathbb{R}$ . Dados  $f$  e  $g \in V$ , então  $f \vee g$  e  $f \wedge g \in V$ .*

**Demonstração.** Conseqüência direta da definição de reticulado vetorial, e do fato de reescrever  $f \vee g$  e  $f \wedge g$  respectivamente como  $\frac{f+g}{2} + \frac{|f-g|}{2}$  e  $\frac{f+g}{2} - \frac{|f-g|}{2}$ . ■

### Propriedades da Integral das Funções Simples

Suponha  $\mu : \mathcal{S} \rightarrow [0, +\infty)$   $\sigma$ -aditiva. Dados  $t_n, t \in [-\infty, +\infty]$ , então  $t_n \nearrow t$  e  $t_n \searrow t$  significam que  $\lim t_n = t$  com  $t_n \leq t_{n+1}$ , respectivamente,  $t_n \geq t_{n+1} \quad \forall n$ .

(i) Se  $h = \sum_{j=1}^n r_j 1_{S_j} \in \mathcal{H}$ , onde  $h \geq 0$  então  $I(h) \geq 0$ . De fato,  $r_j \geq 0$  e  $\mu(S_j) \geq 0$ , logo  $I(h) = \sum_{j=1}^n r_j \mu(S_j) \geq 0$ .

(ii) Se  $h_1 \geq h_2$  estão em  $\mathcal{H}$ , então  $I(h_1) \geq I(h_2)$ . De fato,  $h_1 - h_2 \geq 0$ , logo  $I(h_1 - h_2) \geq 0$ , como  $I(h_1 - h_2) = I(h_1) - I(h_2)$ , então  $I(h_1) - I(h_2) \geq 0$ , ou seja,  $I(h_1) \geq I(h_2)$ .

**Proposição 3** *Se  $h_n \nearrow h$ , com  $h_n$  e  $h \in \mathcal{H}$ , então  $I(h_n) \nearrow I(h)$ .*

**Demonstração.** Consultar (ISNARD, 2007).

**Lema 1** Se  $t_n \nearrow t \in [-\infty, +\infty]$  e  $s_n \nearrow s \in [-\infty, +\infty]$ , então  $t_n \wedge s_n \nearrow t \wedge s$ .

**Demonstração.** De  $t_n \nearrow t$  e  $s_n \nearrow s$ , implica que  $t_n \leq t_{n+1}$  e  $s_n \leq s_{n+1}$ , logo temos  $t_n \wedge s_n \leq t_{n+1} \wedge s_{n+1} \leq t \wedge s$ , assim  $t_n \wedge s_n \nearrow \lim(t_n \wedge s_n) \leq t \wedge s$ . Dado qualquer  $r < t \wedge s$ , então  $r < t$  e  $r < s$ , logo existe algum  $n_0$  tal que para todo  $n > n_0$ ,  $r < t_n$  e  $r < s_n$ , logo  $r < t_n \wedge s_n$ , e assim  $r < \lim(t_n \wedge s_n)$ . Fazendo  $r \rightarrow t \wedge s$ , concluímos que  $t \wedge s \leq \lim(t_n \wedge s_n)$ , donde  $t \wedge s \leq \lim(t_n \wedge s_n) \leq t \wedge s$ . Portanto  $\lim(t_n \wedge s_n) = t \wedge s$ , ou seja,  $t_n \wedge s_n \nearrow t \wedge s$ . ■

**Proposição 4** Sejam  $h$  e  $h_n \in \mathcal{H}$ ,  $h \geq 0$ ,  $h_n \geq 0$ . Se para cada  $x \in \Omega$  vale  $h(x) \leq \sum_{n=1}^{\infty} h_n(x)$ , então  $I(h) \leq \sum_{n=1}^{\infty} I(h_n)$ .

**Demonstração.** Seja  $f_n = (h_1 + \dots + h_n) \wedge h$ , de modo que  $f_n \in \mathcal{H}$ . Afirmamos que  $f_n(x) \nearrow h(x)$ . De fato, pelo Lema anterior, temos

$$(h_1(x) + \dots + h_n(x)) \wedge h(x) \nearrow \lim_{n \rightarrow \infty} \left[ (h_1(x) + \dots + h_n(x)) \wedge h(x) \right],$$

como  $h(x) \leq \sum_{n=1}^{\infty} h_n(x)$ , então

$$\lim_{n \rightarrow \infty} \left[ (h_1(x) + \dots + h_n(x)) \wedge h(x) \right] = \left( \sum_{n=1}^{\infty} h_n(x) \right) \wedge h(x) = h(x),$$

Assim  $f_n(x) \nearrow h(x)$ . Pela Proposição 3, temos que  $I(f_n) \nearrow I(h)$ , mas  $f_n \leq h_1 + \dots + h_n$ , logo  $I(f_n) \leq I(h_1 + \dots + h_n)$ , ou seja,  $I(f_n) \leq I(h_1) + \dots + I(h_n)$ , donde

$$\lim_{n \rightarrow \infty} I(f_n) \leq \lim_{n \rightarrow \infty} [I(h_1) + \dots + I(h_n)]$$

Portanto  $I(h) \leq \sum_{n=1}^{\infty} I(h_n)$ . ■

## A Integral Superior

A integral de Lebesgue é definida primeiro para funções não negativas, em seguida pela decomposição  $f = f^+ - f^-$ , para outras funções.

Dada uma seqüência de funções  $f_n : \Omega \rightarrow [0, +\infty]$  sempre está definida a função  $\sum_{n=1}^{\infty} f_n : x \mapsto \sum_{n=1}^{\infty} f_n(x) \in [0, +\infty]$ . Essa é a vantagem de se trabalhar com funções não negativas no contexto de séries.

**Definição 2** Dada qualquer função  $f : \Omega \rightarrow [0, +\infty]$  definimos a integral superior

$$I^*(f) = \inf \left\{ \sum_{n=1}^{\infty} I(h_n) \mid h_n \in \mathcal{H}, h_n \geq 0, \sum_{n=1}^{\infty} h_n \geq f \right\}. \quad (1)$$

O ínfimo deste conjunto é definido considerando todas as seqüências possíveis de  $h_n \in \mathcal{H}$ ,  $h_n \geq 0$ , tais que  $\sum_{n=1}^{\infty} h_n(x) \geq f(x)$  para todo  $x \in \Omega$ . Caso não exista tais seqüências de  $h_n \in \mathcal{H}$ ,  $h_n \geq 0$ , então, por definição  $I^*(f) = \inf \{\emptyset\} = +\infty$ . Chamamos  $I^*(f)$  de uma integral superior tipo Lebesgue porque é definida através da majoração de  $f$  por funções do tipo  $x \mapsto \sum_{n=1}^{\infty} h_n(x)$ .

**Proposição 5** Vale  $I^*(h) = I(h)$  para todo  $h \in \mathcal{H}$  tal que  $h \geq 0$ .

**Demonstração.** Da definição de  $I^*(h)$  temos que  $\sum_{n=1}^{\infty} h_n \geq h$ , pela Proposição 5,  $I(h) \leq \sum_{n=1}^{\infty} I(h_n)$ , assim

$$I(h) \leq \inf \left\{ \sum_{n=1}^{\infty} I(h_n) \mid h_n \in \mathcal{H}, h_n \geq 0, \sum_{n=1}^{\infty} h_n \geq h \right\},$$

ou seja,  $I(h) \leq I^*(h)$ .

Por outro lado, com  $h_1 = h$  e  $h_n = 0$  para todo  $n \geq 2$  em (1), obtemos que  $h = \sum_{n=1}^{\infty} h_n$ , logo  $I(h) = \sum_{n=1}^{\infty} I(h_n)$ , então

$$I(h) \geq \inf \left\{ \sum_{n=1}^{\infty} I(h_n) \mid h_n \in \mathcal{H}, h_n \geq 0, \sum_{n=1}^{\infty} h_n \geq h \right\},$$

ou seja,  $I(h) \geq I^*(h)$ . Portanto  $I^*(h) = I(h)$ . ■

**Proposição 6** Da definição de  $I^*$ , se verifica facilmente que  $I^*(rf) = rI^*(f)$ , para cada número real  $r \geq 0$  e toda função  $f : \Omega \rightarrow [0, +\infty]$ .

**Proposição 7** Sejam  $f$  e  $f_n : \Omega \rightarrow [0, +\infty]$ . Se  $f(x) \leq \sum_{n=1}^{\infty} f_n(x) \leq +\infty \forall x \in \Omega$ , então  $I^*(f) \leq \sum_{n=1}^{\infty} I^*(f_n)$ . Em particular

$$I^*\left(\sum_{n=1}^{\infty} f_n\right) \leq \sum_{n=1}^{\infty} I^*(f_n),$$

ou seja,  $I^*$  é  $\sigma$ -subaditiva. Segue que

$$I^*(f_1 + f_2) \leq I^*(f_1) + I^*(f_2),$$

isto é,  $I^*$  é sub-aditiva.

**Demonstração.** Se  $\sum_{n=1}^{\infty} I^*(f_n) = +\infty$ , então trivialmente  $I^*(f) \leq \sum_{n=1}^{\infty} I^*(f_n)$ . Suponha  $\sum_{n=1}^{\infty} I^*(f_n) < +\infty$ , então  $\forall \varepsilon > 0, \forall n$ , da definição de  $I^*(f_n)$  segue que existem  $h_{n,j} \in \mathcal{H}, h_{n,j} \geq 0$ , tais que

$$f_n \leq \sum_{j=1}^{\infty} h_{n,j} \text{ e } \sum_{j=1}^{\infty} I(h_{n,j}) < I^*(f_n) + \frac{\varepsilon}{2^n}.$$

Então,  $\forall x \in \Omega$  vale

$$f_n(x) \leq \sum_{j=1}^{\infty} h_{n,j}(x),$$

logo,

$$\sum_{n=1}^{\infty} f_n(x) \leq \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} h_{n,j}(x),$$

ou seja,

$$f(x) \leq \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} h_{n,j}(x) = \sum_{n,j} h_{n,j}.$$

Da Análise real, sabemos que  $\mathbb{N} \times \mathbb{N}$  é enumerável, (LIMA, 1989). Tome  $(n_k, j_k)$  uma enumeração qualquer de  $\mathbb{N} \times \mathbb{N}$ , assim

$$f(x) \leq \sum_{k=1}^{\infty} h_{n_k, j_k} \quad \forall x \in \Omega.$$

Portanto, pela definição de  $I^*$ ,

$$I^*(f) = \inf \left\{ \sum_{k=1}^{\infty} I(h_{n_k, j_k}) \mid h_{n_k, j_k} \in \mathcal{H}, h_{n_k, j_k} \geq 0, \sum_{k=1}^{\infty} h_{n_k, j_k} \geq f \right\},$$

vale

$$I^*(f) \leq \sum_{k=1}^{\infty} I(h_{n_k, j_k}) = \sum_{n=1}^{\infty} \sum_{j=1}^{\infty} I(h_{n,j}) \leq \sum_{n=1}^{\infty} \left( I^*(f_n) + \frac{\varepsilon}{2^n} \right) = \sum_{n=1}^{\infty} I^*(f_n) + \sum_{n=1}^{\infty} \frac{\varepsilon}{2^n}.$$

Como  $\sum_{n=1}^{\infty} \frac{\varepsilon}{2^n} \rightarrow \varepsilon$ , daí

$$I^*(f) \leq \sum_{n=1}^{\infty} I^*(f_n) + \varepsilon, \quad \forall \varepsilon > 0$$

Fazendo  $\varepsilon \rightarrow 0$ , concluímos que

$$I^*(f) \leq \sum_{n=1}^{\infty} I^*(f_n). \quad \blacksquare$$

**Corolário 1** *Se  $0 \leq f \leq f_1$  então  $I^*(f) \leq I^*(f_1)$ .*

## Convergência e Operadores Lineares Limitados

Seja  $V$  um espaço vetorial normado ou semi-normado, (ISNARD, 2007). Dizemos que uma seqüência  $(v_n)$  converge para  $v \in V$  quando  $\lim \|v_n - v\| = 0$ . Para esse fato, usamos as seguintes notações:  $\lim v_n = v$  ou  $v_n \rightarrow v$ .

**Proposição 8** *Seja  $V$  um espaço vetorial normado ou semi-normado, se  $v_n \rightarrow v$  e  $w_n \rightarrow w$  em  $V$ , então se verifica facilmente que  $v_n \pm w_n \rightarrow v \pm w$ ,  $rv_n \rightarrow rv \quad \forall r \in \mathbb{R}$  e  $\|v_n\| \rightarrow \|v\|$ .*

**Definição 3** *Dado  $W \subseteq V$ , o fecho de  $W$  é*

$$\overline{W} = \{v \in V \mid \text{existe uma sequencia } (w_n) \in W \text{ tal que } w_n \rightarrow v\}.$$

Em conseqüência, dado  $W \subseteq V$  então  $\overline{W}$  é fechado em  $V$ , isto é, o fecho de  $\overline{W}$  é  $\overline{W}$ . Isso significa que se  $v_n \in \overline{W}$  e  $v_n \rightarrow v$  em  $V$  então  $v \in \overline{W}$ .

Sejam  $V$  e  $U$  espaços vetoriais normados ou semi-normados. Um operador linear  $T : V \rightarrow U$  é limitado quando existe  $c \in \mathbb{R}$  tal que  $\|T(v)\| \leq c\|v\| \quad \forall v \in V$ . Nesse caso dizemos que  $T : V \rightarrow U$  é contínuo, ou seja, se  $v_n \rightarrow v$  em  $V$  então  $T(v_n) \rightarrow T(v)$  em  $U$ .

**Teorema 1 (Teorema de Extensão)** *Seja  $V$  um espaço vetorial normado ou seminormado. Seja  $W$  um subespaço vetorial de  $V$ . Então:*

(i)  $\overline{W}$  é um subespaço vetorial de  $V$ .

(ii) *Todo operador linear limitado  $T : W \rightarrow \mathbb{R}$  tem uma única extensão a um operador linear  $\overline{T} : \overline{W} \rightarrow \mathbb{R}$  definido por  $\overline{T}(v) = \lim T(w_n)$  se  $w_n \rightarrow v$ ,  $w_n \in W$ .*

**Demonstração.** Temos  $T : W \rightarrow \mathbb{R}$  linear tal que  $|T(w)| \leq c\|w\| \forall w \in W$ . Seja  $w_n \rightarrow v \in \overline{W}$ ,  $w_n \in W$ , do fato de  $T$  ser limitado, a seqüência  $(T(w_n))$  converge em  $\mathbb{R}$ , donde  $(T(w_n))$  é de Cauchy. Definimos  $\overline{T}(v) = \lim T(w_n)$ . Afirmamos que  $\overline{T}(v)$  está bem definida, ou seja, se  $w_n$  e  $u_n \in W$  tais que  $w_n \rightarrow v$  e  $u_n \rightarrow v$ , então  $\lim T(w_n) = \lim T(u_n)$ . De fato,  $w_n - u_n \rightarrow 0 \in W$  pela Proposição 8, e da hipótese de  $T$  ser limitado temos  $T(w_n - u_n) \rightarrow T(0) = 0$ , ou seja,  $\lim T(w_n - u_n) = \lim(T(w_n) - T(u_n)) = 0$ , daí  $\lim T(w_n) = \lim T(u_n)$ .

Para  $w \in W$  vale  $\overline{T}(w) = T(w)$ , pois se  $w_n \rightarrow w$ ;  $w_n \in W$  e do fato de  $T$  ser contínuo em  $W$ , temos  $T(w_n) \rightarrow T(w)$ , ou seja,  $\lim T(w_n) = T(w)$ , daí  $\overline{T}(w) = T(w)$ .

Vamos mostrar que  $\overline{W}$  é subespaço vetorial de  $V$  e que o operador  $\overline{T} : \overline{W} \rightarrow \mathbb{R}$  é um operador linear limitado. Sejam  $w_n \rightarrow v \in \overline{W}$  e  $w'_n \rightarrow v' \in \overline{W}$ , com  $w_n$  e  $w'_n \in W$ . Da Proposição 8,  $w_n + w'_n \rightarrow v + v'$  e  $rw_n \rightarrow rv$ , ou seja,  $v + v'$  e  $rv \in \overline{W}$ , então  $\overline{W}$  é um subespaço vetorial de  $V$ . Agora,

$$\overline{T}(v+v') = \lim T(w_n+w'_n) = \lim(T(w_n)+T(w'_n)) = \lim T(w_n)+\lim T(w'_n) = \overline{T}(v)+\overline{T}(v'),$$

bem como

$$\overline{T}(rv) = \lim T(rw_n) = \lim rT(w_n) = r \lim T(w_n) = r\overline{T}(v).$$

Então  $\overline{T}$  é um operador linear. Finalmente, da hipótese  $|T(w_n)| \leq c\|w_n\|$ , temos pela Desigualdade Triangular, (LIMA, 1989), que  $|T(w_n)| \leq c\|w_n - v\| + c\|v\|$ , daí

$$\lim |T(w_n)| \leq \lim(c\|w_n - v\| + c\|v\|) = \lim c\|w_n - v\| + \lim c\|v\|,$$

logo,  $\lim |T(v)| \leq c\|v\|$  pois  $w_n \rightarrow v$ . Portanto  $|\overline{T}(v)| \leq c\|v\|$ . ■

## $\sigma$ -Álgebras

Seja  $\Omega$  um conjunto. Uma  $\sigma$ -álgebra em  $\Omega$  é uma coleção  $\mathcal{A}$  de subconjuntos de  $\Omega$  tal que  $\emptyset \in \mathcal{A}$  e dados  $A_n \in \mathcal{A}$  então  $\bigcup_{n=1}^{\infty} A_n \in \mathcal{A}$  e  $A_n^c \in \mathcal{A}$ . Em consequência, também  $\bigcap_{n=1}^{\infty} A_n \in \mathcal{A}$ , pois  $\bigcap_{n=1}^{\infty} A_n = (\bigcup_{n=1}^{\infty} A_n^c)^c$ . Vale ainda que  $\Omega = \emptyset^c \in \mathcal{A}$ .

Um espaço de medida é  $(\Omega, \mathcal{A}, \nu)$ , onde  $\Omega$  é um conjunto,  $\mathcal{A}$  é uma  $\sigma$ -álgebra em  $\Omega$  e  $\nu : \mathcal{A} \rightarrow [0, +\infty]$  é uma medida  $\sigma$ -aditiva.

A  $\sigma$ -álgebra  $\mathcal{A}$  é completa relativamente à medida  $\nu : \mathcal{A} \rightarrow [0, +\infty]$  quando dado qualquer  $A \in \mathcal{A}$  tal que  $\nu(A) = 0$ , então para todo  $E \subseteq A$  temos que  $E \in \mathcal{A}$ . Nesse caso diz-se que  $(\Omega, \mathcal{A}, \nu)$  é um espaço de medida completo.

## A $\sigma$ -Álgebra de Lebesgue

No que se segue  $\mu : \mathcal{S} \rightarrow [0, +\infty)$  é uma medida  $\sigma$ -aditiva e  $I^*$  é a integral superior correspondente.

**Teorema 2 (Teorema da Existência)** *Dada qualquer medida  $\sigma$ -aditiva*

*$\mu : \mathcal{S} \rightarrow [0, +\infty)$  existe uma  $\sigma$ -álgebra  $\Lambda \supseteq \mathcal{S}$  que satisfaz as seguintes propriedades:*

(L<sub>1</sub>) *Se  $f$  é uma função de tipo  $f = \sum_{n=1}^{\infty} r_n 1_{A_n} : \Omega \rightarrow [0, +\infty]$ , com  $A_n \in \Lambda$  e  $r_n \geq 0$  real, então  $I^*(f) = \sum_{n=1}^{\infty} r_n I^*(1_{A_n})$ .*

(L<sub>2</sub>) *Definindo  $\nu(A) = I^*(1_A)$  para todo  $A \in \Lambda$  então  $\nu : \Lambda \rightarrow [0, +\infty]$  é uma medida  $\sigma$ -aditiva tal que o espaço de medida  $(\Omega, \Lambda, \nu)$  é completo.*

(L<sub>3</sub>) *Se  $A \subseteq \Omega$  satisfaz  $A \cap S \in \Lambda$  para todo  $S \in \mathcal{S}$ , então  $A \in \Lambda$ .*

**Definição 4** *Chamamos uma  $\sigma$ -álgebra  $\Lambda \supseteq \mathcal{S}$  com as propriedades acima de  $\sigma$ -álgebra de Lebesgue para a medida  $\mu : \mathcal{S} \rightarrow [0, +\infty)$ .*

Provaremos agora a existência de  $\Lambda$ . A extensão de Lebesgue de  $\mu : \mathcal{S} \rightarrow [0, +\infty)$  é a medida  $\nu : \Lambda \rightarrow [0, +\infty]$  definida em (L<sub>2</sub>) por  $\nu(A) = I^*(1_A)$ , pois para todo  $S \in \mathcal{S}$   $\nu(S) = I^*(1_S) = I(1_S) = 1 \cdot \mu(S) = \mu(S)$ .

**Lema 2 (Stone, Bourbaki)** *Seja*

$$\mathfrak{F}^1 = \{f : \Omega \rightarrow \mathbb{R} \mid I^*(|f|) < +\infty\},$$

onde  $|f| : \Omega \rightarrow \mathbb{R}$  é a função  $x \mapsto |f(x)|$ .

Então  $\mathfrak{F}^1$  é uma espaço vetorial, e definindo  $\|f\| = I^*(|f|)$  temos uma semi-norma em  $\mathfrak{F}^1$ .

**Demonstração.** Sejam  $f$  e  $g \in \mathfrak{F}^1$ . Tomando as integrais superiores das funções  $x \mapsto |f(x) + g(x)|$  e  $x \mapsto |f(x)| + |g(x)|$ , do fato de  $|f(x) + g(x)| \leq |f(x)| + |g(x)|$ , pelo Corolário 1, e pela subaditividade de  $I^*$  na Proposição 7, obtemos

$$I^*(|f + g|) \leq I^*(|f| + |g|) \leq I^*(|f|) + I^*(|g|),$$

ou seja,  $I^*(|f + g|) < +\infty$ , logo  $f + g \in \mathfrak{F}^1$  e  $\|f + g\| \leq \|f\| + \|g\|$ . Agora, para cada  $r \in \mathbb{R}$   $|rf(x)| = |r||f(x)|$ , logo  $I^*(|rf|) = |r|I^*(|f|)$  pela Proposição 6, portanto  $rf \in \mathfrak{F}^1$  e  $\|rf\| = |r|\|f\|$ . Finalmente,  $0 \in \mathfrak{F}^1$  trivialmente, e da construção de  $I^*$ , temos  $\|f\| = I^*(|f|) \geq 0$ . ■

**Lema 3** *Considere  $\mathcal{H} = \{h : \Omega \rightarrow \mathbb{R} \mid h \text{ é } \mathcal{S}\text{-simples}\}$ . Da definição de função  $\mathcal{S}$ -simples, vale  $\mathcal{H} \subseteq \mathfrak{F}^1$ . Para todo  $h \in \mathcal{H}$   $|I(h)| \leq I(|h|) = \|h\|$ . Portanto  $I : \mathcal{H} \rightarrow \mathbb{R}$  é um operador linear limitado.*

**Demonstração.** Seja  $h \in \mathcal{H}$ . Então  $|h| \in \mathcal{H}$  e  $I^*(|h|) = I(|h|)$  pela Proposição 5. Da definição de função  $\mathcal{S}$ -simples,  $|h| \in \mathcal{H}$ , logo  $h \in \mathfrak{F}^1$  e  $\|h\| = I^*(|h|) = I(|h|)$ . De  $h = h^+ - h^-$ , onde  $h^+$  e  $h^- \geq 0$  temos  $I(h) = I(h^+ - h^-) = I(h^+) - I(h^-)$ , assim

$$|I(h)| = |I(h^+) - I(h^-)| \leq |I(h^+) + I(h^-)| = I(h^+ + h^-) = I(|h|) = \|h\|. \blacksquare$$

**Definição 5** *Seja*

$$\overline{\mathcal{H}} = \{f \in \mathfrak{F}^1 \mid \exists (h_n) \in \mathcal{H} \text{ tal que } \|h_n - f\| \rightarrow 0\}.$$



**Lema 4 (Stone)**  $\overline{\mathcal{H}}$  é um reticulado vetorial de funções. Existe um operador linear limitado  $\overline{I} : \overline{\mathcal{H}} \rightarrow \mathbb{R}$  estendendo  $I : \mathcal{H} \rightarrow \mathbb{R}$  e tal que se  $f \geq 0$  e  $f \in \overline{\mathcal{H}}$ , então  $\overline{I}(f) = I^*(f)$ .

**Demonstração.** Pelo Teorema 1, e o Lema 3, sabemos imediatamente que  $\overline{\mathcal{H}}$  é um espaço vetorial e que existe um operador linear limitado  $\overline{I} : \overline{\mathcal{H}} \rightarrow \mathbb{R}$  estendendo  $I : \mathcal{H} \rightarrow \mathbb{R}$ , definido por

$$\overline{I}(f) = \lim I(h_n) \text{ se } \|h_n - f\| \rightarrow 0, h_n \in \mathcal{H}, f \in \overline{\mathcal{H}} \quad (2)$$

Fixando  $n$ , considere as funções

$$x \mapsto |h_n(x) - f(x)| \text{ e } x \mapsto ||h_n(x)| - |f(x)||,$$

como  $||h_n(x)| - |f(x)|| \leq |h_n(x) - f(x)|$ , logo pelo Corolário 1,

$$I^*(||h_n| - |f||) \leq I^*(|h_n - f|),$$

pelo Lema 2,  $|||h_n| - |f|| \leq \|h_n - f\|$ . De  $\|h_n - f\| \rightarrow 0$ , temos que  $|||h_n| - |f|| \rightarrow 0$  pela Proposição 8, assim de (2) e do fato de  $|h_n| \in \mathcal{H}$  temos  $|f| \in \overline{\mathcal{H}}$  e  $\overline{I}(|f|) = \lim I(|h_n|)$ . Portanto  $\overline{\mathcal{H}}$  é um reticulado vetorial, e  $\overline{I}(|f|) = \lim I(|h_n|) = \lim \|h_n\|$ , daí  $\overline{I}(|f|) = \|f\|$ , pois  $|||h_n| - |f|| \rightarrow 0$ . Em particular, se  $f \geq 0$  e  $f \in \overline{\mathcal{H}}$ , então

$$\overline{I}(f) = \overline{I}(|f|) = \|f\| = I^*(|f|) = I^*(f). \blacksquare$$

**Lema 5** Dados  $(f_n) \in \overline{\mathcal{H}}$ ,  $f_n \geq 0$ . Seja  $f = \sum_{n=1}^{\infty} I^*(f_n) : \Omega \rightarrow [0, +\infty]$ , então:

(i)  $I^*(f) = \sum_{n=1}^{\infty} I^*(f_n)$ ;

(ii) Se  $I^*(f) < +\infty$  e  $f(x) \in \mathbb{R} \forall x \in \Omega$ , então  $f \in \overline{\mathcal{H}}$ .

**Demonstração.** (i) Com efeito, do Lema anterior,  $\overline{I}(f_n) = I^*(f_n)$ . Da linearidade de  $\overline{I}$ , temos

$$I^*(f_1) + \cdots + I^*(f_n) = I^*(f_1 + \cdots + f_n) \leq I^*(f),$$

pelo Corolário 1, pois  $f_1 + \cdots + f_n \leq \sum_{n=1}^{\infty} f_n = f$ . Fazendo  $n \rightarrow +\infty$ , vale

$$\sum_{n=1}^{\infty} I^*(f_n) \leq I^*(f).$$

Agora, da Proposição 7, concluímos que  $\sum_{n=1}^{\infty} I^*(f_n) = I^*(f)$ .

(ii) Com efeito,

$$0 \leq f - (f_1 + \cdots + f_n) = \sum_{j=1}^{\infty} f_{n+j} : \Omega \rightarrow [0, +\infty],$$

de (i), vale

$$I^*(f - (f_1 + \cdots + f_n)) = \sum_{j=1}^{\infty} I^*(f_{n+j}),$$

do Lema 3, temos

$$\begin{aligned} \|f - (f_1 + \cdots + f_n)\| &= \sum_{j=1}^{\infty} I^*(f_{n+j}) \\ &= \sum_{j=1}^{\infty} I^*(f_j) - \sum_{j=1}^n I^*(f_j), \end{aligned}$$

logo,

$$\begin{aligned} \lim_{n \rightarrow \infty} \|f - (f_1 + \cdots + f_n)\| &= \lim_{n \rightarrow \infty} \left( \sum_{j=1}^{\infty} I^*(f_j) - \sum_{j=1}^n I^*(f_j) \right) \\ &= \lim_{n \rightarrow \infty} \sum_{j=1}^{\infty} I^*(f_j) - \lim_{n \rightarrow \infty} \sum_{j=1}^n I^*(f_j) \\ &= 0, \end{aligned}$$

devido a convergência da série em (i). Do fato de  $\overline{\mathcal{H}}$  ser um espaço vetorial, Lema 4,  $f_1 + \cdots + f_n \in \overline{\mathcal{H}}$ , então  $f \in \overline{\mathcal{H}}$ , pois  $\overline{\mathcal{H}}$  é fechado em  $\mathfrak{F}^1$ , consequência da Definição 3. ■

**Definição 6** *Seja*

$$\Lambda_0 = \{A \subseteq \Omega \mid 1_A \in \overline{\mathcal{H}}\}.$$

Sendo assim  $\mathcal{S} \subseteq \Lambda_0$ , pois  $1_S \in \mathcal{H} \subseteq \overline{\mathcal{H}} \forall S \in \mathcal{S}$ , ou seja,  $\mathcal{S} \subseteq \Lambda_0$ . Para todo  $A \in \Lambda_0$   $I^*(1_A) < +\infty$ , devido  $\overline{\mathcal{H}} \subseteq \mathfrak{F}^1$ . Os conjuntos de  $\Lambda_0$  serão chamados conjuntos mensuráveis com medida finita.

**Lema 6**

(i)  $\Lambda_0$  é uma anel;

(ii) Seja  $A = \bigcup_{n=1}^{\infty} A_n$ , com  $A_n \in \Lambda_0 \forall n$ . Se  $I^*(1_A) < +\infty$ , então  $A \in \Lambda_0$ .

**Demonstração.** (i) Dados  $A$  e  $B \in \Lambda_0$  então  $1_A$  e  $1_B \in \mathcal{H}$ , como  $\overline{\mathcal{H}}$  é um reticulado vetorial, Lema 4, então pela Proposição 2,  $1_A \vee 1_B = 1_{A \cup B}$  e  $1_A \wedge 1_B = 1_{A \cap B} \in \overline{\mathcal{H}}$ , assim da Definição 6,  $A \cup B$  e  $A \cap B \in \Lambda_0$ . De  $A \cap B^c = A - (A \cap B)$ , vale  $1_{A \cap B^c} = 1_A - 1_{A \cap B} \in \overline{\mathcal{H}}$ , pois  $\overline{\mathcal{H}}$  é um espaço vetorial que contém  $1_A$  e  $1_{A \cap B}$ , então  $A \cap B^c \in \Lambda_0$ . Finalmente,  $\emptyset \in \overline{\mathcal{H}}$ , pois  $\overline{\mathcal{H}}$  é um espaço vetorial, onde  $1_{\emptyset} = 0 \in \overline{\mathcal{H}}$ , ou seja,  $\emptyset \in \Lambda_0$ . Portanto  $\Lambda_0$  é um anel.

(ii) Definindo  $B_1 = A_1$  e  $B_n = A_n \cap (A_1 \cup \dots \cup A_{n-1})^c$  para  $n \geq 2$ , então trivialmente  $B_i \cap B_j = \emptyset$  se  $i \neq j$ . Agora, note que

$$\begin{aligned} \sum_{n=1}^m B_n &= A_1 \cup \left( \bigcup_{n=2}^m (A_n \cap A_1^c \cap \dots \cap A_{n-1}^c) \right) = \\ &= (A_1 \cup A_2) \cup \left( \bigcup_{n=3}^m (A_1 \cup A_n) \cap (A_1 \cup A_2^c) \cap \dots \cap (A_1 \cup A_{n-1}^c) \right). \end{aligned}$$

Procedendo dessa maneira, temos

$$\sum_{n=1}^m B_n = \left( \bigcup_{n=1}^{m-1} A_n \right) \cup \left[ \left( \left( \bigcup_{n=1}^{m-2} A_n \right) \cup A_m \right) \cap \left( \left( \bigcup_{n=1}^{m-2} A_n \right) \cup A_{m-1}^c \right) \right] = \bigcup_{n=1}^m A_n.$$

Fazendo  $m \rightarrow \infty$ , vale

$$A = \bigcup_{n=1}^{\infty} A_n = \sum_{n=1}^{\infty} B_n, \text{ com } B_n \in \Lambda_0.$$

Daí  $1_A = \sum_{n=1}^{\infty} 1_{B_n} : \Omega \rightarrow \mathbb{R}$ , com  $1_{B_n} \in \overline{\mathcal{H}}$ , pois  $B_n \in \Lambda_0$ , logo segue do Lema 5, que  $1_A \in \overline{\mathcal{H}}$ , pois  $I^*(1_A) < +\infty$ . Portanto  $A \in \Lambda_0$ . ■

**Definição 7** *Seja*

$$\Lambda = \{A \subseteq \Omega \mid A \cap S \in \Lambda_0 \forall S \in \mathcal{S}\}.$$

Vale  $\Lambda_0 \subseteq \Lambda$ , pois para todo  $S \in \mathcal{S}$  e todo  $A \in \Lambda_0$ ,  $A \cap S \in \Lambda_0$  porque  $\Lambda_0$  é um anel.

**Lema 7**  $\Lambda$  é uma  $\sigma$ -álgebra.

**Demonstração.** Dados  $A_n \in \Lambda$ , seja  $A = \bigcup_{n=1}^{\infty} A_n$ . Para cada  $S \in \mathcal{S}$  temos que  $A \cap S = \bigcup_{n=1}^{\infty} (A_n \cap S)$ , com  $A_n \cap S \in \Lambda_0$ . Para todo  $S \in \mathcal{S}$  vale que  $A \cap S \subseteq S$ , assim  $0 \leq 1_{A \cap S} \leq 1_S : \Omega \rightarrow \mathbb{R}$ , logo pelo Corolário 1,  $I^*(1_{A \cap S}) \leq I^*(1_S)$ , e pela Proposição 5, concluímos que  $I^*(1_{A \cap S}) \leq I(1_S) = \mu(S) < +\infty$ . Portanto  $A \cap S \in \Lambda_0$ , o que prova que  $A \in \Lambda$ . Agora, para todo  $A \in \Lambda$ ,  $A \cap S \in \Lambda_0$ , assim

$$S \cap (S \cap A)^c = S \cap (S^c \cup A^c) = S \cap A^c \in \Lambda_0,$$

pois  $\Lambda_0$  é um anel, logo  $A^c \in \Lambda$ . Finalmente  $\emptyset \in \Lambda$ , pois  $\emptyset \in \Lambda_0$ . ■

**Lema 8**

$$\Lambda_0 = \{A \in \Lambda \mid I^*(1_A) < +\infty\}.$$

**Demonstração.** Seja  $A \in \Lambda$  tal que  $I^*(1_A) < +\infty$ . Então existem  $h_n \in \mathcal{H}$ ,  $h_n \geq 0$ , tais que  $1_A \leq \sum_{n=1}^{\infty} h_n$ . Em conseqüência, se  $x \in A$ , então existe  $n_0 \in \mathbb{N}$  tal que  $x \in [h_{n_0} > 0]$ , ou seja,  $A \subseteq \sum_{n=1}^{\infty} [h_n > 0]$ . Cada conjunto  $[h_n > 0]$  é um conjunto do tipo  $S_1 \cup \dots \cup S_p$ , com  $S_j \in \mathcal{S}$ , portanto  $A \cap [h_n > 0] = ((A \cap S_1) \cup \dots \cup (A \cap S_p)) \in \Lambda_0$ , já que  $A \cap S_j \in \Lambda_0$  e  $\Lambda_0$  é um anel.

Definindo  $A_n = A \cap [h_n > 0]$ , afirmamos que  $A = \bigcup_{n=1}^{\infty} A_n$ . De fato, trivialmente  $\bigcup_{n=1}^{\infty} A_n \subseteq A$ . Agora, se  $x \in A$ , então do fato de  $A \subseteq \sum_{n=1}^{\infty} [h_n > 0]$  existe  $n_0$  tal que  $x \in [h_{n_0} > 0]$ , ou seja  $x \in A \cap [h_{n_0} > 0]$ . Portanto  $A = \bigcup_{n=1}^{\infty} A_n$ , com  $A_n \in \Lambda_0$ . Assim pelo Lema 6,  $A \in \Lambda_0$ , pois  $I^*(1_A) < +\infty$ . ■

Para concluir a Demonstração do Teorema 2, provaremos que  $\Lambda$  satisfaz as condições  $(L_1)$ ,  $(L_2)$  e  $(L_3)$ .

**Demonstração.**  $(L_1)$  Se  $r_n 1_{A_n} \in \overline{\mathcal{H}}$  para todo  $n$ , então pelo Lema 5, vale a igualdade. Agora, se  $r_n 1_{A_n} \notin \overline{\mathcal{H}}$  para algum  $n$ , então por  $\overline{\mathcal{H}}$  ser um reticulado vetorial, Lema 4, podemos supor  $r_n > 0$  e  $1_{A_n} \notin \overline{\mathcal{H}}$ , então  $A_n \notin \Lambda_0$ , logo pelo Lema 8,  $I^*(1_{A_n}) = +\infty$ . Do fato de  $f = \sum_{n=1}^{\infty} r_n 1_{A_n}$  com  $r_n \geq 0$ , temos  $f \geq r_n 1_{A_n} \geq 0$ , daí pelo Corolário 1, e a Proposição 6, vale

$$I^*(f) \geq I^*(r_n 1_{A_n}) = r_n I^*(1_{A_n}) = +\infty,$$

Portanto, vale

$$I^*(f) = +\infty = \sum_{n=1}^{\infty} r_n I^*(1_{A_n}).$$

$(L_2)$  Se  $A = \sum_{n=1}^{\infty} A_n$ ,  $A_n \in \Lambda$ , então

$$1_A = \sum_{n=1}^{\infty} 1_{A_n} : \Omega \rightarrow [0, +\infty],$$

e por  $(L_1)$ , segue que  $I^*(1_A) = \sum_{n=1}^{\infty} I^*(1_{A_n})$ , ou seja,  $\nu$  é  $\sigma$ -aditiva.

Seja  $A \in \Lambda$  tal que  $\nu(A) = 0$  e seja  $E \subseteq A$ . Segue-se do fato de  $E \subseteq A$  que  $0 \leq 1_E \leq 1_A$ , então pelo Corolário 1,  $0 \leq I^*(1_E) \leq I^*(1_A) = 0$ , logo  $I^*(1_E) = 0$ . Considere a seqüência  $(h_n)$  tal que  $h_n = 0 \in \mathcal{H}$  para todo  $n \in \mathbb{N}$ . Daí

$$\lim_{n \rightarrow \infty} \|h_n - 1_E\| = \|1_E\| = I^*(1_E) = 0.$$

Logo  $1_E \in \overline{\mathcal{H}}$ . Segue que  $E \in \Lambda_0$ , assim  $E \in \Lambda$ . Portanto  $(\Omega, \Lambda, \nu)$  é completo.

$(L_3)$  De  $A \cap S \in \Lambda$  segue-se que  $(A \cap S) \cap S \in \Lambda_0$ , ou seja,  $A \cap S \in \Lambda_0$ . Portanto  $A \in \Lambda$ . ■

### Referências Bibliográficas:

- LIMA, E. L.: **Análise Real**, Impa, Rio de Janeiro, (1989).  
 ISNARD, C.: **Introdução à Medida e Integração**, Impa, Rio de Janeiro, (2007).

# FUNÇÕES ELÍPTICAS

Kelly Roberta Mazzutti Lübeck

Universidade Estadual do Oeste do Paraná - UNIOESTE, kellyrobertaml@gmail.com

## Resumo:

Uma *função elíptica* é qualquer função meromorfa definida sobre um toro, que como veremos é o quociente do plano complexo  $\mathbb{C}$  por algum reticulado  $G$ . Dentre todas as funções elípticas a mais importante é a clássica função  $\wp$  de Weierstrass. Neste trabalho apresentaremos a *função  $\wp$  de Weierstrass simétrica*, que foi introduzida por H. Karcher. A relação algébrica entre elas, quando definidas sobre o mesmo toro, é um reescalonamento  $a$  seguido por uma translação  $b$ , onde  $a$  e  $b$  são constantes complexas. Entretanto, as funções  $\wp$  de Weierstrass simétricas são mais fáceis de manipular quando observadas do ponto de vista geométrico. Mostraremos que o comportamento delas pode ser estudado usando apenas argumentos geométricos. Além disso, teremos mais informações do que na teoria padrão da  $\wp$  de Weierstrass clássica.

**Palavras-chave:** Funções Elípticas, Função de Weierstrass, Toros planos.

## Apresentação:

Algumas definições e teoremas básicos serão apresentados a seguir. Para maiores detalhes veja Karcher, H. e Nitsche, J. C. C.

**Definição 1.** Uma *superfície de Riemann* é um par  $(X, \Sigma)$ , onde  $X$  é uma variedade 2-dimensional conexa e  $\Sigma$  é uma estrutura complexa em  $X$ .

O toro plano ou toro, exemplo de superfície de Riemann que será manipulado no decorrer do texto, é o quociente do plano complexo por algum reticulado  $\Gamma$ , sendo que

um reticulado é dado por

$$\Gamma := \{nw_1 + mw_2 \mid n, m \in \mathbb{Z}\}$$

para  $w_1, w_2 \in \mathbb{C}$  linearmente independentes sobre  $\mathbb{R}$ . Considerando  $\pi : \mathbb{C} \rightarrow \mathbb{C}/\Gamma$  a projeção canônica podemos induzir uma estrutura complexa em  $\mathbb{C}/\Gamma$  do seguinte modo: seja  $V \subset \mathbb{C}$  um subconjunto aberto tal que quaisquer dois de seus pontos não são equivalentes módulo  $\Gamma$ . Então  $U := \pi(V)$  é aberto e  $\pi|_V : V \rightarrow U$  é um homeomorfismo. Sua inversa  $\phi : U \rightarrow V$  é uma carta complexa. A coleção de todas estas cartas forma o atlas complexo  $\mathcal{U}$ , que cobre  $\mathbb{C}/\Gamma$  e determina, assim, uma estrutura complexa nesta superfície.

**Teorema 1.** (Fórmula de Riemann-Hurwitz). *Sejam  $X$  e  $Y$  superfícies de Riemann compactas e  $f : X \rightarrow Y$  uma aplicação holomorfa não constante. Se  $b_f$  é a ordem total dos ramos de  $f$ , então*

$$gen(X) = \frac{b_f}{2} + \#(f)(gen(Y) - 1) + 1. \quad (1)$$

Uma fórmula equivalente a (1) é

$$\chi(X) = -b_f + \#(f) \cdot \chi(Y). \quad (2)$$

Consideramos agora  $R$  uma superfície topológica compacta e  $J : R \rightarrow R$  um homeomorfismo, cujo conjunto de pontos fixos é discreto, tal que  $J^m = id$ ,  $m > 1$ , e  $J^{m-k} \neq id$  para  $1 \leq k < m$ . Dado  $Q \in R$ , definimos

- o grupo de isotropia de  $Q$  em  $\langle J \rangle$  como  $\mathcal{F}(Q) = \{A \in \langle J \rangle \mid A(Q) = Q\}$ ,
- $\mu(Q) = |\mathcal{F}(Q)|$ ,
- a órbita de  $Q$  associada a  $\langle J \rangle$  como  $orb(Q) = \{Q, J(Q), \dots, J^{m-1}(Q)\}$ .

Notamos que  $\#(\text{orb}(Q)) \cdot \mu(Q) = \text{ord}(J)$ . Para o recobrimento ramificado  $\zeta : R \rightarrow R/\langle J \rangle$  obtemos

$$\chi(R) = \text{ord}(J) \cdot \chi(R/\langle J \rangle) - \sum_{Q \in R} (\mu(Q) - 1). \quad (3)$$

Se  $Q \in R$  com  $\mu(Q) > 1$ , observamos que  $\text{orb}(Q) = \text{orb}(J(Q)) = \dots = \text{orb}(J^{m-1}(Q))$ . Neste caso existem exatamente  $s$  conjuntos disjuntos, com  $s \in \mathbb{N}^*$ , e cada um possui uma cardinalidade  $m_i$ ,  $i = 1, \dots, s$ . Então, podemos reescrever (3) como

$$\chi(R/\langle J \rangle) = \frac{\chi(R)}{\text{ord}(J)} + \frac{\sum_{i=1}^s (\text{ord}(J) - m_i)}{\text{ord}(J)}. \quad (4)$$

Quando  $J^2 = \text{id}$ , temos  $m_i = 1$  e  $s$  igual ao número de pontos fixos de  $R$ . Dessa forma,

$$\chi(R/\langle J \rangle) = \frac{\chi(R)}{2} + \frac{s}{2}.$$

A equação (3) é conhecida como *fórmula de Euler-Poincaré*.

**Teorema 2.** *Sejam  $X, \tilde{X}, Y$  e  $\tilde{Y}$  variedades diferenciáveis, e  $r : \tilde{Y} \rightarrow Y, s : \tilde{X} \rightarrow X$  recobrimentos  $C^\infty$ . Se  $F : \tilde{X} \rightarrow \tilde{Y}$  é um difeomorfismo  $C^\infty$  que preserva fibras, então existe  $f : X \rightarrow Y$  difeomorfismo de classe  $C^\infty$  tal que  $f \circ s = r \circ F$ .*

*Reciprocamente, se  $\tilde{X}$  e  $\tilde{Y}$  são simplesmente conexos e temos  $f : X \rightarrow Y$  difeomorfismo de classe  $C^\infty$ , então existe  $F : \tilde{X} \rightarrow \tilde{Y}$  difeomorfismo  $C^\infty$  que preserva fibras, com  $f \circ s = r \circ F$ .*

Agora introduzimos resultados importantes que não estão estritamente relacionados com funções elípticas. Estes resultados constituem os argumentos fundamentais para podermos analisar o comportamento das funções mencionadas acima. Primeiramente, precisamos da seguinte definição:

**Definição 2.** Uma *involução* é uma aplicação contínua  $I : S \rightarrow S$  que satisfaz  $I \circ I = \text{id}$  (a identidade em  $S$ ). Quando  $S$  é uma superfície de Riemann compacta, a involução é



chamada *hiperelíptica* se  $gen(S/I) = 0$ .

Da definição é imediato ver que toda involução é uma bijeção.

**Teorema 3.** *Toda involução conforme ou anti-conforme em  $\widehat{\mathbb{C}}$  é dada por uma transformação de Möbius  $M$  ou sua conjugada  $\overline{M}$ .*

Prova: Tome uma involução  $I : \widehat{\mathbb{C}} \rightarrow \widehat{\mathbb{C}}$ . Se é conforme, então é biholomorfa e assim uma transformação de Möbius. Se é anti-conforme, então  $\overline{I}$  é biholomorfa e assim  $I$  é a conjugada de uma transformação de Möbius.  $\square$

Note que a recíproca do Teorema 3 **não** é válida, pois  $z \rightarrow 2z$  não é uma involução.

**Teorema 4.** *Sejam  $S$  e  $R$  superfícies de Riemann,  $I : S \rightarrow S$  uma involução e  $f : S \rightarrow R$  uma função contínua, aberta e sobrejetora. Então, existe uma única involução  $J : R \rightarrow R$  tal que  $J \circ f = f \circ I$  se, e somente se, sempre que  $f(x) = f(y)$  temos  $f \circ I(x) = f \circ I(y)$ .*

Prova: As hipóteses do teorema foram formuladas para garantir a comutatividade do diagrama:

$$\begin{array}{ccc} S & \xrightarrow{I} & S \\ f \downarrow & & \downarrow f \\ R & \xrightarrow{J} & R \end{array}$$

Admitimos que  $J \circ f = f \circ I$ . Se  $f(x) = f(y)$  então

$$f \circ I(x) = J \circ f(x) = J \circ f(y) = f \circ I(y).$$

Suponhamos agora que  $f \circ I(x) = f \circ I(y)$  quando  $f(x) = f(y)$ . Para cada  $z \in R$

definimos  $J(z) := f \circ I(x)$ , para algum  $x \in S$  tal que  $f(x) = z$  (este elemento existe pois  $f$  é sobrejetora). Por hipótese  $J : S \rightarrow S$  está bem definida. Além disso,  $J \circ f = f \circ I$ . A função  $J$  é contínua por causa do seguinte argumento: para qualquer subconjunto aberto  $U \subset R$  temos que  $I^{-1}(f^{-1}(U))$  é aberto em  $S$ . Mas  $I^{-1}(f^{-1}(U)) = f^{-1}(J^{-1}(U))$ , e conseqüentemente  $J^{-1}(U)$  é aberto em  $R$ . Ainda temos

$$J \circ J \circ f = J \circ f \circ I = f \circ I \circ I = f.$$

Assim,  $J$  é uma involução. □

É imediato concluir que dado  $F \subset R$  temos  $J(F) = F$  se, e somente se,  $I(f^{-1}(F)) = f^{-1}(F)$ .

## A função $\wp$ de Weierstrass simétrica

Como qualquer toro  $T$  é o quociente de  $\mathbb{C}$  por algum reticulado  $G \subset \mathbb{C}$ , podemos descrever isto por meio de números complexos não nulos  $w_1$  e  $w_2$ , com  $\frac{w_1}{w_2} \notin \mathbb{R}$ . Assim,  $G := \{(2n+1)w_1 + (2m+1)w_2 : n, m \in \mathbb{Z}\}$  e  $T := \mathbb{C}/G$ .

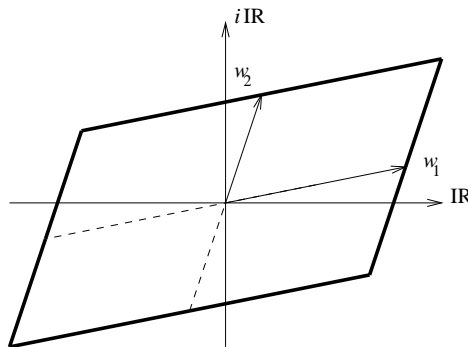


Figura 1: Toro  $T$ .

No toro  $T$  a involução  $I(z) = -z$  tem exatamente quatro pontos fixos:  $0, w_1, w_2$

e  $w_1 + w_2$ . Usando a fórmula de Euler-Poincaré temos:

$$\chi(T/I) = \frac{\chi(T)}{2} + \frac{4}{2} = 2.$$

Logo  $T/I$  tem gênero zero e pelo teorema de Koebe existe um biholomorfismo  $\mathcal{B} : T/I \rightarrow \widehat{\mathbb{C}}$ . A menos de uma transformação de Möbius a função  $\mathcal{B}$  está bem definida.

**Definição 3.** A função  $\wp$  de Weierstrass simétrica é a composta  $\wp := \mathcal{B} \circ (\cdot/I) : T \rightarrow \widehat{\mathbb{C}}$  tal que

$$\wp(0) = 0, \quad \wp(w_1 + w_2) = \infty \quad \text{e} \quad \wp\left(\frac{w_1 + w_2}{2}\right) = i.$$

Devido à função quociente  $\cdot/I : T \rightarrow T/I$ , é imediato concluir que  $\wp(-z) = \wp(z)$ , para qualquer  $z \in T$ . Logo  $\wp\left(-\frac{w_1 + w_2}{2}\right) = i$ . Além disso,  $\deg(\wp) = 2$  e seus pontos de ramo são exatamente os quatro pontos fixos de  $I$ .

Da Definição 2 vemos que  $I$  é hiperelíptica. Vamos agora estudar outra função hiperelíptica em  $T$ , que é dada por  $\mathcal{H}(z) = -z + w_1 + w_2$  (rotação em torno de  $\frac{w_1 + w_2}{2}$ ). Seu conjunto de pontos fixos é exatamente  $\left\{ \pm \frac{w_1 + w_2}{2}, \pm \frac{w_1 - w_2}{2} \right\}$ . Do Teorema 4,  $\mathcal{H}$  e  $\wp$  induzem outra involução em  $\widehat{\mathbb{C}}$ , e como ambas são holomorfas, a involução induzida também será holomorfa. Do Teorema 3 esta involução é uma transformação de Möbius. Como  $\mathcal{H}$  intercambia 0 e  $w_1 + w_2$  e deixa fixo  $\frac{w_1 + w_2}{2}$ , se  $J$  é a induzida por  $\wp$  temos:

$$\begin{aligned} J(0) &= J(\wp(0)) = \wp(\mathcal{H}(0)) = \wp(w_1 + w_2) = \infty; \\ J(\infty) &= J(\wp(w_1 + w_2)) = \wp(\mathcal{H}(w_1 + w_2)) = \wp(0) = 0; \\ J(i) &= J(\wp\left(\frac{w_1 + w_2}{2}\right)) = \wp(\mathcal{H}\left(\frac{w_1 + w_2}{2}\right)) = \wp\left(\frac{w_1 + w_2}{2}\right) = i. \end{aligned}$$

Assim, a transformação de Möbius associada é

$$\wp \rightarrow -\frac{1}{\wp}. \tag{5}$$

Uma vez que os pontos  $\pm \frac{w_1 - w_2}{2}$  permanecem fixados pela  $\mathcal{H}$ , neles temos  $\wp^2 = -1$ . Como o grau da  $\wp$  é 2 e  $\mathcal{H}(\pm \frac{w_1 + w_2}{2}) \neq \mathcal{H}(\pm \frac{w_1 - w_2}{2})$ , temos

$$\wp\left(\pm \frac{w_1 - w_2}{2}\right) = -i.$$

Além disso, (5) implica que todo segmento em  $T$  com extremos  $p, q$  e cujo centro é um ponto fixo de  $\mathcal{H}$ , satisfaz  $\wp(q) = -1/\wp(p)$ . Veja a figura abaixo.

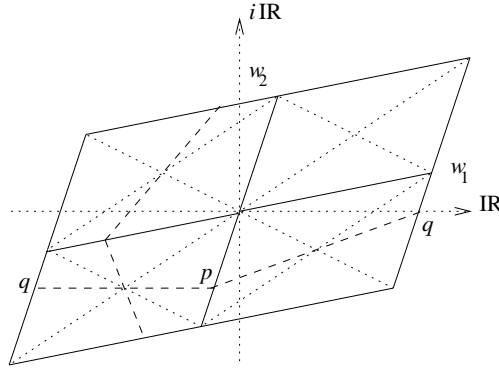


Figura 2:  $p, q$  diagonalmente opostos na subgrade,  $\wp(q) = -\frac{1}{\wp(p)}$ .

De fato, como  $J$  é induzida de  $\mathcal{H}$  pela  $\wp$ , temos  $(\wp \circ \mathcal{H})(p) = (J \circ \wp)(p) = \frac{-1}{\wp(p)}$ . Assim, basta mostrarmos que  $\mathcal{H}(p) = q$ . Tomamos  $v$  um ponto fixo de  $\mathcal{H}$  e  $p, q$  os extremos de um segmento com ponto médio  $v$ . Os pontos  $p, q$  e  $v$  satisfazem:  $|\vec{pq}| = 2|\vec{vq}|$ , donde  $p = 2v - q$ . Como  $2v - w_1 - w_2 \in G$ , para  $v \in \left\{ \pm \frac{w_1 + w_2}{2}, \pm \frac{w_1 - w_2}{2} \right\}$ , verificamos o resultado.

No caso especial de um toro retangular, a menos de biholomorfismo ou anti-biholomorfismo podemos assumir que  $w_1 \in \mathbb{R}_+$  e  $w_2 \in i\mathbb{R}_+$ . Algumas involuções adicionais também vão satisfazer as hipóteses do Teorema 4, a saber:

$$I_1 : z \rightarrow \bar{z};$$

$$I_2 : z \rightarrow -\bar{z};$$

$$I_3 : z \rightarrow \bar{z} + w_2;$$

$$I_4 : z \rightarrow -\bar{z} + w_1.$$

Elas estão representadas na seguinte figura:

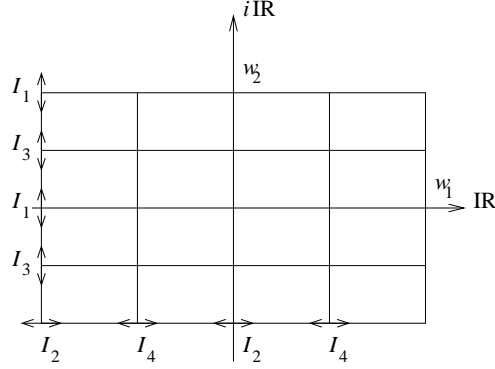


Figura 3: As involuções  $I_{1,2,3,4}$  no toro retangular  $T$ .

Junto com  $\wp$ , elas induzem involuções anti-holomorfas  $J_{1,2,3,4}$  em  $\widehat{\mathbb{C}}$  que identificaremos com o auxílio do Teorema 3.

Consideramos  $I_1$  e  $I_2$ . Elas fixam alguns pontos especiais como  $0, w_1 + w_2$  e intercambiam  $\frac{w_1 + w_2}{2}$  e  $\frac{w_1 - w_2}{2}$ . Uma vez que  $\wp(0) = 0$ ,  $\wp(w_1 + w_2) = \infty$  e  $\wp(\frac{w_1 + w_2}{2}) = -\wp(\frac{w_1 - w_2}{2}) = i$ , elas induzem a mesma involução em  $\widehat{\mathbb{C}}$ , que fixa  $0, \infty$  e intercambia  $i$  com  $-i$ . Do Teorema 3, esta involução é  $\wp \rightarrow \bar{\wp}$ . Em particular, isto significa que a imagem de  $\wp$  do conjunto dos pontos fixos de  $I_{1,2}$  é **real**. Assim,  $\wp(w_1) = \tan \alpha$  para algum  $\alpha \in (-\frac{\pi}{2}, \frac{\pi}{2}) \setminus \{0\}$ . Pela Figura 2,  $w_1$  e  $w_2$  são os extremos de um segmento com ponto médio  $\frac{w_1 + w_2}{2}$ , e concluímos que  $\wp(w_2) = -\cot \alpha$ . Como  $\tan \alpha = -\cot(\alpha + \frac{\pi}{2})$ , a menos de um anti-biholomorfismo podemos escolher o nosso toro de modo que  $\alpha > 0$ . Dessa forma, consideramos  $\alpha \in (0, \frac{\pi}{2})$ .

A involução  $I_4$  intercambia  $0$  com  $w_1$  e  $w_1 + w_2$  com  $w_2$ . Logo, se  $J_4$  for a

involução induzida por  $\wp$  (de  $I_4$ ), ela será uma transformação de Möbius anti-holomorfa que satisfaz:

$$\begin{aligned} J_4(0) &= J_4(\wp(0)) = \wp(I_4((0)) = \wp(w_1) = \tan \alpha; \\ J_4(\infty) &= J_4(\wp(w_1 + w_2)) = \wp(I_4((w_1 + w_2)) = \wp(w_2) = -\cot \alpha; \\ J_4(-\cot \alpha) &= J_4(\wp(w_2)) = \wp(I_4((w_2)) = \wp(w_1 + w_2) = \infty. \end{aligned}$$

Portanto  $J_4$  é dada por

$$\wp \rightarrow \frac{\tan \alpha - \bar{\wp}}{1 + \tan \alpha \cdot \bar{\wp}}.$$

Os cálculos para determinar a involução induzida por  $I_3$  são análogos. O próximo desenho representa  $\wp(T)$  com as suas involuções induzidas (foram dados os mesmos nomes).

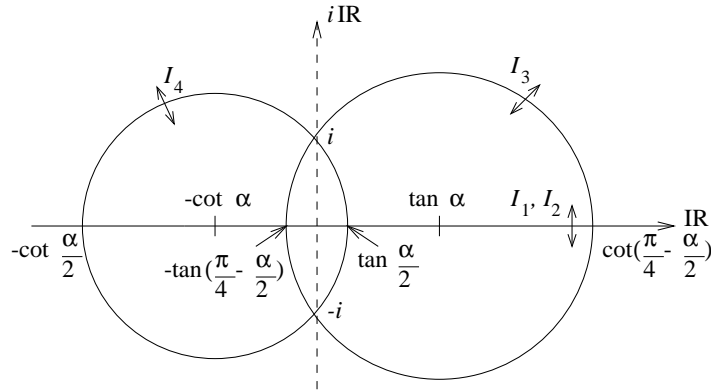


Figura 4: A imagem  $\wp(T)$  com as involuções induzidas.

Uma equação algébrica para esse toro é dada por  $\wp'^2 = c_1 \wp(\wp - \tan \alpha)(\wp + \cot \alpha)$ , sendo  $c_1$  uma constante complexa não-nula. Usualmente escrevem-se os valores de  $\wp$  sobre o toro, como na próxima figura.

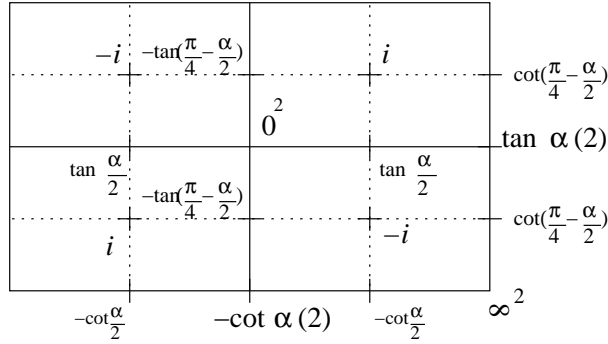


Figura 5: Os valores de  $\varphi$  no toro retangular  $T$ .

Agora estudaremos o caso especial do toro rômboico. A menos de um biholomorfismo ou anti-biholomorfismo podemos assumir que  $Arg(w_1) \in (0, \frac{\pi}{2})$  e  $w_2 = -\bar{w}_1$ . As involuções definidas anteriormente  $I_{3,4}$  não são mais válidas aqui, mas  $I_{1,2}$  e outras duas involuções ainda vão satisfazer as hipóteses do Teorema 4, a saber:

$$I_5 : z \rightarrow \bar{z} + w_1 + w_2;$$

$$I_6 : z \rightarrow -\bar{z} + w_1 - w_2.$$

Elas estão representadas na Figura:

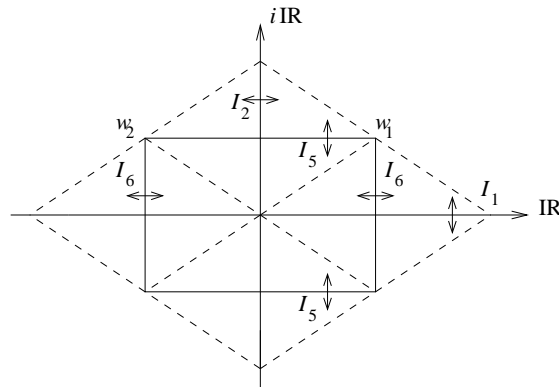


Figura 6: As involuções  $I_{1,2,5,6}$  no toro rômboico  $T$ .

Em  $\widehat{\mathbb{C}}$ ,  $I_{1,2}$  induzem a mesma involução, a saber  $\varphi \rightarrow -\bar{\varphi}$ . O mesmo ocorre com

$I_{5,6}$  que também induzem a mesma involução:  $\wp \rightarrow 1/\bar{\wp}$ . Uma consequência importante da involução  $\wp \rightarrow -\bar{\wp}$  em  $\widehat{\mathbb{C}}$  é que a imagem pela  $\wp$  da diagonal de  $T$  é o eixo imaginário. Como escolhemos  $\wp\left(\frac{w_1 + w_2}{2}\right) = i$ , a imagem pela  $\wp$  da diagonal horizontal de  $T$  é  $i\mathbb{R}_-$ , enquanto que a imagem pela  $\wp$  da diagonal vertical de  $T$  é  $i\mathbb{R}_+$ . Da involução  $\wp \rightarrow 1/\bar{\wp}$  em  $\widehat{\mathbb{C}}$ , temos que a imagem do sub-retângulo representado na Figura 6 cobre o  $S^1$ . Dessa forma,  $\wp(w_1) = e^{i\rho}$  para algum  $\rho \in (-\frac{\pi}{2}, \frac{\pi}{2})$ . Similarmente ao caso do toro retângulo, sem perder a generalidade, podemos tomar  $\rho \geq 0$  e agora considerar  $\rho \in [0, \frac{\pi}{2})$ . Da Figura 2 concluímos que  $\wp(w_2) = -e^{-i\rho}$ .

Uma equação algébrica para este toro é dada por  $\wp'^2 = c_2\wp(\wp - e^{i\rho})(\wp + e^{-i\rho})$ , sendo  $c_2$  uma constante complexa não nula.

No caso particular do toro quadrado, que é simultaneamente rômbo e retangular, temos  $\alpha = \frac{\pi}{4}$  e  $\rho = 0$ . A próxima figura descreve os valores de  $\wp$  diretamente sobre o toro.

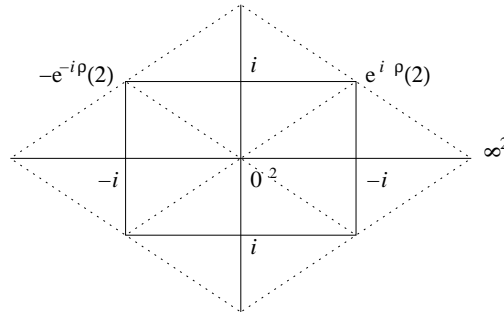


Figura 7: Os valores de  $\wp$  no toro rômbo  $T$ .

**Observação 1:** Cálculos análogos poderiam ter sido desenvolvidos sobre o toro  $T := \mathbb{C}/\Gamma$ , com  $\Gamma = \{nw_1 + mw_2 | n, m \in \mathbb{Z}\}$ ,  $w_1, w_2 \in \mathbb{C}^*$  e  $\frac{w_1}{w_2} \notin \mathbb{R}$ . Uma equação algébrica para este toro é  $\wp'^2 = c\wp(\wp - x)(\wp + \frac{1}{x})$  com  $x \in \mathbb{R}_+$  e  $x$  só dependendo de  $\frac{w_1}{w_2}$ .



**Observação 2:** Quaisquer dois toros com mesmo  $\frac{w_1}{w_2}$  são biholomorfos, e reciprocamente. Basta observarmos que se  $\frac{w_1}{w_2} = \frac{\tilde{w}_1}{\tilde{w}_2}$ , então existe  $g : \mathbb{C} \rightarrow \mathbb{C}$  com  $g(z) = az$  e  $a \in \mathbb{C}^*$  tal que  $g(w_1) = \tilde{w}_1$  e  $g(w_2) = \tilde{w}_2$ . Do Teorema 2 temos que  $T$  é biholomorfa a  $\tilde{T}$ . A recíproca também é consequência imediata deste teorema 2.

Agora consideramos  $T := \mathbb{C}/\Gamma$  e  $\tilde{T} := \mathbb{C}/\tilde{\Gamma}$ , onde  $\tilde{w}_1 = g(w_1)$  e  $\tilde{w}_2 = g(w_2)$  para  $g(z) = az + b$  com  $a, b \in \mathbb{C}$  e  $a \neq 0$ , conforme a Figura 8. Sejam  $\beta(t) := t(w_1 + w_2)$ ,  $0 \leq t \leq 1$  e  $\tilde{\beta} = t(\tilde{w}_1 + \tilde{w}_2) + b = a\beta(t) + b$ ,  $0 \leq t \leq 1$ . Como vimos anteriormente,  $\wp(\beta(t))$  é uma curva que liga 0 a  $\infty$  e passa por  $i = \wp(\beta(\frac{1}{2}))$ . Logo,  $\wp(\beta(t))'|_{\frac{1}{2}} = y \in \mathbb{C}^*$ , sendo que  $y$  depende apenas de  $\frac{w_1}{w_2}$ . Por outro lado,  $[\wp(\beta(t))'|_{\frac{1}{2}}]^2 = y^2 \Rightarrow \wp'(\beta(\frac{1}{2})) \cdot [\beta'(\frac{1}{2})]^2 = y^2 \Rightarrow ci(i-x)(i+\frac{1}{x}) \cdot (w_1 + w_2)^2 = y^2 \Rightarrow c = \frac{w_2^{-2}y^2}{\left(1 + \frac{w_1}{w_2}\right)^2 \left(x - \frac{1}{x} - 2i\right)}$ .

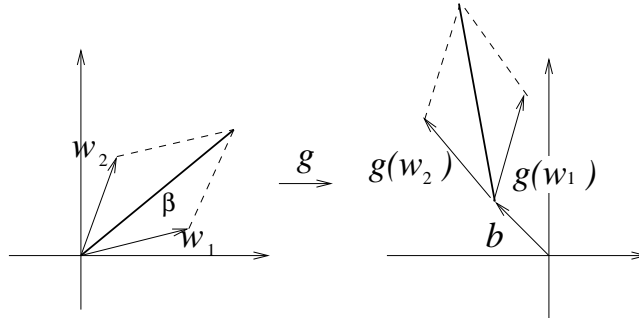


Figura 8: A aplicação  $g$ .

A relação entre  $c$  e o reticulado  $\Gamma$  para cada  $u = \frac{w_1}{w_2}$  fixo é única, pois para cada par  $(w_1, w_2), (\tilde{w}_1, \tilde{w}_2)$  com  $\frac{w_1}{w_2} = \frac{\tilde{w}_1}{\tilde{w}_2} = u$  temos:

$$c|_{(w_1, w_2)} = c|_{(\tilde{w}_1, \tilde{w}_2)} \iff (w_1 + w_2)^2 = (\tilde{w}_1 + \tilde{w}_2)^2 \iff (w_1, w_2) = \pm(\tilde{w}_1, \tilde{w}_2),$$

que por sua vez define o mesmo reticulado  $\Gamma$ .

Logo, a menos de um biholomorfismo ou anti-biholomorfismo, todos os toros estão representados num quadrante de círculo, conforme a figura abaixo.

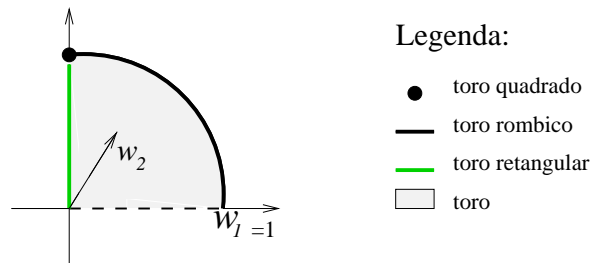


Figura 9: Representação dos tipos de toros.

### Referências Bibliográficas:

KARCHER, H.: **Construction of minimal surfaces**, *Surveys in geometry*, University of Tokyo, (1989), 1-96 and Lecture Notes 12, SFB256, Bonn, (1989).

MARTÍN, F. and RODRÍGUES, D.: **A characterization of the periodic Callahan-Hoffman-Meeks surfaces in terms of their symmetries**, *Duke Math. J.*, vol. 89, 3, (1997), 445-463.

NITSCHKE, J.C.C.: **Lectures on minimal surfaces**, Cambridge University Press, Cambridge, (1989).

## O MÉTODO SIMPLEX E O MODELO DE SISTEMA DE TRANSPORTES

Emanuele Furlan, Adriano Almeida, Ana Carvalho, Carlos Rocha, Túlio O. Carvalho

Universidade Estadual do Oeste do Paraná – UNIOESTE, [manufurlan@hotmail.com](mailto:manufurlan@hotmail.com);

Universidade Estadual do Oeste do Paraná – UNIOESTE, [adbaal@yahoo.com.br](mailto:adbaal@yahoo.com.br);

Universidade Estadual do Oeste do Paraná – UNIOESTE, [anatuccicarv@bol.com.br](mailto:anatuccicarv@bol.com.br);

Universidade Estadual do Oeste do Paraná – UNIOESTE, [rrocha41@hotmail.com](mailto:rrocha41@hotmail.com);

Universidade Estadual de Londrina – UEL, [tcarvalho@uel.br](mailto:tcarvalho@uel.br)

### RESUMO

O estudo do problema de planejamento de sistemas de transmissão recai na resolução de um problema de programação linear que pode ser resolvido através do algoritmo simplex. O algoritmo simplex é considerado fundamental na programação matemática, pois se mostra muito eficiente na resolução de problemas de programação linear e funciona como base para implementação de algoritmos para problemas de programação linear mais sofisticados. Pretendemos desenvolver uma adaptação do algoritmo Simplex, implementando-a em linguagem computacional outra que não Fortran, mais adequada sob o ponto de vista computacional atual. Desta forma, o interesse de nosso estudo está nesta interdisciplinaridade entre engenharia elétrica e matemática aplicada, uma vez que utilizaremos o método simplex para desenvolver um algoritmo que solucione o problema de determinar o número de linhas de transmissão a serem construídas entre cada barra de forma a equilibrar o sistema, ou seja, atender todas as cargas com o menor custo.

**PALAVRAS-CHAVE:** Algoritmo Simplex, Sistema de Transportes, Modelagem Matemática.

### APRESENTAÇÃO

A questão de otimizar funções lineares sujeita a restrições lineares foi considerado por Fourier em 1826. No entanto, considera-se Kantorovich, em 1939, o primeiro a reconhecer sua importância prática, ao criar um algoritmo para a solução de um problema que buscava a melhor produção possível com base na utilização ótima dos recursos disponíveis. A análise de fluxos de cargas, distribuídas em veículos de transportes, ficou conhecido como problema dos transportes no sistema de planejamento de transmissão.

Segundo CASTILLO ET. AL. (2002), ao considerarmos o planejamento de sistema de transmissão, estamos considerando um sistema elétrico no qual se deseja encontrar o plano de expansão ótimo (construção de novos circuitos) dentro de um planejamento exequível, de acordo com uma demanda específica e que opere adequadamente. Desta forma, buscar a otimização do funcionamento de tal sistema

envolve a necessidade dupla de modelar o problema e encontrar métodos de busca para uma solução ótima.

A Programação Linear (PL) trata especialmente os problemas com função objetivo e restrições representadas por funções lineares. Assim, o objetivo é especificado com uma função linear de várias variáveis, que deverá ser maximizada ou minimizada. As variáveis de decisão, por sua vez, possuem restrições de seus valores, representadas por equações ou inequações também lineares. O conjunto dessas funções é o chamado Problema de Programação Linear. Representado na seguinte forma

canônica:

$$\max Z = \sum_{j=1}^n c_j x_j$$

Sujeito a:

$$\sum_{j=1}^n a_{ij} x_j \leq b_i, \quad (i = 1, 2, \dots, m)$$

$$x_1, x_2, \dots, x_n \geq 0,$$

em que  $n$  é o número de variáveis;  $m$  é o número de restrições do problema;  $i$  é o índice de uma determinada restrição;  $j$  é o índice de uma determinada variável;  $c_i$  é o coeficiente da variável  $x_i$  da função objetivo;  $a_{ij}$  é o coeficiente da variável  $x_i$  da  $j$ -ésima restrição.

A solução do Problema de Programação Linear consiste em determinar o valor das variáveis  $x_1, x_2, \dots, x_n$  que torna máximo o valor da função  $Z$  (o problema de minimização também pode ser considerado, sem que a forma padrão acima deixe de ser aplicável, bastando para isto considerar a função  $-Z$ ).

Um dos métodos de solução mais antigos, e também um dos mais eficazes e intensamente utilizados na prática, é o Algoritmo Simplex. Através de um processo iterativo efetuado de modo que o valor da função objetivo decresça no caso de quisermos minimizá-la, ou cresça, caso desejemos maximizá-la em cada iteração, o Algoritmo Simplex pesquisa entre as soluções admissíveis (solução onde todas as restrições são satisfeitas) a solução ótima do sistema (solução admissível que maximiza ou minimiza a função objetivo). O algoritmo simplex foi proposto por G. Dantzig em 1947. No caso de haver apenas duas variáveis de decisão, é possível fazer uma interpretação gráfica. Para mais de duas variáveis, embora não seja prática, pode-se trabalhar com conceitos semelhantes, sendo fundamentais as propriedades de poliedros convexos. Cada variável é uma dimensão do gráfico. No caso de três variáveis, cada

restrição é descrita por um semi-espaço, e a interseção desses semi-espaços descreve a região possível. Considerando  $n$  variáveis, teremos um espaço  $n$  dimensional e a região formada pelos semi-espaços será chamada de simplex. Devido a convexidade, a solução ótima acontece num vértice do simplex.

Essas interpretações baseiam-se nos teoremas a seguir (LACHTERMACHER, 2007, pp34).

**Teorema 1:** O conjunto de todas as soluções viáveis de um modelo de Programação Linear é um conjunto convexo.

**Teorema 2:** Toda solução compatível básica (solução inicial) do sistema de equações lineares de um modelo de Programação Linear é um ponto extremo do conjunto de soluções viáveis, isto é, do conjunto convexo de soluções.

**Teorema 3:** Se uma função objetivo possui um único ponto ótimo finito, então este é um ponto extremo do conjunto de soluções admissíveis.

**Teorema 4:** Se a função objetivo assume valor ótimo em mais de um ponto do conjunto de soluções admissíveis, então ela assume este valor para pelo menos dois pontos extremos do conjunto convexo e para qualquer combinação convexa desses pontos extremos.

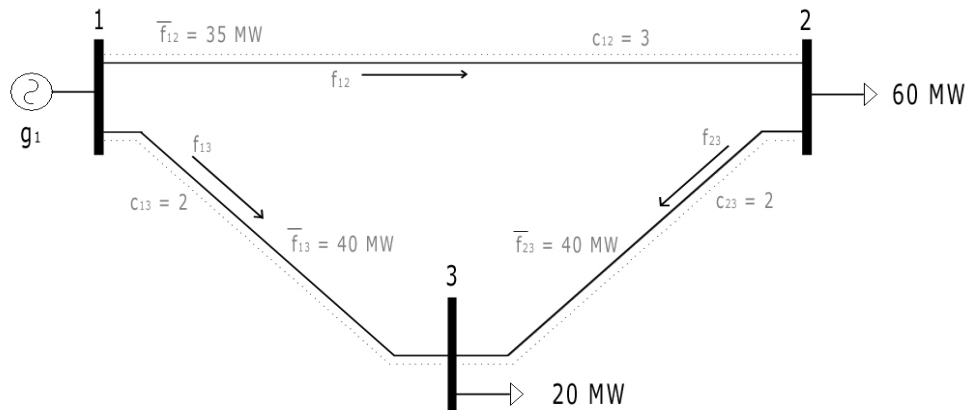
Segundo CORMEN ET AL. (2002), o Algoritmo Simplex para a minimização de uma função objetivo toma como entrada um problema linear e retorna uma solução ótima. Ele começa em algum vértice do simplex e executa uma seqüência de iterações. Em cada iteração, ele se move ao longo de uma aresta do simplex a partir de um vértice atual até um vértice vizinho cujo valor da função objetivo não é maior que o vértice atual (e normalmente menor). O Algoritmo Simplex termina quando alcança um mínimo local, que é um vértice a partir do qual todos os vértices vizinhos têm um valor de objetivo maior. Como a região possível é convexa e a função objetivo é linear, esse valor ótimo local é realmente um valor ótimo global.

No ERMAC 2007, apresentamos um trabalho sobre modelagem matemática ao problema de expansão de linhas de transmissão (BUBIAK ET AL, 2007). Neste trabalho, o objetivo é realizar um estudo teórico sobre a interface dos problemas de planejamento de transmissão e a programação linear. Pretendemos desenvolver uma adaptação do algoritmo Simplex, implementando-a em linguagem computacional outra que não Fortran; por exemplo, a linguagem JAVA, por se mostrar mais adequada sob o ponto de vista computacional atual. Desta forma, o interesse de nosso estudo está nesta interdisciplinaridade entre engenharia elétrica e matemática aplicada. Exemplificamos o

problema que iremos tratar no caso de um sistema de transmissão de três barras (CASTILLO ET. AL., 2002).

### EXEMPLO DE SISTEMA DE 3 BARRAS

A figura abaixo é usada para apresentar o modelo de transporte, no qual existem 3 barras e 3 linhas de transmissão.



Custo das linhas:  $c_{12}=3$ ,  $c_{13}=2$  e  $c_{23}=2$  U\$\$  
 Geração máxima e carga:  $g=80$ ,  $d_2=60$  e  $d_3=20$  MW  
 Fluxo máximo por linha:  $f_{12}=35$ ,  $f_{13}=40$  e  $f_{23}=40$  MW

Os sentidos positivos adotados para os fluxos, em cada caminho, são da barra de numeração menor para a barra de numeração maior.

Aplicando a primeira lei de Kirchhof em cada barra, considerando positivo o fluxo que entra e negativo o que sai da barra, obtemos as seguintes equações:

$$\begin{aligned} \text{Barra1:} & \quad -f_{12} - f_{13} + g_1 = 0 \\ \text{Barra2:} & \quad f_{12} - f_{23} - 60 = 0 \\ \text{Barra3:} & \quad f_{13} + f_{23} - 20 = 0 \end{aligned}$$

As inequações correspondentes aos limites de capacidade de transmissão produzem as seguintes relações:

$$\begin{aligned} \text{Linha 1-2:} & \quad |f_{12}| \leq 35(1 + n_{12}) \\ \text{Linha 1-3:} & \quad |f_{13}| \leq 40(1 + n_{13}) \\ \text{Linha 2-3:} & \quad |f_{23}| \leq 40(1 + n_{23}) \end{aligned}$$

A função objetivo assume a seguinte forma:  $\min v = 3n_{12} + 2n_{13} + 2n_{23}$

Assim para o exemplo, a modelagem matemática assume a seguinte forma:

$$\begin{aligned} \min v &= 3n_{12} + 2n_{13} + 2n_{23} \\ \text{s.a.} & \\ -f_{12} - f_{13} & \quad + g_1 = 0 \\ f_{12} & \quad - f_{23} = 60 \end{aligned}$$

$$\begin{aligned}
 f_{13} + f_{23} &= 20 \\
 |f_{12}| &\leq 35(1 + n_{12}) \\
 |f_{13}| &\leq 40(1 + n_{13}) \\
 |f_{23}| &\leq 40(1 + n_{23}) \\
 n_{12} &\geq 0 \\
 n_{13} &\geq 0 \\
 n_{23} &\geq 0 \\
 f_{12}, f_{13} \text{ e } f_{23} &\text{ irrestritos.}
 \end{aligned}$$

As variáveis de decisão no problema são  $n_{12}$ ,  $n_{13}$ ,  $n_{23}$ ,  $f_{12}$ ,  $f_{13}$  e  $f_{23}$ .

A função objetivo  $v$  representa o somatório dos custos de cada linha multiplicado pelos respectivos números de circuitos,  $n_{ij}$  e  $f_{ij}$  representam o número de circuitos e o fluxo do caminho  $i$ - $j$ , respectivamente.

Do ponto de vista de pesquisa operacional, o modelo que apresentamos acima é um modelo de programação linear inteiro misto, cuja resolução não é simples, especialmente para sistemas elétricos de grande porte. Permitindo-se que os  $n_{ij}$  assumam valores contínuos, o problema se transforma em um problema de PL, mesmo para o caso de sistemas de grande porte.

O problema consiste em determinar o número de linhas de transmissão a serem construídas entre cada barra de forma a equilibrar o sistema, ou seja, atender todas as cargas com o menor custo. Assim, nosso objetivo é, baseados nas idéias do processo de funcionamento do algoritmo simplex, desenvolver um algoritmo que permita a resolução do problema de planejamento de sistemas de transmissão, através da programação linear.

## REFERÊNCIAS BIBLIOGRÁFICAS

- BUBIAK, G. R.; SOUZA, F.A.; ROCHA, C.R.M. **Modelagem Matemática (Modelo DC) aplicada em problemas de Planejamento de Expansão de Sistemas de Transmissão em longo Prazo**. XI ERMAC, Curitiba, 06 – 10 agosto, UFPR. In ANAIS...,(2007)
- CASTILLO, E.; CONEJO, A.J. ; PEDREGAL, P. ;GARCÍA, R.; ALGUACIL, N. **Formulación y Resolución de Modelos de Programación Matemática em Ingeniería y Ciência**. Wiley (2002).
- CORMEN, T. H; LEISERSON, C. E.; RIVEST, R. L.; STEIN, C. **Algoritmos: teoria e prática**. 2ª ed., Rio de Janeiro, LTC, (2008).
- LACHTERMACHER, G. **Pesquisa operacional na tomada de decisões: modelagem em Excel**. Rio de Janeiro, Elsevier, (2007).

# EXISTENCE OF SOLUTIONS IN WEIGHTED SOBOLEV SPACES FOR SOME DEGENERATE QUASILINEAR ELLIPTIC EQUATIONS<sup>1</sup>

Albo Carlos Cavalheiro

Universidade Estadual de Londrina, accava@gmail.com

**Abstract:** In this paper we prove an existence result for Dirichlet problem associated to the degenerate quasilinear elliptic equations

$$-\sum_{j=1}^n D_j[\omega(x)\mathcal{A}_j(x, u, \nabla u)] = f(x), \quad \text{on } \Omega$$

in the setting of the weighted Sobolev spaces  $W_0^{1,p}(\Omega, \omega)$ .

**Keywords:** degenerate quasilinear elliptic equations, weighted Sobolev spaces

## INTRODUCTION

In this paper we prove the existence of (weak) solutions in the weighted Sobolev spaces  $W_0^{1,p}(\Omega, \omega)$  (see Definition 2.1) for the Dirichlet problem

$$(P) \begin{cases} Lu(x) = f(x), & \text{on } \Omega \\ u(x) = 0, & \text{on } \partial\Omega \end{cases}$$

where  $L$  is the partial differential operator

$$Lu(x) = -\sum_{j=1}^n D_j[\omega(x)\mathcal{A}_j(x, u(x), \nabla u(x))], \quad (1)$$

where  $\omega$  is a weight function (see section 2),  $\Omega$  is a bounded open set in  $\mathbb{R}^n$  and

$\mathcal{A}_j : \Omega \times \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}$  are given functions satisfying the following conditions:

**(H1)**  $x \mapsto \mathcal{A}_j(x, \eta, \xi)$  is measurable on  $\Omega$  for all  $(\eta, \xi) \in \mathbb{R} \times \mathbb{R}^n$

$(\eta, \xi) \mapsto \mathcal{A}_j(x, \eta, \xi)$  is continuous on  $\mathbb{R} \times \mathbb{R}^n$  for almost all  $x \in \Omega$ .

---

<sup>1</sup>Georgian Mathematical Journal (to appear)



(H2)  $[\mathcal{A}(x, \eta, \xi) - \mathcal{A}(x, \eta', \xi')].(\xi - \xi') \geq 0$ , whenever  $\xi, \xi' \in \mathbb{R}^n$ ,  $\xi \neq \xi'$ , where

$$\mathcal{A}(x, \eta, \xi) = (\mathcal{A}_1(x, \eta, \xi), \dots, \mathcal{A}_n(x, \eta, \xi)).$$

(H3)  $\mathcal{A}(x, \eta, \xi). \xi \geq \lambda|\xi|^p - g_1(x)|\eta| - g_2(x)$ , with  $1 < p < \infty$ ,  $g_1 \in L^{p'}(\Omega, \omega)$  and  $g_2 \in L^1(\Omega, \omega)$ .

(H4)  $|\mathcal{A}(x, \eta, \xi)| \leq K(x) + h_1(x)|\eta|^{p/p'} + h_2(x)|\xi|^{p/p'}$ , where  $K, h_1$  and  $h_2$  are positive functions, with  $h_1 \in L^\infty(\Omega)$ ,  $h_2 \in L^\infty(\Omega)$  and  $K \in L^{p'}(\Omega, \omega)$ .

By a *weight*, we shall mean a locally integrable function  $\omega$  on  $\mathbb{R}^n$  such that  $0 < \omega(x) < \infty$  for a.e.  $x \in \mathbb{R}^n$ . Every weight  $\omega$  gives rise to a measure on the measurable subsets on  $\mathbb{R}^n$  through integration. This measure will also be denoted by  $\omega$ . Thus,  $\omega(E) = \int_E \omega(x) dx$  for measurable sets  $E \subset \mathbb{R}^n$ .

In general, the Sobolev spaces  $W^{k,p}(\Omega)$  without weights occur as spaces of solutions for elliptic and parabolic partial differential equations. For degenerate partial differential equations, i.e., equations with various types of singularities in the coefficients, it is natural to look for solutions in weighted Sobolev spaces (see [1], [2] and [3]).

A class of weights, which is particularly well understood, is the class of  $A_p$  weights (or Muckenhoupt class) that was introduced by B. Muckenhoupt (see [4]). These classes have found many useful applications in harmonic analysis (see [5]). Another reason for studying  $A_p$ -weights is the fact that powers of distance to submanifolds of  $\mathbb{R}^n$  often belong to  $A_p$  (see [6]). There are, in fact, many interesting examples of weights (see [7] for  $p$ -admissible weights).

The following theorem will be proved in section 3.

**THEOREM 1.1** Assume (H1), (H2), (H3) and (H4). If  $\omega \in A_p$ ,  $1 < p < \infty$ , and  $f/\omega \in L^{p'}(\Omega, \omega)$  then problem (P) has a solution  $u \in W_0^{1,p}(\Omega, \omega)$ .

## 0.1 DEFINITIONS AND BASIC RESULTS

Let  $\omega$  be a locally integrable nonnegative function in  $\mathbb{R}^n$  and assume that  $0 < \omega < \infty$  almost everywhere. We say that  $\omega$  belongs to the Muckenhoupt class  $A_p$ ,  $1 < p < \infty$ , or that  $\omega$  is an  $A_p$ -weight, if there is a constant  $C = C_{p,\omega}$  such that

$$\left( \frac{1}{|B|} \int_B \omega(x) dx \right) \left( \frac{1}{|B|} \int_B \omega^{1/(1-p)}(x) dx \right)^{p-1} \leq C$$

for all balls  $B \subset \mathbb{R}^n$ , where  $|\cdot|$  denotes the  $n$ -dimensional Lebesgue measure in  $\mathbb{R}^n$ . If  $1 < q \leq p$ , then  $A_q \subset A_p$  (see [5],[7] or [8] for more informations about  $A_p$ -weights). The weight  $\omega$  satisfies the doubling condition if  $\omega(2B) \leq C\omega(B)$ , for all balls  $B \subset \mathbb{R}^n$ , where  $\omega(B) = \int_B \omega(x) dx$  and  $2B$  denotes the ball with the same center as  $B$  which is twice as large. If  $\omega \in A_p$ , then  $\omega$  is doubling (see corollary 15.7 in [7]).

As an example of  $A_p$ -weight, the function  $\omega(x) = |x|^\alpha$ ,  $x \in \mathbb{R}^n$ , is in  $A_p$  if and only if  $-n < \alpha < n(p-1)$  (see corollary 4.4, chapter IX in [5]).

Given a open subset  $\Omega$  of  $\mathbb{R}^n$ , we will denote by  $L^p(\Omega, \omega)$  ( $1 \leq p < \infty$ ), the Banach space of all measurable functions  $f$  defined on  $\Omega$  for which

$$\|f\|_{L^p(\Omega, \omega)} = \left( \int_\Omega |f(x)|^p \omega(x) dx \right)^{1/p} < \infty.$$

If  $\omega \in A_p$ ,  $1 < p < \infty$ , then  $\omega^{-1/(p-1)}$  is locally integrable and we have

$$L^p(\Omega, \omega) \subset L^1_{\text{loc}}(\Omega)$$

for every open set  $\Omega$  (see Remark 1.2.4 in [9]). It thus makes sense to talk about weak derivatives of functions in  $L^p(\Omega, \omega)$ .

**DEFINITION 2.1** Let  $\Omega \subset \mathbb{R}^n$  be open,  $1 < p < \infty$ ,  $k$  a nonnegative interger and  $\omega \in A_p$ . We define the weighted Sobolev space  $W^{k,p}(\Omega, \omega)$  as the set of functions  $u \in L^p(\Omega, \omega)$  with

weak derivatives  $D^\alpha u \in L^p(\Omega, \omega)$ ,  $1 \leq |\alpha| \leq k$ . The norm of  $u$  in  $W^{k,p}(\Omega, \omega)$  is defined by

$$\|u\|_{W^{k,p}(\Omega, \omega)} = \left( \int_{\Omega} |u(x)|^p \omega(x) dx + \sum_{1 \leq |\alpha| \leq k} \int_{\Omega} |D^\alpha u(x)|^p \omega(x) dx \right)^{1/p}. \quad (2)$$

We also define  $W_0^{k,p}(\Omega, \omega)$  as the closure of  $C_0^\infty(\Omega)$  with respect to the norm

$$\|u\|_{W_0^{k,p}(\Omega, \omega)} = \left( \sum_{1 \leq |\alpha| \leq k} \int_{\Omega} |D^\alpha u(x)|^p \omega(x) dx \right)^{1/p}.$$

If  $\omega \in A_p$ , then  $W^{k,p}(\Omega, \omega)$  is the closure of  $C^\infty(\Omega)$  with respect to the norm (2.1) (see Proposition 3.5 in [10] or Theorem 2.1.4 in [9]).

The spaces  $W^{k,p}(\Omega, \omega)$  and  $W_0^{k,p}(\Omega, \omega)$  are Banach spaces, and for  $k = 1$  and  $p = 2$  the spaces  $W^{1,2}(\Omega, \omega)$  and  $W_0^{1,2}(\Omega, \omega)$  are Hilbert spaces.

Since  $W^{k,p}(\Omega, \omega)$  is isomorphic to the Sobolev space  $W^{k,p}(\Omega)$  for a weight  $\omega$  with  $0 < C_1 \leq \omega(x) \leq C_2 < \infty$ , we suppose that  $\omega$  either vanishes somewhere in  $\bar{\Omega}$  or is not bounded (or both).

In this paper we use frequently the following two theorems.

**THEOREM 2.2**(The weighted Sobolev inequality) Let  $\Omega \subset \mathbb{R}^n$  be a bounded open set and  $\omega \in A_p$ ,  $1 < p < \infty$ . Then there exist positive constants  $C_\Omega$  and  $\delta$  such that for all  $f \in C_0^\infty(\Omega)$  and  $1 \leq \theta \leq n/(n-1) + \delta$ ,

$$\|f\|_{L^{\theta p}(\Omega, \omega)} \leq C_\Omega \|\nabla f\|_{L^p(\Omega, \omega)},$$

where the symbol  $\nabla$  indicates the gradient.

**Proof.** See Theorem 1.3 in [2]. □

**THEOREM 2.3** Let  $\omega \in A_p$ ,  $1 < p < \infty$ , and let  $\Omega$  be a bounded open set in  $\mathbb{R}^n$ . If  $u_n \rightarrow u$  in  $L^p(\Omega, \omega)$  then there exist a subsequence  $\{u_{n_k}\}$  and a function  $v \in L^p(\Omega, \omega)$  such that

(i)  $u_{n_k}(x) \rightarrow u(x)$ ,  $n_k \rightarrow \infty$ ,  $\omega$ -a.e. on  $\Omega$ .

(ii)  $|u_{n_k}(x)| \leq v(x)$ ,  $\omega$ -a.e. on  $\Omega$ .

**Proof.** The proof of this theorem follows the lines of Theorem 2.8.1 in [6]. □

**DEFINITION 2.4** Let  $1 < p < \infty$ . We say that an element  $u \in W_0^{1,p}(\Omega, \omega)$  is a (weak) solution of problem (P) if

$$\sum_{j=1}^n \int_{\Omega} \omega(x) \mathcal{A}_j(x, u(x), \nabla u(x)) D_j \varphi(x) dx = \int_{\Omega} f(x) \varphi(x) dx$$

for all  $\varphi \in W_0^{1,p}(\Omega, \omega)$ .

## 0.2 PROOF OF THEOREM 1.1

The basic idea is to reduce the problem (P) to an operator equation  $Au = T$  and apply the theorem below.

**THEOREM 3.1** Let  $A : X \rightarrow X^*$  be a monotone, coercive and hemicontinuous operator on the real, separable, reflexive Banach space  $X$ . Then for each  $T \in X^*$  the equations  $Au = T$  has a solution  $u \in X$ .

**Proof.** See Theorem 26.A in [11]. □

We define

$$B : W_0^{1,p}(\Omega, \omega) \times W_0^{1,p}(\Omega, \omega) \rightarrow \mathbb{R} \text{ and } T : W_0^{1,p}(\Omega, \omega) \rightarrow \mathbb{R}$$

by

$$B(u, \varphi) = \sum_{j=1}^n \int_{\Omega} \omega \mathcal{A}_j(x, u, \nabla u) D_j \varphi dx = \int_{\Omega} \mathcal{A}(x, u, \nabla u) \cdot \nabla \varphi(x) \omega dx$$

$$T(\varphi) = \int_{\Omega} f \varphi dx.$$

Then  $u \in W_0^{1,p}(\Omega, \omega)$  is a (weak) solution to problem (P) if

$$B(u, \varphi) = T(\varphi), \quad \text{for all } \varphi \in W_0^{1,p}(\Omega, \omega).$$

**Step 1.** For  $j = 1, \dots, n$  we define the operator

$$\begin{aligned} F_j &: W_0^{1,p}(\Omega, \omega) \rightarrow L^{p'}(\Omega, \omega) \\ (F_j u)(x) &= \mathcal{A}_j(x, u(x), \nabla u(x)). \end{aligned}$$

We have that the operator  $F_j$  is bounded and continuous. In fact:

(i) Using (H4) and the weighted Sobolev inequality (with  $\theta = 1$ ) we obtain

$$\begin{aligned} & \|F_j u\|_{L^{p'}(\Omega, \omega)}^{p'} \\ &= \int_{\Omega} |F_j u(x)|^{p'} \omega \, dx \\ &= \int_{\Omega} |\mathcal{A}_j(x, u, \nabla u)|^{p'} \omega \, dx \\ &\leq \int_{\Omega} \left( K + h_1 |u|^{p/p'} + h_2 |\nabla u|^{p/p'} \right)^{p'} \omega \, dx \\ &\leq C_1 \int_{\Omega} (K^{p'} + h_1^{p'} |u|^p + h_2^{p'} |\nabla u|^p) \omega \, dx \\ &= C_1 \left[ \int_{\Omega} K^{p'} \omega \, dx + \int_{\Omega} h_1^{p'} |u|^p \omega \, dx + \int_{\Omega} h_2^{p'} |\nabla u|^p \omega \, dx \right] \\ &\leq C_1 \left[ \|K\|_{L^{p'}(\Omega, \omega)}^{p'} + \|h_1\|_{\infty}^{p'} \int_{\Omega} |u|^p \omega \, dx + \|h_2\|_{\infty}^{p'} \int_{\Omega} |\nabla u|^p \omega \, dx \right] \\ &\leq C_1 \left[ \|K\|_{L^{p'}(\Omega, \omega)}^{p'} + \|h_1\|_{\infty}^{p'} C_{\Omega} \int_{\Omega} |\nabla u|^p \omega \, dx + \|h_2\|_{\infty}^{p'} \int_{\Omega} |\nabla u|^p \omega \, dx \right] \\ &= C_1 \left[ \|K\|_{L^{p'}(\Omega, \omega)}^{p'} + (C_{\Omega} \|h_1\|_{\infty}^{p'} + \|h_2\|_{\infty}^{p'}) \|u\|_{W_0^{1,p}(\Omega, \omega)}^p \right]. \end{aligned}$$

Therefore

$$\|F_j u\|_{L^{p'}(\Omega, \omega)} \leq C_2 \left( \|K\|_{L^{p'}(\Omega, \omega)} + C_3 \|u\|_{W_0^{1,p}(\Omega, \omega)}^{p/p'} \right). \quad (3)$$

(ii) Let  $u_n \rightarrow u$  in  $W_0^{1,p}(\Omega, \omega)$  as  $n \rightarrow \infty$ . We need to show that  $F_j u_n \rightarrow F_j u$  in  $L^{p'}(\Omega, \omega)$ . If  $u_n \rightarrow u$  in  $W_0^{1,p}(\Omega, \omega)$ , then  $u_n \rightarrow u$  in  $L^p(\Omega, \omega)$  (by Theorem 2.2). Using Theorem 2.3, there exist a subsequence  $\{u_{n_k}\}$  and functions  $v_1$  and  $v_2$  in  $L^p(\Omega, \omega)$  such that

$$\begin{aligned} u_{n_k}(x) &\rightarrow u(x), \quad \omega - \text{a.e. in } \Omega, \\ |u_{n_k}(x)| &\leq v_1(x), \quad \omega - \text{a.e. in } \Omega, \\ |\nabla u_{n_k}(x)| &\rightarrow |\nabla u(x)|, \quad \omega - \text{a.e. in } \Omega, \\ |\nabla u_{n_k}(x)| &\leq v_2(x), \quad \omega - \text{a.e. in } \Omega. \end{aligned}$$

Hence, we obtain

$$\begin{aligned} &\|F_j u_{n_k} - F_j u\|_{L^{p'}(\Omega, \omega)}^{p'} \\ &= \int_{\Omega} |F_j u_{n_k}(x) - F_j u(x)|^{p'} \omega \, dx \\ &= \int_{\Omega} |\mathcal{A}_j(x, u_{n_k}, \nabla u_{n_k}) - \mathcal{A}_j(x, u, \nabla u)|^{p'} \omega \, dx \\ &\leq C \int_{\Omega} \left( |\mathcal{A}_j(x, u_{n_k}, \nabla u_{n_k})|^{p'} + |\mathcal{A}_j(x, u, \nabla u)|^{p'} \right) \omega \, dx \\ &\leq C \left[ \int_{\Omega} \left( K + h_1 |u_{n_k}|^{p/p'} + h_2 |\nabla u_{n_k}|^{p/p'} \right)^{p'} \omega \, dx \right. \\ &\quad \left. + \int_{\Omega} \left( K + h_1 |u|^{p/p'} + h_2 |\nabla u|^{p/p'} \right)^{p'} \omega \, dx \right] \\ &\leq 2C \int_{\Omega} \left( K + h_1 v_1^{p/p'} + h_2 v_2^{p/p'} \right)^{p'} \omega \, dx \\ &\leq C_p \int_{\Omega} [K^{p'} + h_1^{p'} v_1^p + h_2^{p'} v_2^p] \omega \, dx \\ &\leq C_p \left( \|K\|_{L^{p'}(\Omega, \omega)}^{p'} + \|h_1\|_{\infty}^{p'} \|v_1\|_{L^p(\Omega, \omega)}^p + \|h_2\|_{\infty}^{p'} \|v_2\|_{L^p(\Omega, \omega)}^p \right). \end{aligned}$$

By condition (H1) we have

$$F_j u_n(x) = \mathcal{A}_j(x, u_n(x), \nabla u_n(x)) \rightarrow \mathcal{A}_j(x, u(x), \nabla u(x)) = F_j u(x), \quad \text{as } n \rightarrow +\infty.$$

Therefore, by Dominated Convergence Theorem, we obtain

$$\|F_j u_{n_k} - F_j u_n\|_{L^{p'}(\Omega, \omega)} \rightarrow 0,$$

that is,  $F_j u_{n_k} \rightarrow F_j u$  in  $L^{p'}(\Omega, \omega)$ . By Convergence principle in Banach spaces, we have

$$F_j u_n \rightarrow F_j u \quad \text{in } L^{p'}(\Omega, \omega). \quad (4)$$

**Step 2.** Using the weighted Sobolev inequality (with  $\theta = 1$ ), we obtain

$$\begin{aligned} |T(\varphi)| &\leq \int_{\Omega} |f| |\varphi| \, dx \\ &= \int_{\Omega} \frac{|f|}{\omega} |\varphi| \omega \, dx \\ &\leq \left[ \int_{\Omega} \left( \frac{|f|}{\omega} \right)^{p'} \omega \, dx \right]^{1/p'} \left[ \int_{\Omega} |\varphi|^p \omega \, dx \right]^{1/p} \\ &\leq C_{\Omega} \|f/\omega\|_{L^{p'}(\Omega, \omega)} \|\varphi\|_{W_0^{1,p}(\Omega, \omega)}. \end{aligned}$$

Moreover, by (3.1), we also have

$$\begin{aligned} |B(u, \varphi)| &\leq \sum_{j=1}^n \int_{\Omega} |\mathcal{A}_j(x, u, \nabla u)| |D_j \varphi| \omega \, dx \\ &= \sum_{j=1}^n \int_{\Omega} |F_j u| |D_j \varphi| \omega \, dx \\ &\leq \sum_{j=1}^n \left( \int_{\Omega} |F_j u|^{p'} \omega \, dx \right)^{1/p'} \left( \int_{\Omega} |D_j \varphi|^p \omega \, dx \right)^{1/p} \\ &\leq C_{\Omega} \sum_{j=1}^n \|F_j u\|_{L^{p'}(\Omega, \omega)} \|\varphi\|_{W_0^{1,p}(\Omega, \omega)} \\ &\leq C_{\Omega} C_2 \left( \|K\|_{L^{p'}(\Omega, \omega)} + C_3 \|u\|_{W_0^{1,p}(\Omega, \omega)}^{p/p'} \right) \|\varphi\|_{W_0^{1,p}(\Omega, \omega)}. \end{aligned}$$

Since  $B(u, \cdot)$  is linear, for each  $u \in W_0^{1,p}(\Omega, \omega)$ , there exists a linear and continuous operator  $A : W_0^{1,p}(\Omega, \omega) \rightarrow [W_0^{1,p}(\Omega, \omega)]^*$  such that

$$\langle Au, \varphi \rangle = B(u, \varphi), \quad \text{for all } u, \varphi \in W_0^{1,p}(\Omega, \omega)$$

(where  $\langle f, x \rangle$  denotes the value of the linear functional  $f$  at point  $x$ ) and

$$\|Au\|_* \leq C_\Omega C_2 \left( \|K\|_{L^{p'}(\Omega, \omega)} + C_3 \|u\|_{W_0^{1,p}(\Omega, \omega)}^{p/p'} \right).$$

Consequently, problem (P) is equivalent to the operator equation  $Au = T$ ,  $u \in W_0^{1,p}(\Omega, \omega)$ .

**Step 3.** Using condition (H2) we have

$$\begin{aligned} & \langle Au_1 - Au_2, u_1 - u_2 \rangle \\ &= B(u_1, u_1 - u_2) - B(u_2, u_1 - u_2) \\ &= \int_{\Omega} \omega \mathcal{A}(x, u_1, \nabla u_1) \cdot \nabla(u_1 - u_2) dx \\ &\quad - \int_{\Omega} \omega \mathcal{A}(x, u_2, \nabla u_2) \cdot \nabla(u_1 - u_2) dx \\ &= \int_{\Omega} \omega \left( \mathcal{A}(x, u_1, \nabla u_1) - \mathcal{A}(x, u_2, \nabla u_2) \right) \cdot \nabla(u_1 - u_2) dx \\ &\geq 0. \end{aligned}$$

Therefore the operator  $A$  is monotone. Moreover, using (H3) we obtain

$$\begin{aligned} \langle Au, u \rangle &= B(u, u) \\ &= \int_{\Omega} \omega \mathcal{A}(x, u, \nabla u) \cdot \nabla u dx \\ &\geq \int_{\Omega} \left( \lambda |\nabla u|^p - g_1 |u| - g_2 \right) \omega dx \\ &= \lambda \int_{\Omega} |\nabla u|^p \omega dx - \int_{\Omega} g_1 |u| \omega dx - \int_{\Omega} g_2 \omega dx \\ &\geq \lambda \|u\|_{W_0^{1,p}(\Omega, \omega)}^p - \|g_1\|_{L^{p'}(\Omega, \omega)} \|u\|_{L^p(\Omega, \omega)} - \|g_2\|_{L^1(\Omega, \omega)} \\ &\geq \lambda \|u\|_{W_0^{1,p}(\Omega, \omega)}^p - C_\Omega \|g_1\|_{L^{p'}(\Omega, \omega)} \|u\|_{W_0^{1,p}(\Omega, \omega)} - \|g_2\|_{L^1(\Omega, \omega)}. \end{aligned}$$



Hence, since  $1 < p < \infty$ , we have

$$\frac{\langle Au, u \rangle}{\|u\|_{W_0^{1,p}(\Omega, \omega)}} \rightarrow +\infty, \quad \text{as } \|u\|_{W_0^{1,p}(\Omega, \omega)} \rightarrow +\infty,$$

that is,  $A$  is coercive.

**Step 4.** We need to show that the operator  $A$  is continuous. Let  $u_n \rightarrow u$  in  $W_0^{1,p}(\Omega, \omega)$  as  $n \rightarrow \infty$ . We have,

$$\begin{aligned} |B(u_n, \varphi) - B(u, \varphi)| &\leq \sum_{j=1}^n \int_{\Omega} |\mathcal{A}_j(x, u_n, \nabla u_n) - \mathcal{A}_j(x, u, \nabla u)| |D_j \varphi| \omega \, dx \\ &= \sum_{j=1}^n \int_{\Omega} |F_j u_n - F_j u| |D_j \varphi| \omega \, dx \\ &\leq \sum_{j=1}^n \|F_j u_n - F_j u\|_{L^{p'}(\Omega, \omega)} \|\varphi\|_{W_0^{1,p}(\Omega, \omega)}, \end{aligned}$$

for all  $\varphi \in W_0^{1,p}(\Omega, \omega)$ . Hence,

$$\|Au_n - Au\|_* \leq \sum_{j=1}^n \|F_j u_n - F_j u\|_{L^{p'}(\Omega, \omega)}.$$

Therefore, using (3.2) we have  $\|Au_n - Au\|_* \rightarrow 0$  as  $n \rightarrow +\infty$ .

Therefore, by Theorem 3.1, the operator equation  $Au = T$  has a solution  $u \in W_0^{1,p}(\Omega, \omega)$  and it is the solution for problem (P).  $\square$

**Example.** Let  $\Omega = \{(x, y) \in \mathbb{R}^n : x^2 + y^2 < 1\}$ ,  $\omega(x, y) = (x^2 + y^2)^{-1/2}$  and consider the operator

$$\begin{aligned} Lu(x, y) = & - \frac{\partial}{\partial x} \left[ \omega(x, y) \left( \frac{\partial u}{\partial x} + \frac{1}{1 + |\nabla u|^2} \frac{\partial u}{\partial x} \right) \right] \\ & - \frac{\partial}{\partial y} \left[ \omega(x, y) \left( \frac{\partial u}{\partial y} + \frac{1}{1 + |\nabla u|^2} \frac{\partial u}{\partial y} \right) \right]. \end{aligned}$$

We have that  $\omega \in A_2$  ( $p = 2$ ), and for  $\xi = (\xi_1, \xi_2) \in \mathbb{R}^2$  we define

$$\begin{aligned}\mathcal{A}_1(x, \eta, \xi) &= \xi_1 + \frac{\xi_1}{1 + |\xi|^2} \\ \mathcal{A}_2(x, \eta, \xi) &= \xi_2 + \frac{\xi_2}{1 + |\xi|^2},\end{aligned}$$

that is,

$$\mathcal{A}(x, \eta, \xi) = (\mathcal{A}_1(x, \eta, \xi), \mathcal{A}_2(x, \eta, \xi)) = \xi + \frac{\xi}{1 + |\xi|^2},$$

satisfies the conditions (H1), (H2), (H3) (with  $\lambda = 1$ ,  $p = 2$ ,  $g_1(x) = 0$  and  $g_2(x) = 0$ ) and (H4) (with  $K(x) = 0$ ,  $h_1(x) = 0$ ,  $h_2(x) = 2$ ,  $p = p' = 2$ ). Therefore, by Theorem 1.1, the problem

$$(P) \begin{cases} Lu = \cos(x + y), & \text{on } \Omega \\ u = 0, & \text{on } \partial\Omega \end{cases}$$

has a solution  $u \in W_0^{1,2}(\Omega, \omega)$ . □

## REFERENCES

- [1] E. Fabes, D. Jerison and C. Kenig, *The Wiener test for degenerate elliptic equations*, Annls Inst. Fourier 32 (1982), 151-182.
- [2] E. Fabes, C. Kenig, R. Serapioni, *The local regularity of solutions of degenerate elliptic equations*, Comm. PDEs 7 (1982), 77-116.
- [3] B. Franchi and R. Serapioni, *Pointwise Estimates for a Class of Strongly Degenerate Elliptic Operators: a Geometrical Approach*, Ann. Scuola Norm. Sup. Pisa, 14 (1987), 527-568.
- [4] B. Muckenhoupt, *Weighted norm inequalities for the Hardy maximal function*, Trans. Am. Math. Soc. 165 (1972), 207-226.
- [5] A. Torchinsky, *Real-variable methods in Harmonic Analysis*, Academic Press, San Diego, 1986.

- [6] A. Kufner and O. John and S. Fučík, *Function Spaces*, Noordhoff International Publishing, Leyden, 1977.
- [7] J. Heinonen, T. Kilpeläinen and O. Martio, *Nonlinear Potential Theory of Degenerate Elliptic Equations*, Oxford Math. Monographs, Clarendon Press, 1993.
- [8] J. Garcia-Cuerva and J.L. Rubio de Francia, *Weighted Norm Inequalities and Related Topics*, North-Holland Mathematics Studies 116, 1985.
- [9] B.O. Turesson, *Nonlinear Potential Theory and Weighted Sobolev Spaces*, Lectures notes in Mathematics, vol. 1736, Springer-Verlag, 2000.
- [10] V. Chiadò Piat and F. Serra Cassano, *Relaxation of Degenerate Variational Integrals*, Nonlinear Anal. 22, (1994), 409-429.
- [11] E. Zeidler, *Nonlinear Functional Analysis and its Applications*, vol.II/B, Springer-Verlag, 1990.

# USO DOS ÂNGULOS E DAS TENSÕES NODAIS COMO PARÂMETRO NA SOLUÇÃO DE SISTEMAS DE EQUAÇÕES NÃO LINEARES DOS SISTEMAS DE POTÊNCIA

Alfredo Bonini Neto, Dilson Amâncio Alves.

Universidade Estadual Paulista - UNESP, [alfredoneto@aluno.feis.unesp.br](mailto:alfredoneto@aluno.feis.unesp.br),

Universidade Estadual Paulista - UNESP, [dalves@deee.feis.unesp.br](mailto:dalves@deee.feis.unesp.br).

## Resumo:

Este trabalho apresenta uma técnica de parametrização para encontrar soluções de sistemas de equações não lineares do fluxo de carga continuado (FCC). Esta técnica possibilita o traçado completo da trajetória de soluções (TS) das equações não lineares, bem como elimina o mau-condicionamento da matriz Jacobiana (J) no ponto crítico da TS. É neste ponto que a matriz J torna-se singular, ou seja, o determinante é nulo. A técnica apresentada é de fácil implementação, onde é utilizado o método de Newton-Raphson (NR) com pequenas modificações.

**Palavras-chave:** Ponto Crítico, Fluxo de Carga Continuado, Parametrização.

## 1-Apresentação

O cálculo de fluxo de carga em uma rede de energia elétrica consiste essencialmente na determinação do estado da rede. Esta rede é composta de uma barra de referência (barra  $V\theta$ ) que fornece a referência angular, barras de geração (barras PV) que geram energia elétrica, como uma usina hidroelétrica, barras de carga (barras PQ) que consome energia, como uma cidade, linhas de transmissão (LT), que ligam uma barra a outra e transformadores que elevam e diminuem a tensão. O problema do fluxo de carga pode ser formulado por um sistema de equações e inequações algébricas não lineares que correspondem, respectivamente, às leis de Kirchhoff e a um conjunto de restrições operacionais da rede elétrica e de seus componentes (MONTICELLI, 1983). Na formulação mais simples do problema (formulação básica), a cada barra da rede são associadas quatro variáveis, sendo que duas delas entram no problema como dados e duas como incógnitas.

$V_k$  \_ magnitude da tensão nodal  
 $\theta_k$  \_ ângulo da tensão nodal  
 $P_k$  \_ geração líquida de potência ativa  
 $Q_k$  \_ injeção líquida de potência reativa

Dependendo de quais variáveis nodais entram como dados e quais são consideradas como incógnitas, definem-se três tipos de barras:

PQ \_ são dados  $P_k$  e  $Q_k$ , e calculados  $V_k$  e  $\theta_k$   
 PV \_ são dados  $P_k$  e  $V_k$ , e calculados  $Q_k$  e  $\theta_k$   
 V $\theta$  \_ são dados  $V_k$  e  $\theta_k$ , e calculados  $P_k$  e  $Q_k$

O conjunto de equações do problema do fluxo de carga é formado por duas equações para cada barra, embora este trabalho tem por objetivo calcular apenas o subsistema 1, ou seja, calcular  $V$  e  $\theta$  para todas as barras, cada uma delas representando o fato de as potências ativas e reativas injetadas em uma barra serem iguais à soma dos fluxos correspondentes que deixam a barra através de linhas de transmissão, transformadores, etc. Isso corresponde à imposição das Leis de Kirchhoff e pode ser expressa matematicamente como se segue:

$$\begin{aligned}
 P_k(\underline{q}, \underline{V}) &= G_k V_k^2 - V_k \sum_{l \in \Omega_k} V_l (g_{kl} \cos \theta_{kl} + b_{kl} \sin \theta_{kl}) \\
 Q_k(\underline{q}, \underline{V}) &= -B_k V_k^2 - V_k \sum_{l \in \Omega_k} V_l (g_{kl} \sin \theta_{kl} - b_{kl} \cos \theta_{kl})
 \end{aligned} \quad (1)$$

onde  $\Omega_k$  é o conjunto de todas as barras diretamente conectadas à barra  $k$ . Lembrando que quanto maior forem os números de barras (usinas, cidades) maior é o sistema de equações não lineares e maiores são as equações. Em (MONTICELLI, 1983) pode-se ver com maiores detalhes as deduções das equações.

## 2-Método de NR para Solução das Equações Não Lineares do Fluxo de Carga

Subsistema 1 (dimensão: 2 NPQ+NPV)

Neste subproblema são dados  $P_k$  e  $Q_k$  nas barras  $PQ$  e  $P_k$  e  $V_k$  nas barras  $PV$ ; pretende-se calcular  $V_k$  e  $\theta_k$  nas barras  $PQ$ , e  $\theta_k$  nas barras  $PV$ . Ou seja, trata-se de um sistema de  $2 \text{ NPQ} + \text{NPV}$  equações algébricas não lineares com o mesmo número de incógnitas

$$\text{Seja } G(\underline{\theta}, \underline{V}) = 0 \quad (2)$$

ou:

$$\begin{aligned} \underline{\Delta P} &= \underline{P}^{esp} - \underline{P}(\underline{q}, \underline{V}) = 0 \\ \underline{\Delta Q} &= \underline{Q}^{esp} - \underline{Q}(\underline{q}, \underline{V}) = 0 \end{aligned} \quad (3)$$

onde  $G(\theta, V)$  são as equações básicas do FC,  $V$  é o vetor das magnitudes das tensões nodais e  $\theta$  é o vetor do ângulo das tensões nodais,  $P_{esp} = P_g^{esp} - P_c^{esp}$  é a diferença entre as potências ativas geradas e consumidas para as barras de carga ( $PQ$ ) e de geração ( $PV$ ) e  $Q_{esp} = Q_g^{esp} - Q_c^{esp}$  é a diferença entre as potências reativas geradas e consumidas para as barras  $PQ$ .  $P(\theta, V)$  e  $Q(\theta, V)$  é o sistema de equação (1).

Este sistema de equações algébricas não lineares pode ser resolvido por um número muito grande de métodos (RUGGIERO E LOPES, 1996), sendo que o mais eficiente é o método de NR, onde será aplicado na resolução do subsistema 1. A seguir segue o fluxograma do Método de Newton Raphson para soluções de sistemas de equações não lineares do FCC, Figura 1:

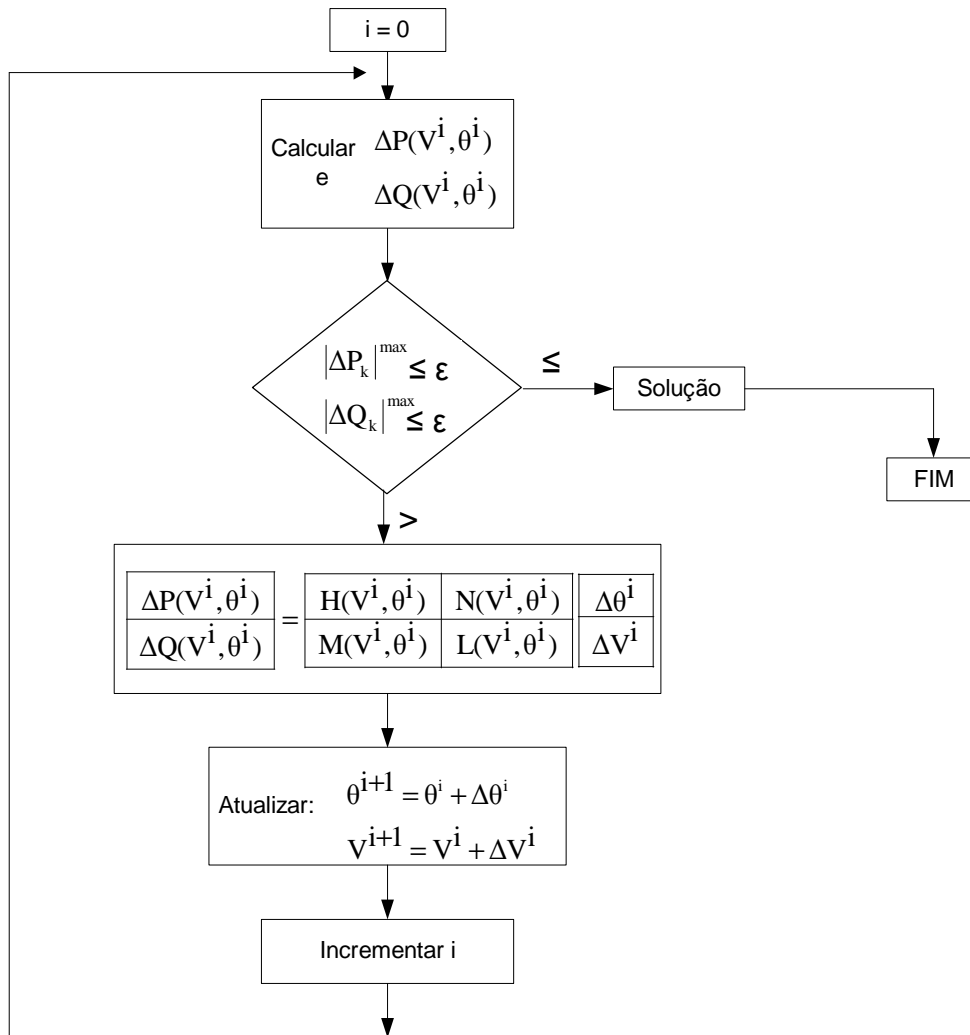


Figura 1: Fluxograma do Método de Newton

Supondo que na segunda iteração o método de NR anteriormente visto convergisse, ou seja, se  $\text{Max} \{ |\Delta P^v| \} \leq e_p$  e  $\text{Max} \{ |\Delta Q^v| \} \leq e_q$ , então os valores de  $\underline{\theta}$  e  $\underline{V}$  seria, ver Figura 2.

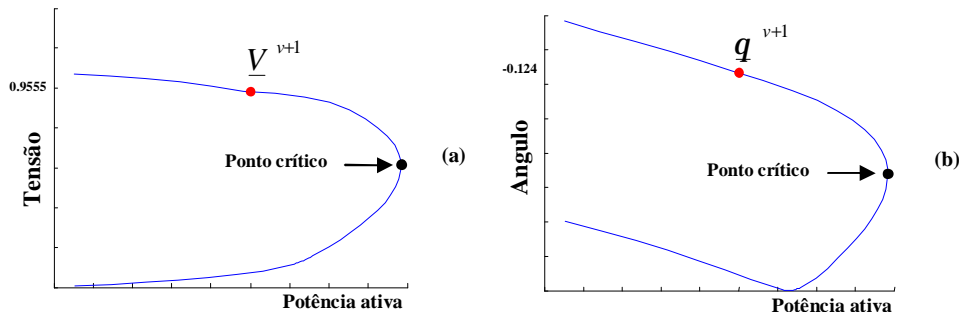


Figura 2: Curvas TS de uma rede de duas barras ( $V\theta$  e  $PQ$ ): (a)  $V$  em função de  $P$ , (b)  $\theta$  em função de  $P$ , (obter os  $V_k$  e os  $\theta_k$  de todas as equações não lineares, subsistema 1) pelo método de NR.

### 3- Sucessivas Soluções do Método de NR para Obtenção da TS das Equações Não Lineares

A TS das equações não lineares pode ser obtida por meio de sucessivas soluções do método de NR, a partir de um caso base até próximo ao ponto crítico, para incrementos graduais da carga. Em geral a equação do fluxo de carga (FC), equação (2) e (3) pode ser reescrita como:

$$G(\underline{\theta}, \underline{V}, I) = 0 \quad (4)$$

ou ainda:

$$\begin{aligned} \Delta \underline{P} &= I \underline{P}^{esp} - \underline{P}(q, \underline{V}) = 0 \\ \Delta \underline{Q} &= I \underline{Q}^{esp} - \underline{Q}(q, \underline{V}) = 0 \end{aligned} \quad (5)$$

estas equações diferem das equações (2) e (3) devido à inclusão da variável  $\lambda$ , onde  $\lambda$  é o fator de carregamento. O traçado da trajetória de solução é efetuado através de sucessivos incrementos dessa nova variável, ver Figura 3, (a) e (b). Mas esta técnica não é muito viável para efeitos de precisão, pois o método diverge próximo ao ponto crítico, devido à singularidade da matriz  $J$  do método de NR, Figura 3(c).

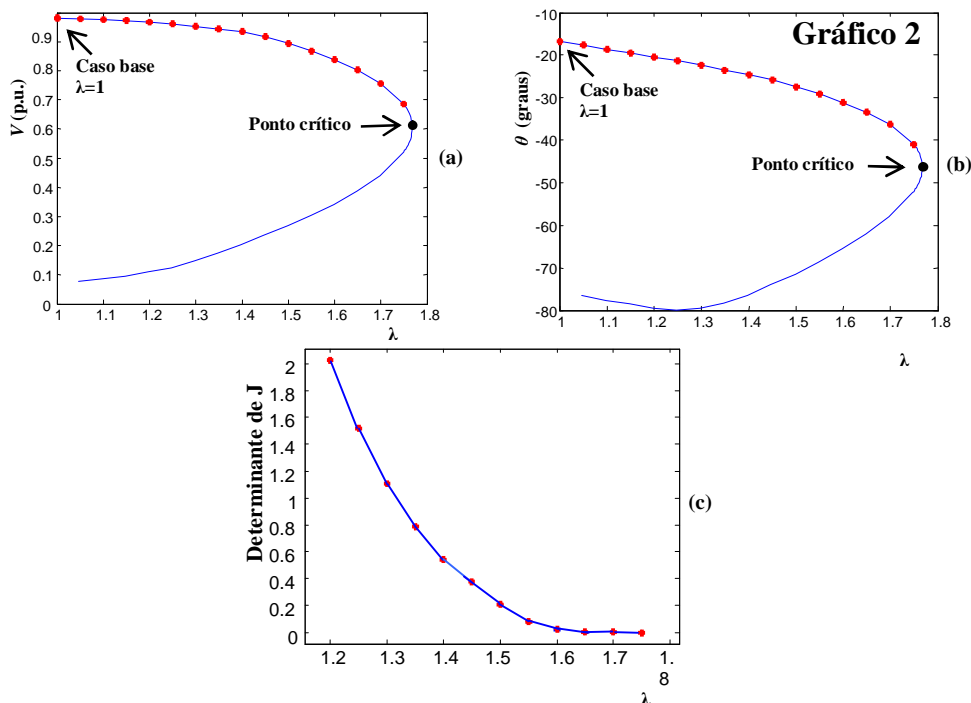


Figura 3: Curvas TS utilizando sucessivas soluções do método NR: (a)  $V$  em função de  $\lambda$ , (b)  $\theta$  em função de  $\lambda$ , (c) determinante da matriz  $J$ .



#### 4-Técnicas de Parametrizações Propostas

A técnica é inicializada utilizando as equações (5), mas como agora é especificado um parâmetro. Utiliza-se então o método de Newton-Raphson ligeiramente modificado (NRM), como o sistema de equações (5) possui uma incógnita a mais que o número de equações, será introduzido ao sistema à equação  $e_k^* \Delta_k = 0$ , onde o vetor  $e_k$  conterá 1 apenas na coluna correspondente ao novo parâmetro  $y = \underline{\theta}$ ,  $\underline{V}$  ou  $\lambda$ , ou seja, na coluna  $G_\theta$ ,  $G_V$  ou  $G_\lambda$  (AJJARAPU E CHRISTY, 1992) e (BONINI, 2006).

$$\begin{bmatrix} G_\theta & G_V & -G_\lambda \\ & e_k & \end{bmatrix} \begin{bmatrix} \Delta\theta \\ \Delta V \\ \Delta\lambda \end{bmatrix} = J_m \begin{bmatrix} \Delta\theta \\ \Delta V \\ \Delta\lambda \end{bmatrix} = \begin{bmatrix} \Delta P \\ \Delta Q \\ 0 \end{bmatrix} \quad (6)$$

onde  $G_\theta = [\partial P/\partial \theta^T \quad \partial Q/\partial \theta^T]^T$ ,  $G_V = [\partial P/\partial V^T \quad \partial Q/\partial V^T]^T$  e  $G_\lambda = [P^{esp} \quad Q^{esp}]^T$  são as derivadas parciais de  $G$  em relação a  $\underline{q}$ ,  $\underline{V}$  e  $\lambda$ , respectivamente.  $G_\theta$  e  $G_V$  compõem a matriz  $J$  do método de NR convencional. Acresce-se a  $J$  uma coluna ( $-G_\lambda$ ) correspondente a nova variável  $\lambda$

Calculado um valor para o caso base  $\lambda=1$ , obtém-se o valor de  $y$  ( $\underline{\theta}$ ,  $\underline{V}$  ou  $\lambda$ ) inicial basta dar o valor para o passo de  $y$  ( $y_{k+1} = y_k - \text{passo}$ ) e obter a curva TS desejada para o parâmetro escolhido, então para cada passo, a equação (6) fornecerá o novo valor de  $y$  obtendo a curva TS. A escolha do parâmetro é de muita importância. Para estas equações se fosse escolhido o parâmetro  $\lambda$  haveria singularidade da matriz  $J$  da mesma forma que o método de NR convencional.

Supondo que agora o  $\underline{\theta}$  fosse utilizado como parâmetro, a Figura 4 mostra a TS de  $\theta_k$  para cada  $\lambda$ , a equação (6) passaria a ser:

$$\begin{bmatrix} G_\theta & G_V & -G_\lambda \\ [0, \dots, 1] & 0 & 0 \end{bmatrix} \begin{bmatrix} \Delta\theta \\ \Delta V \\ \Delta\lambda \end{bmatrix} = J_m \begin{bmatrix} \Delta\theta \\ \Delta V \\ \Delta\lambda \end{bmatrix} = \begin{bmatrix} \Delta P \\ \Delta Q \\ 0 \end{bmatrix} \quad (7)$$

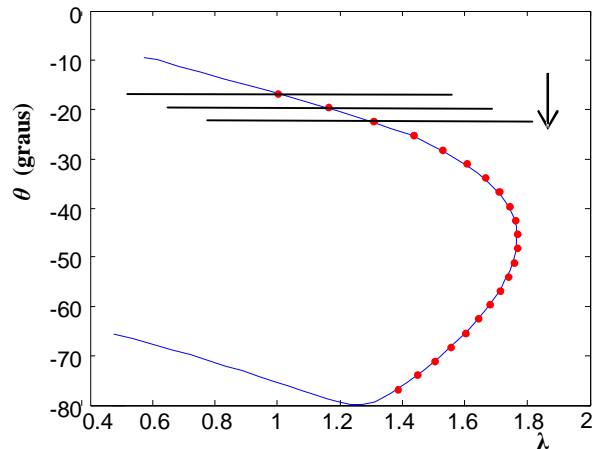


Figura 4: Curvas TS:  $\theta$  em função de  $\lambda$  pelo método de NRM.

O mesmo procedimento é utilizado quando utiliza o  $V$  ou  $\lambda$  como parâmetro, a diferença é no vetor  $e_k$ , ou seja, a posição do número 1.

## 5- Aplicações

Exemplo do método de NR para uma rede de duas barras

Barra	Tipo	P	Q	V	$\theta$
1	V $\theta$	----	----	1.1	0
2	PQ	-0.15	-0.1	*	*

Tabela 1: Dados da barra

L.T.	$r$	$x$
1___2	0.2	1

Tabela 1: Dados da linha

O sistema de equações não lineares é:

$$P_2 = V_2^2 G_{22} + V_2 V_1 (G_{21} \cos q_{21} + B_{21} \sin q_{21})$$

$$Q_2 = -V_2^2 B_{22} + V_2 V_1 (G_{21} \sin q_{21} - B_{21} \cos q_{21})$$

A matriz Jacobiana é:

$$J(q_2 V_2) = \begin{bmatrix} H_{22} = \frac{\partial P_2}{\partial q_2} & N_{22} = \frac{\partial P_2}{\partial V_2} \\ M_{22} = \frac{\partial Q_2}{\partial q_2} & L_{22} = \frac{\partial Q_2}{\partial V_2} \end{bmatrix}$$

Os valores obtidos na segunda iteração da resolução do subsistema 1 pelo método de NR são:

$$\begin{aligned} * \quad q_2^{(2)} &= -0.1240 \text{ rad. ou } -7.1047 \text{ graus} \\ V_2^{(2)} &= 0.9555 \end{aligned}$$

Agora com graduais aumentos (passo de 0.05) na potência ativa, observe a tabela 1.

Tabela 1: TS do sistema de equações não lineares do FCC obtidos pelo NR

<b>P</b>	<b>V<sub>2</sub></b>	<b>θ<sub>2</sub> (graus)</b>
0.15	0.9555	-7.1047
0.20	0.9329	-10.1012
0.25	0.9043	-13.3671
0.35	0.8135	-21.6406
0.40	0.6981	-29.6563
0.45	<b>O programa não convergiu</b>	

Ao aplicar o método NRM com  $V_2$  como parâmetro para a rede de duas barras do exemplo anterior, pode-se notar que não haverá mais problema com a singularidade devido à linha e a coluna acrescida a matriz  $J$ . A tabela 2 mostra os valores das incógnitas  $V_2$  e  $\theta_2$  obtidas pelo método de NRM com passo de 0.05.

Com isso a matriz  $J$  passa a ser:

$$J(q_2 V_2 I) = \begin{bmatrix} H_{22} = \frac{\partial P_2}{\partial q_2} & N_{22} = \frac{\partial P_2}{\partial V_2} & \frac{\partial \Delta P_2}{\partial I} \\ M_{22} = \frac{\partial Q_2}{\partial q_2} & L_{22} = \frac{\partial Q_2}{\partial V_2} & \frac{\partial \Delta Q_2}{\partial I} \\ 0 & 1 & 0 \end{bmatrix}$$

Tabela 2: TS do sistema de equações não lineares do FCC obtidos pelo NRM

$\lambda$	$V_2$	$\theta_2$ (graus)
1.0000	0.9555	-7.1049
1.2518	0.9055	-9.4033
1.4602	0.8555	-11.6374
1.6276	0.8055	-13.8156
1.7561	0.7555	-15.9447
1.8478	0.7055	-18.0308
1.9042	0.6555	-20.0788
<b>**1.9270</b>	<b>0.6055</b>	<b>-22.0934</b>
1.9177	0.5555	-24.0785
1.8776	0.5055	-26.0376
1.8079	0.4555	-27.9740
1.7099	0.4055	-29.8905
1.5847	0.3555	-31.7898
1.4333	0.3055	-33.6745
1.2569	0.2555	-35.5469
1.0563	0.2055	-37.4091
0.8327	0.1555	-39.2632
0.5870	0.1055	-41.1114

\*\* ponto crítico

Desempenho para um sistema de três barras.

Barra	Tipo	P	Q	V	$\theta$
1	V $\theta$	-----	-----	1.0	0.0
2	PQ	-0.05	-0.02	*	*
3	PV	-0.15	-----	0.98	*

O sistema de equações não lineares na resolução do subsistema 1 (2 NPQ+NPV)

é:

$$P_2 = V_2^2 G_{22} + V_2 V_1 (G_{21} \cos q_{21} + B_{21} \text{sen } q_{21}) + V_2 V_3 (G_{23} \cos q_{23} + B_{23} \text{sen } q_{23})$$

$$P_3 = V_3^2 G_{33} + V_3 V_1 (G_{31} \cos q_{31} + B_{31} \text{sen } q_{31}) + V_3 V_2 (G_{32} \cos q_{32} + B_{32} \text{sen } q_{32})$$

$$Q_2 = -V_2^2 B_{22} + V_2 V_1 (G_{21} \text{sen } q_{21} - B_{21} \cos q_{21}) + V_2 V_3 (G_{23} \text{sen } q_{23} - B_{23} \cos q_{23})$$

A matriz  $J$  do método de NR agora passa a ser:

$$J_2(q_2^{(0)} q_3^{(0)} V_2^{(0)}) = \begin{bmatrix} H_{22} = \frac{\partial P_2}{\partial q_2} & H_{23} = \frac{\partial P_2}{\partial q_3} & N_{22} = \frac{\partial P_2}{\partial V_2} \\ H_{32} = \frac{\partial P_3}{\partial q_2} & H_{33} = \frac{\partial P_3}{\partial q_3} & N_{32} = \frac{\partial P_3}{\partial V_2} \\ M_{22} = \frac{\partial Q_2}{\partial q_2} & M_{23} = \frac{\partial Q_2}{\partial q_3} & L_{22} = \frac{\partial Q_2}{\partial V_2} \end{bmatrix}$$

### 5.1-Resolução pelo método de NR

A Figura 5(a) mostra o valores da incógnita  $V_2$  e a Figura 5(b) mostra os valores das incógnitas  $\theta_2$  e  $\theta_3$  do sistema de equações não lineares resolvido pelo método de NR com passo de 0.05 para o incremento da potência ativa.

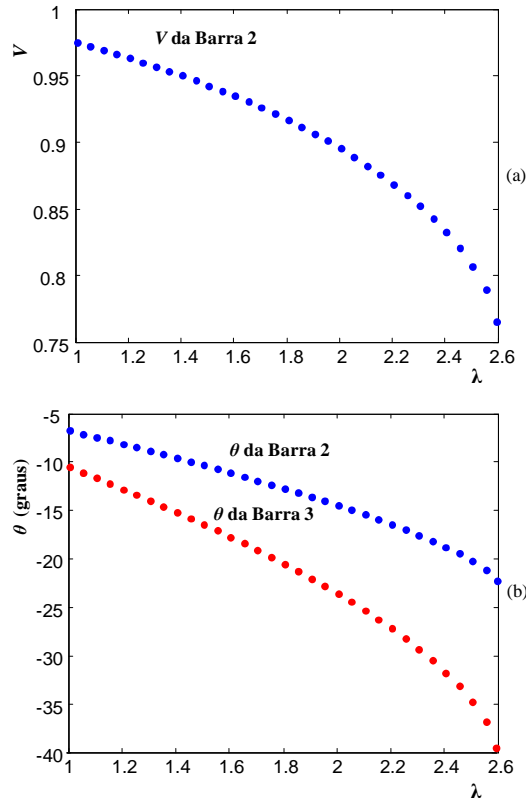


Figura 5: Curvas TS de uma rede de três barras ( $V\theta$ ,  $PQ$  e  $PV$ ): (a)  $V_2$  em função de  $\lambda$ , (b)  $\theta_2$  e  $\theta_3$  em função de  $\lambda$ , (obter os  $V_k$  e os  $\theta_k$  de todas as equações não lineares, subsistema 1) pelo método de NR.

## 5.2-Resolução pelo método de NRM

A matriz  $J$  do método de NRM é:

$$J_2(q_2^{(0)}q_3^{(0)}V_2^{(0)}I) = \begin{bmatrix} H_{22} = \frac{\partial P_2}{\partial q_2} & H_{23} = \frac{\partial P_2}{\partial q_3} & N_{22} = \frac{\partial P_2}{\partial V_2} & \frac{\partial \Delta P_2}{\partial I} \\ H_{32} = \frac{\partial P_3}{\partial q_2} & H_{33} = \frac{\partial P_3}{\partial q_3} & N_{32} = \frac{\partial P_3}{\partial V_2} & \frac{\partial \Delta P_3}{\partial I} \\ M_{22} = \frac{\partial Q_2}{\partial q_2} & M_{23} = \frac{\partial Q_2}{\partial q_3} & L_{22} = \frac{\partial Q_2}{\partial V_2} & \frac{\partial \Delta Q_2}{\partial I} \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

A Figura 6(a) mostra o valores da incógnita  $V_2$  obtidas com o uso do parâmetro  $\theta_3$  e a Figura 6(b) mostra os valores das incógnitas  $\theta_2$  e  $\theta_3$  obtidos também com o parâmetro  $\theta_3$  resolvido pelo método de NRM com passo de 0.05 para o incremento.

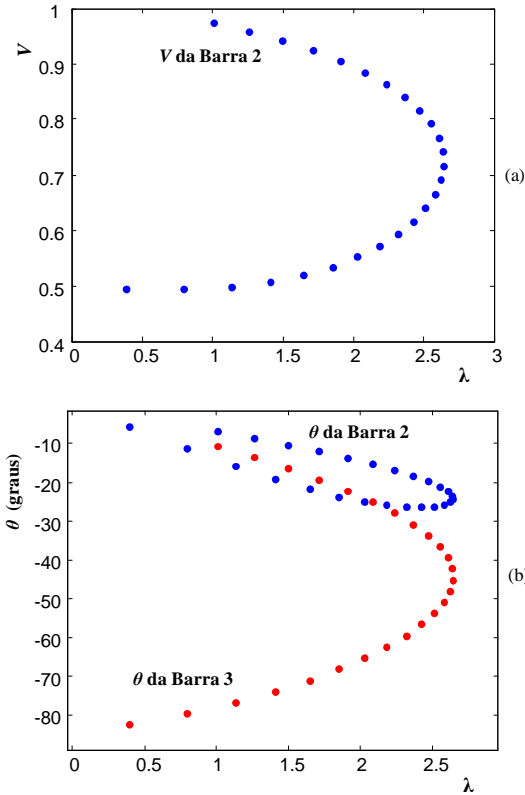


Figura 6: Curvas TS de uma rede de três barras ( $V\theta$ ,  $PQ$  e  $PV$ ): (a)  $V_2$  em função de  $\lambda$ , (b)  $\theta_2$  e  $\theta_3$  em função de  $\lambda$ , (obter os  $V_k$  e os  $\theta_k$  de todas as equações não lineares, subsistema 1) pelo método de NRM.

### 5.2.1-Desempenho para uma rede de 30 barras (IEEE)

É composto por 6 geradores (barras  $PV$ ), 24 cargas (barras  $PQ$ ), 37 linhas de transmissão e 4 transformadores.

A Figura 7(a) apresenta o traçado das TS para todas as incógnitas do sistema de equações não lineares obtidas com o uso do parâmetro  $\theta_{30}$ , bem como o número de iterações gastas para cada ponto obtido, pode-se observar o número baixo de iterações, tanto para o método de NR quanto para o método de NRM, ver Figura 7(b). Pode ser visto o mesmo procedimento na Figura 8(a), mas agora com o  $V_{30}$  como parâmetro. Se observar, na Figura 8(b) é apresentado o número de iterações utilizadas pelo NRM e pelo NR, onde o NR ficou restrito apenas na parte superior da curva. Na rede de 30 barras o sistema passa a ter (2 NPQ+NPV) equações não lineares, ou seja, contém 54 equações não lineares. E a matriz  $J$  do NRM é uma matriz de ordem 55 x 55.

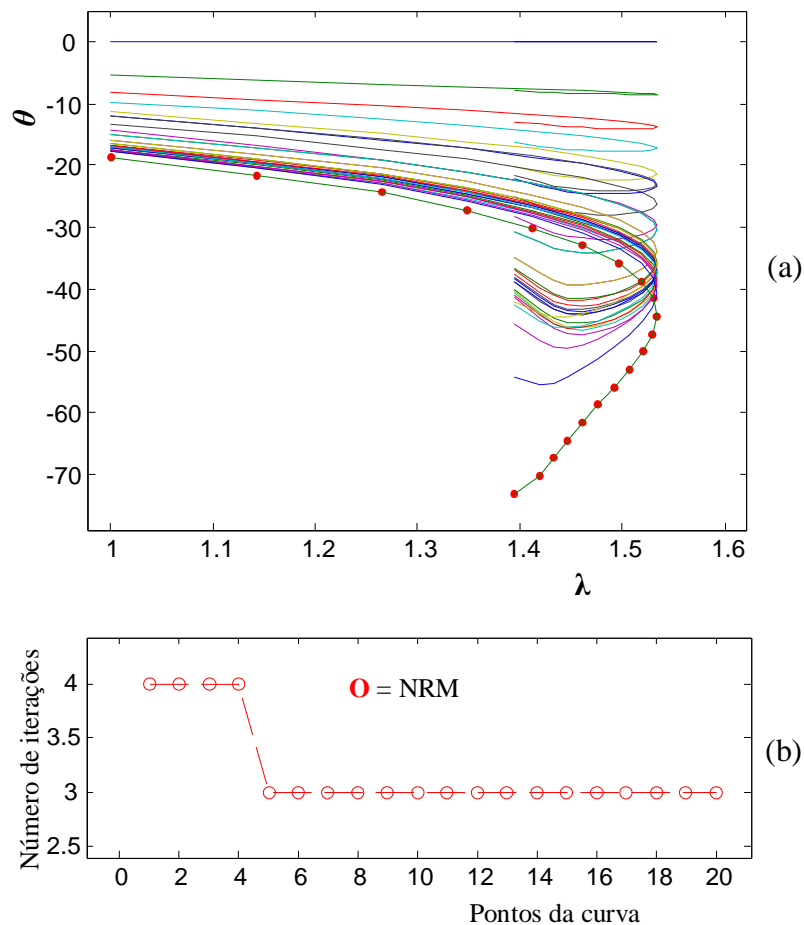


Figura 7: Curvas TS: (a)  $\theta$  em função de  $\lambda$  pelo método de NRM com  $\theta_{14}$  como parâmetro, (b) número de iterações gastas para cada ponto da curva TS.

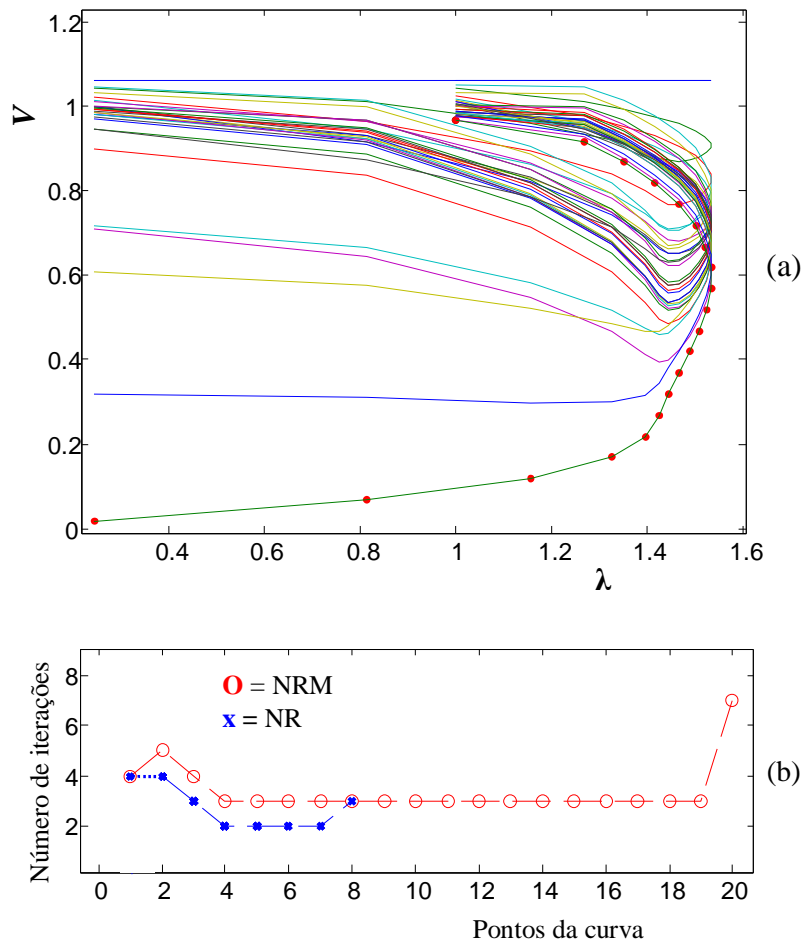


Figura 8: Curvas TS: (a)  $V$  em função de  $\lambda$  pelo método de NRM com  $V_{14}$  como parâmetro, (b) número de iterações gastas para cada ponto da curva TS.

## 6-Conclusão

Este trabalho apresentou uma técnica de parametrização geométrica para a resolução de sistemas de equações não lineares do fluxo de carga. O método de NR é considerado inadequado para a solução destas equações não lineares em torno do ponto crítico, pois o determinante é nulo. Este trabalho também mostrou que o método de NRM obteve toda a curva TS sem problemas de mau condicionamento da matriz  $J$ , ou seja, com a inclusão de uma linha e uma coluna na matriz  $J$  o determinante da matriz jacobiana modificada ( $J_m$ ) não se anula em torno do ponto crítico, com isso toda a TS foi obtida sem nenhuma dificuldade. Observa-se em termos de atuação do método de NRM comparado com o de NR, que não houve grandes mudanças na estrutura do método, em termos de número de iterações os métodos são praticamente iguais.



Com isso pretende-se modelar o método para sistemas algébricos de grande porte e com isso verificar a TS para cada incógnita. O método se torna robusto, pois pode-se utilizar qualquer uma das incógnitas como parâmetro.

### **Agradecimentos**

Agradecemos ao CNPq pelo apoio financeiro.

### **Referências**

MONTICELLI, A., “Fluxo de Carga em Redes de Energia Elétrica”, Edgard Blücher, São Paulo, 164, (1983)

RUGGIERO M. A. G., LOPES V. L. R., “Cálculo Numérico: Aspectos Teóricos e Computacionais”, 2<sup>a</sup> ed. Rio de Janeiro. Makron Books, (1996).

AJJARAPU V., CHRISTY E C., The Continuation Power Flow: a Tool for Steady State Voltage Stability Analysis, IEEE Trans. on Power Systems, vol. 7, 1, 416-423, (1992).

BONINI NETO, A. “Proposição de uma Técnica de Parametrização Geométrica para o Fluxo de Carga Continuado Baseado nas Variáveis Tensão e Fator de Carregamento”. Dissertação de mestrado. UNESP / DEE. Universidade Estadual Paulista, (2006).

# **TÉCNICA GEOMÉTRICA DO PREDITOR SECANTE APLICADA AO FLUXO DE CARGA CONTINUADO NO PLANO MAGNITUDE DE TENSÃO NODAL E FATOR DE CARREGAMENTO**

Alfredo Bonini Neto, Dilson Amâncio Alves.

Universidade Estadual Paulista - UNESP, [alfredoneto@aluno.feis.unesp.br](mailto:alfredoneto@aluno.feis.unesp.br),

Universidade Estadual Paulista - UNESP, [dalves@deee.feis.unesp.br](mailto:dalves@deee.feis.unesp.br).

## **Resumo:**

Este trabalho apresenta um esquema de parametrização geométrica para o fluxo de carga continuado que possibilita o traçado completo das curvas P-V, e o cálculo do ponto de máximo carregamento (PMC) de sistemas de potência, sem os problemas de mau-condicionamento da matriz Jacobiana ( $J$ ). O objetivo é o de apresentar de uma forma clara e didática, os passos envolvidos no desenvolvimento de uma técnica de parametrização geométrica a partir da observação das trajetórias de solução do fluxo de carga (FC). A técnica associa robustez com facilidade de compreensão. A singularidade da matriz  $J$  é eliminada pelo preditor secante que utiliza as soluções atuais e anteriores para se estimar a próxima solução, e estes dois primeiros pontos são obtidos pelo método do preditor tangente, ou seja, utiliza uma aproximação do vetor tangente. Isto é feito no plano formado nas variáveis magnitude da tensão nodal de uma barra  $k$  escolhida e o fator de carregamento. Os resultados obtidos para esta Metodologia em um sistema de 2 barras mostram que as características do método convencional são melhoradas e a região de convergência ao redor da singularidade é ampliada.

**Palavras-chave:** Colapso de Tensão, Fluxo de Carga Continuado, Ponto de Máximo Carregamento.

## **Apresentação:**

Entre os diversos estudos relacionados com o planejamento e a operação de sistemas elétricos, a análise de estabilidade estática de tensão vem, ao longo do tempo, ganhando um grande destaque. Este interesse é uma consequência direta do crescimento contínuo da demanda que, associado às restrições econômicas e ambientais, e à desregulamentação do setor elétrico, tem levado os sistemas elétricos a operar próximo

de seus limites operacionais, PMC, estando propícios à instabilidade de tensão e grandes blecautes como o ocorrido dia 14 de agosto de 2003 nos EUA causando prejuízos estimados em cerca de US\$ 6 bilhões. A análise de estabilidade estática de tensão pode ser realizada através da obtenção do perfil de tensão das barras em função de seu carregamento (curvas P-V). Entre outras aplicações, estes perfis são usados para determinar os limites de transferência de potência entre as áreas de um sistema, ajustar as margens, observar o comportamento das tensões das barras do sistema em análise e comparar estratégias de planejamento, ver figura 1.

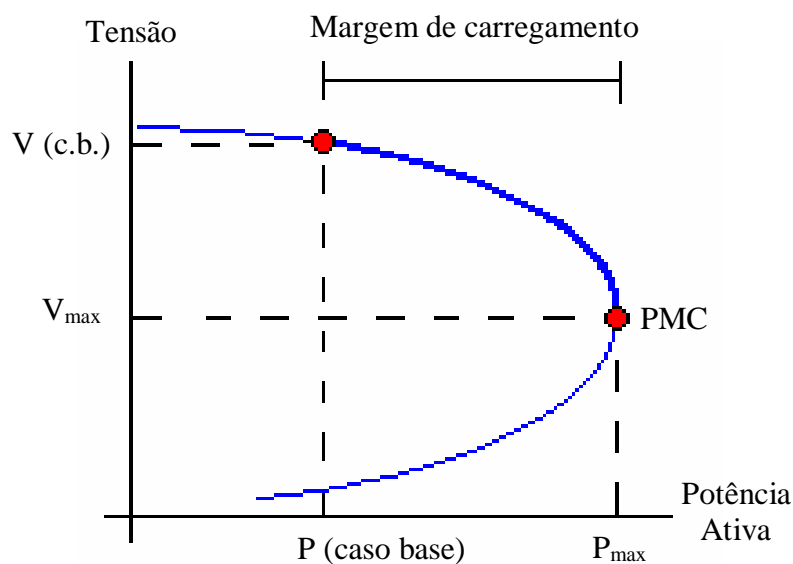


Figura 1: Curva P-V.

Um dos principais objetivos da análise estática de estabilidade de tensão é a obtenção do ponto de máximo carregamento (PMC) do sistema de energia elétrica (SEE). Sua obtenção é importante tanto para o cálculo de margens de estabilidade quanto para a realização da análise modal, uma vez que é neste ponto que ela fornece informações para a determinação de medidas efetivas para o reforço do sistema, já que o PMC define a fronteira entre as regiões de operação estável e instável.

Tradicionalmente, a obtenção de sucessivas soluções do fluxo de carga (FC) tem sido realizada através da variação manual do carregamento do sistema. Este procedimento é realizado até que o processo iterativo pare de convergir. Para fins práticos, este ponto é considerado como sendo o PMC. Entretanto, sabe-se que os problemas de convergência encontrados pelo FC convencional para a obtenção do PMC são conseqüentes das dificuldades numéricas associadas à singularidade da matriz

Jacobiana ( $J$ ). Esta instabilidade numérica é observada mesmo com o uso de cálculo com dupla precisão ou de algoritmos anti-divergentes.

Assim sendo, o uso dos métodos convencionais de FC para a obtenção das curvas P-V fica restrito à sua parte superior (correspondendo a operação estável). Além dos problemas relacionados com a singularidade de  $J$ , a solução do FC também dependerá das características comuns aos processos de solução de equações algébricas não-lineares, tais como do método utilizado na resolução, da existência da solução, das múltiplas soluções existentes, da estimativa inicial, etc. Em geral, a inicialização a partir de um caso base se deve à existência de múltiplas soluções. Por isso, o processo iterativo a ser utilizado tem que ser cuidadosamente inicializado para assegurar que uma solução apropriada seja obtida, uma vez que uma inicialização inadequada poderá resultar em não convergência, ou convergência para soluções não desejadas.

Os métodos de fluxo de carga continuado (FCC) superam as dificuldades numéricas anteriormente mencionadas pela adição de equações parametrizadas (SEYDEL, 1994), (KUNDUR, 1993), (VAN CUTSEM E VOURNAS, 1998), (CAÑIZARES et al., 1992), (ALVES et al., 2000), (AJJARAPU E CHRISTY, 1992), (CHIANG et al., 1995), (LONG E AJJARAPU, 1996), (IBA et al., 1991), (MANSOUR, 1993). As equações do FC são reformuladas visando à eliminação da singularidade da matriz  $J$  no PMC e, conseqüentemente, dos problemas numéricos que ocorrem em torno deste. Dessa forma, possibilitam que mesmo os algoritmos com simples-precisão sejam usados tanto na vizinhança quanto no próprio PMC. O traçado completo do perfil de tensão é efetuado variando automaticamente o valor de um determinado parâmetro do sistema, sem preocupação com as singularidades das equações do fluxo de carga. A diferença entre os métodos de FCC está no modo como o novo parâmetro é escolhido e como a singularidade é eliminada.

### **Método de Newton-Raphson para Solução do Fluxo de Carga**

A seguir se apresenta o FC utilizando o método de Newton-Raphson para a resolução das equações do FC (obter os valores de  $V_k$  e  $q_k$  para todas as barras, e é denominado por caso base).

Algoritmo (MONTICELLI, 1983):

$$\text{Seja } G(\underline{\theta}, \underline{V}) = 0 \quad (1)$$

ou:

$$\begin{aligned}\Delta \underline{P} &= \underline{P}^{esp} - \underline{P}(\underline{q}, \underline{V}) = 0 & \text{para as barras } PQ \text{ e } PV \\ \Delta \underline{Q} &= \underline{Q}^{esp} - \underline{Q}(\underline{q}, \underline{V}) = 0 & \text{para as barras } PQ\end{aligned}\quad (2)$$

onde  $G(\underline{\theta}, \underline{V})$  são as equações básicas do FC,  $\underline{V}$  é o vetor das magnitudes das tensões nodais e  $\underline{\theta}$  é o vetor do ângulo das tensões nodais,  $\underline{P}^{esp} = \underline{P}_g^{esp} - \underline{P}_c^{esp}$  é a diferença entre as potências ativas geradas e consumidas para as barras de carga ( $PQ$ ) e de geração ( $PV$ ) e  $\underline{Q}^{esp} = \underline{Q}_g^{esp} - \underline{Q}_c^{esp}$  é a diferença entre as potências reativas geradas e consumidas para as barras  $PQ$

As equações de injeção de potência ativa e reativa na barra  $k$ , obtidas pela primeira Lei de Kirchhoff são:

$$\begin{aligned}P_k(\underline{q}, \underline{V}) &= G_k V_k^2 - V_k \sum_{l \in \Omega_k} V_l (g_{kl} \cos \theta_{kl} + b_{kl} \sin \theta_{kl}) \\ Q_k(\underline{q}, \underline{V}) &= -B_k V_k^2 - V_k \sum_{l \in \Omega_k} V_l (g_{kl} \sin \theta_{kl} - b_{kl} \cos \theta_{kl})\end{aligned}\quad (3)$$

sendo que  $\Omega_k$  é o conjunto de todas as barras diretamente conectadas à barra  $k$ .

- i) Fazer  $v = 0$  e escolher os valores iniciais do ângulo das tensões das barras  $PQ$  e  $PV$  ( $\underline{q} = \underline{q}^0$ ), e as magnitudes das tensões das barras  $PQ$  ( $\underline{V} = \underline{V}^0$ ).
- ii) Calcular  $\underline{P}(\underline{q}^v, \underline{V}^v)$  para as barras  $PQ$  e  $PV$ , e  $\underline{Q}(\underline{q}^v, \underline{V}^v)$  para as barras  $PQ$ , e determinar os mismatches  $\Delta \underline{P}^v$  e  $\Delta \underline{Q}^v$ .
- iii) Testar convergência: se  $\text{Max} \{|\Delta \underline{P}^v|\} \leq e_p$  e  $\text{Max} \{|\Delta \underline{Q}^v|\} \leq e_q$ , o processo iterativo convergiu para a solução  $(\underline{\theta}^v, \underline{V}^v)$ ; caso contrário passar para (iv).
- iv) Calcular a matriz  $J$

$$J(\underline{q}^v, \underline{V}^v) = \begin{bmatrix} H(\underline{q}^v, \underline{V}^v) & N(\underline{q}^v, \underline{V}^v) \\ M(\underline{q}^v, \underline{V}^v) & L(\underline{q}^v, \underline{V}^v) \end{bmatrix}\quad (4)$$

onde as componentes das submatrizes Jacobianas  $H$ ,  $N$ ,  $M$  e  $L$  são dadas por:

$$H \left\{ \begin{array}{l} H_{kl} = \partial P_k / \partial q_l = V_k V_l (G_{kl} \sin q_{kl} - B_{kl} \cos q_{kl}) \\ H_{kk} = \partial P_k / \partial q_k = -V_k^2 B_{kk} - V_k \sum_{l \in K} V_l (G_{kl} \sin q_{kl} - B_{kl} \cos q_{kl}) \end{array} \right\}\quad (5)$$

$$N \left\{ \begin{array}{l} N_{kl} = \partial P_k / \partial V_l = V_k (G_{kl} \cos q_{kl} + B_{kl} \sin q_{kl}) \\ N_{kk} = \partial P_k / \partial V_k = V_k G_{kk} + \sum_{l \in K} V_l (G_{kl} \cos q_{kl} + B_{kl} \sin q_{kl}) \end{array} \right\} \quad (6)$$

$$M \left\{ \begin{array}{l} M_{kl} = \partial Q_k / \partial q_l = -V_k V_l (G_{kl} \cos q_{kl} + B_{kl} \sin q_{kl}) \\ M_{kk} = \partial Q_k / \partial q_k = -V_k^2 G_{kk} + V_k \sum_{l \in K} V_l (G_{kl} \cos q_{kl} + B_{kl} \sin q_{kl}) \end{array} \right\} \quad (7)$$

$$L \left\{ \begin{array}{l} L_{kl} = \partial Q_k / \partial V_l = V_k (G_{kl} \sin q_{kl} - B_{kl} \cos q_{kl}) \\ L_{kk} = \partial Q_k / \partial V_k = -V_k B_{kk} + \sum_{l \in K} V_l (G_{kl} \sin q_{kl} - B_{kl} \cos q_{kl}) \end{array} \right\} \quad (8)$$

v) Determinar a nova solução ( $\underline{q}^{v+1}, \underline{V}^{v+1}$ ):

$$\begin{aligned} \underline{q}^{v+1} &= \underline{q}^v + \Delta \underline{q}^v \\ \underline{V}^{v+1} &= \underline{V}^v + \Delta \underline{V}^v \end{aligned} \quad (9)$$

sendo  $\Delta \underline{q}^v$  e  $\Delta \underline{V}^v$  determinados resolvendo-se o sistema linear

$$\begin{bmatrix} \Delta \underline{P}(\underline{q}^v, \underline{V}^v) \\ \Delta \underline{Q}(\underline{q}^v, \underline{V}^v) \end{bmatrix} = \begin{bmatrix} H(\underline{q}^v, \underline{V}^v) & N(\underline{q}^v, \underline{V}^v) \\ M(\underline{q}^v, \underline{V}^v) & L(\underline{q}^v, \underline{V}^v) \end{bmatrix} \cdot \begin{bmatrix} \Delta \underline{q}^v \\ \Delta \underline{V}^v \end{bmatrix} \quad (10)$$

vi) Fazer  $v+1 \rightarrow v$  e voltar para o passo (ii).

Supondo que na segunda iteração o método de Newton anteriormente visto convergisse, ou seja, se  $Max \{|\Delta \underline{P}^v|\} \leq e_p$  e  $Max \{|\Delta \underline{Q}^v|\} \leq e_q$ , então os valores de  $\underline{\theta}$  e  $\underline{V}$  seria, ver figura 2.

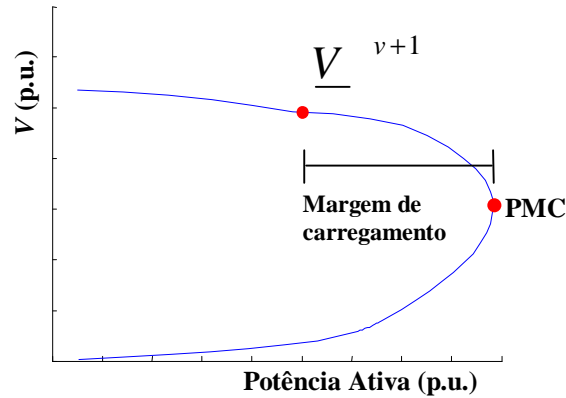


Figura 2: Curvas P-V com o cálculo do subsistema 1 (obter os  $V_k$  e os  $\theta_k$  de todas as barras) pelo método de Newton.

Fica fácil entender que basta ir aumentando a Potência Ativa do sistema e que para cada aumento, utiliza-se o Método de Newton para calcular o novo ponto ( $V$  e  $\theta$ ), com isso obtêm-se a margem de carregamento do sistema.

### Método de Newton para Obtenção da Margem de Carregamento

As curvas P-V podem ser obtidas por meio de sucessivas soluções do FC, a partir de um caso base até próximo ao PMC, para incrementos graduais da carga. Em geral a equação do FC, equação (1) pode ser reescrita como:

$$G(\underline{\theta}, \underline{V}, I) = 0 \quad (11)$$

ou ainda:

$$\begin{aligned} \Delta \underline{P} &= I \underline{P}^{esp} - \underline{P}(\underline{q}, \underline{V}) = 0 \\ \Delta \underline{Q} &= I \underline{Q}^{esp} - \underline{Q}(\underline{q}, \underline{V}) = 0 \end{aligned} \quad (12)$$

Estas equações diferem das equações (1) e (2) devido à inclusão da variável  $\lambda$ , onde  $\lambda$  é o fator de carregamento. O traçado da curva P-V é efetuado através de sucessivos incrementos desse novo parâmetro; a solução do FC para  $\lambda = 1$  fornece o ponto na curva P-V correspondente ao caso base, a Figura 2 mostra o traçado da curva P-V da barra de carga. Onde para sucessivas soluções do fluxo de carga o método de Newton-Raphson obtém a margem de carregamento do sistema Figura 3(a). Como pode-se observar na Figura 3(b), o método de Newton-Raphson diverge próximo ao PMC, devido à singularidade da matriz  $J$ .

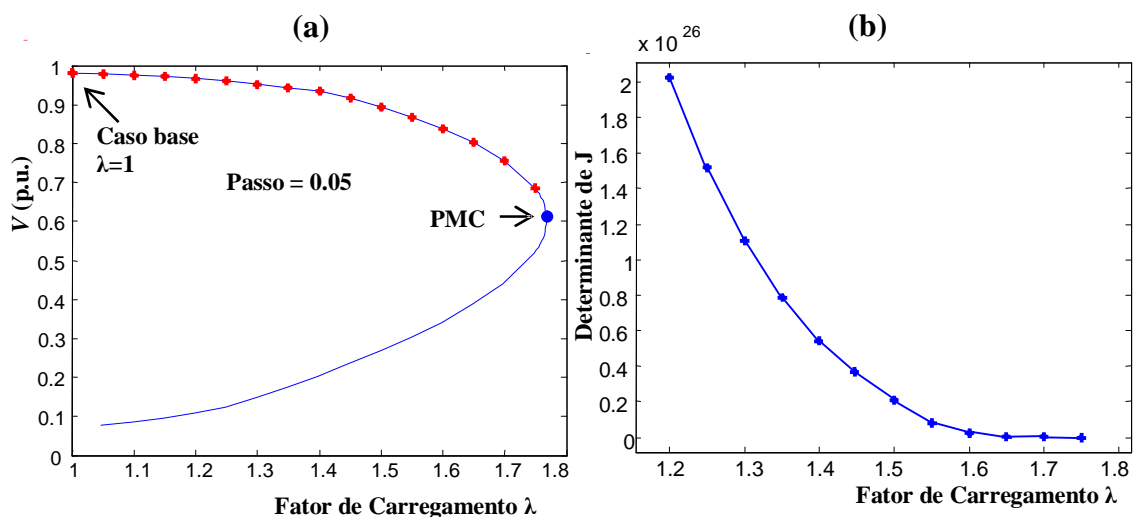


Figura 3: Sistema de duas barras, (a) Margem de Carregamento (b) Singularidade da Matriz Jacobiana do Método de Newton-Raphson

A Figura 4 a seguir ilustra um sistema de 2 barras, considerando os valores de  $V_1 = 1.19 \angle 0^\circ$  p.u. (barra slack ou referência),  $z_{12} = 0,3 + j 1,0$  p.u., e  $b_{12}^{sh} = b_2^{sh} = 0$ .

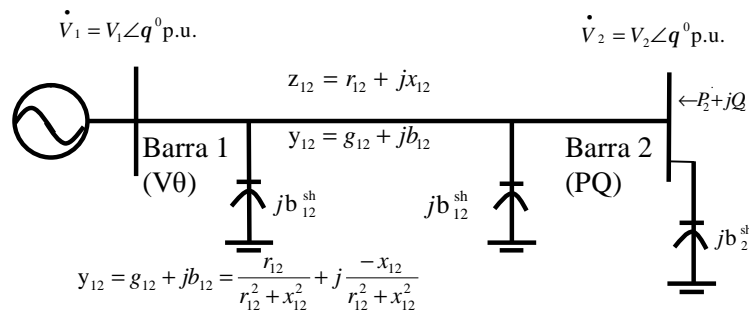


Figura 4: Sistema de duas barras

Um sistema é composto de uma barra de referência (barra  $V\theta$ ), barras de geração (barras  $PV$ ), barras de carga (barras  $PQ$ ), linhas de transmissão (LT) e transformadores. Do sistema da Figura 4, a barra 1 é uma barra de referência  $V\theta$ , por exemplo, usina hidrelétrica a qual fornece a referência angular, a barra 2 é uma barra de carga (PQ), exemplo, uma cidade (consome energia elétrica). Com isso, percebe-se que existem sistemas grandes, como o sistema Sul-Sudeste brasileiro, compostos respectivamente de 1768 barras. Em um sistema grande as demais usinas hidrelétricas interligadas são barras  $PV$  (barras de geração).

Para contornar o problema de singularidade da matriz  $J$  surgiram os métodos da continuação os quais utilizam técnicas de parametrização, ou seja, encontrar um parâmetro adequado que não só elimina o problema de singularidade, mas também obtém a curva P-V completa de todas as barras do sistema com um número baixo de iterações. Entre os diversos métodos da continuação descritos na literatura da área de sistemas de potência, o mais amplamente utilizado consiste de três elementos básicos:

- Ø Um passo preditor e parametrização
- Ø Um controle de passo
- Ø Um passo corretor e parametrização



## Metodologia

### Técnicas de Parametrizações

Parametrização é o caminho matemático para identificar cada solução na curva correspondente a trajetória de soluções (curva P-V) de forma que a próxima solução, ou a solução prevista possa ser calculada.

### Preditor

Existem dois tipos de preditores, o preditor tangente e o preditor secante, o qual vai ser utilizado neste trabalho (secante de ordem um).

O método do preditor tangente utiliza a solução atual, para se estimar a próxima solução (consiste em definir um vetor tangente partindo da solução atual).

O método do preditor secante de ordem um utiliza a solução atual e anterior, para se estimar a próxima solução, sendo que os dois primeiros pontos são obtidos pelo vetor tangente (preditor tangente).

Já o método do preditor secante de ordem zero, utiliza a solução atual e um passo fixo no parâmetro como uma estimativa para a próxima solução. O caso dos feixes de retas visto em Bonini (2006) é um preditor secante de ordem zero.

### Preditor Secante de Ordem Um.

O método do preditor secante de ordem um utiliza as soluções atuais e anteriores, para se estimar a próxima solução, e estes dois primeiros pontos são obtidos pelo método do preditor tangente, ou seja, usa uma aproximação do vetor tangente, ver Figura 5.

$$\underline{x} = (\underline{q}, \underline{V})$$

Os métodos polinomiais estão baseados em um polinômio de ordem variada que intersecta a solução atual e soluções prévias  $(\underline{x}^j, I^j)$ ,  $(\underline{x}^{j-1}, I^{j-1})$ , para prover um ponto de aproximação para a próxima solução  $(\underline{x}^{j+1}, I^{j+1})$ .

$$(\underline{x}^{j+1}, I^{j+1}) = (\underline{x}^j, I^j) + \mathbf{S} (\underline{x}^j - \underline{x}^{j-1}, I^j - I^{j-1}) \quad (13)$$

onde  $\mathbf{S}$  é o tamanho de passo apropriado.

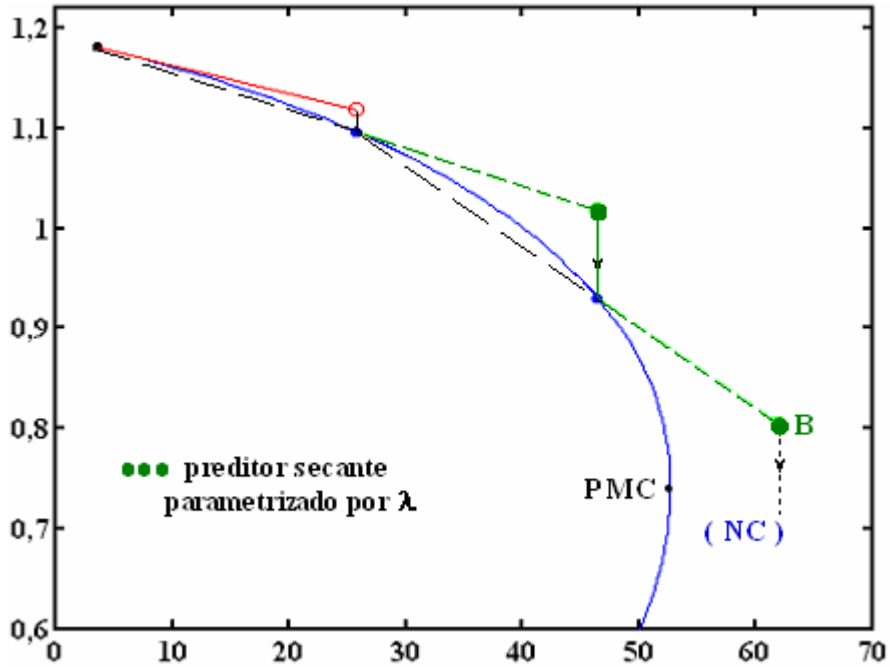


Figura 5: Preditor secante de ordem um

### Controle do Passo Preditor $\sigma$

O método da continuação para o traçado da curva P-V será tão eficiente quanto for à estratégia usada no controle do passo preditor. Para sistemas pouco carregados, passos maiores. Para sistemas altamente carregados, passos menores. O ideal seria se o tamanho do passo se adaptasse às condições reais de convergência.

O método baseado no preditor secante de ordem um  $p$  é escolhido como sendo a componente que apresentar a máxima variação relativa (SEYDEL, 1994):

$$p \leftarrow \max \left\{ \left( \frac{|q^{j+1} - q^j|}{|q^{j+1}|} \right), \left( \frac{|V^{j+1} - V^j|}{|V^{j+1}|} \right), \left( \frac{|I^{j+1} - I^j|}{|I^{j+1}|} \right) \right\} \quad (14)$$

onde  $j$  refere-se ao ponto da curva. A expectativa é de que a escolha de  $p$  com base nestas duas últimas equações implique no bom desempenho do algoritmo. Este procedimento enfatiza a confiabilidade, mas não garante que o parâmetro  $p$  conduza a um processo mais rápido (SEYDEL, 1994).

### Passo Corretor e Parametrização

Uma vez feita à previsão, torna-se necessário realizar a correção da solução aproximada para obter-se a solução correta, evitando que esse erro se acumule. Uma vez que o ponto obtido por um bom preditor encontra-se muito próximo da solução correta, poucas iterações serão necessárias para alcançar a curva trajetória de soluções (curva P-V) dentro da precisão desejada. O FC utilizando o método de Newton-Raphson é o mais utilizado no passo corretor, embora qualquer outro método numérico possa ser utilizado. Neste passo uma equação do tipo  $y - y^e = 0$ , onde  $y$  e  $y^e$  correspondem à variável escolhida como parâmetro de continuação e seu respectivo valor estimado, obtido pelo passo preditor, é acrescentada ao sistema de equações (11). Assim, o sistema de equações da etapa de correção passa a ser:

$$\begin{aligned} G(\underline{\theta}, \underline{V}, I) &= 0 \\ y - y^e &= 0, \end{aligned} \quad (15)$$

o qual é resolvido por um método de Newton-Raphson ligeiramente modificado devido inclusão da última equação. O sistema (20), expandido em série de Taylor fornece:

$$\begin{bmatrix} G_{\underline{\theta}} & G_{\underline{V}} & -G_{\lambda} \\ & e_k & \end{bmatrix} \begin{bmatrix} \Delta \underline{\theta} \\ \Delta \underline{V} \\ \Delta \lambda \end{bmatrix} = J_m \begin{bmatrix} \Delta \underline{\theta} \\ \Delta \underline{V} \\ \Delta \lambda \end{bmatrix} = \begin{bmatrix} \Delta P \\ \Delta Q \\ 0 \end{bmatrix} \quad (16)$$

observe que as submatrizes  $G_{\underline{\theta}}$  e  $G_{\underline{V}}$  formam a matriz  $J$ , e o vetor  $e_k$  conterá  $I$  apenas na coluna correspondente ao novo parâmetro  $y$ , ou seja, como  $V_k$  é escolhido como parâmetro, a coluna  $k$  da submatriz  $G_{\underline{V}}$  conterá  $I$ .

### Resultados

Para todos os testes realizados, a tolerância adotada para os mismatches utilizando o Método de Newton foi de  $10^{-4}$ . O primeiro ponto de cada curva é obtido com o método de FC convencional, denominado caso base.

A figura 6 mostra o desempenho do FCCP para o traçado da curva P-V da barra de carga 2 do sistema apresentado na Figura 4, o passo para  $\sigma$  foi de 1,4 e o parâmetro adotado foi a magnitude da tensão da barra 2.

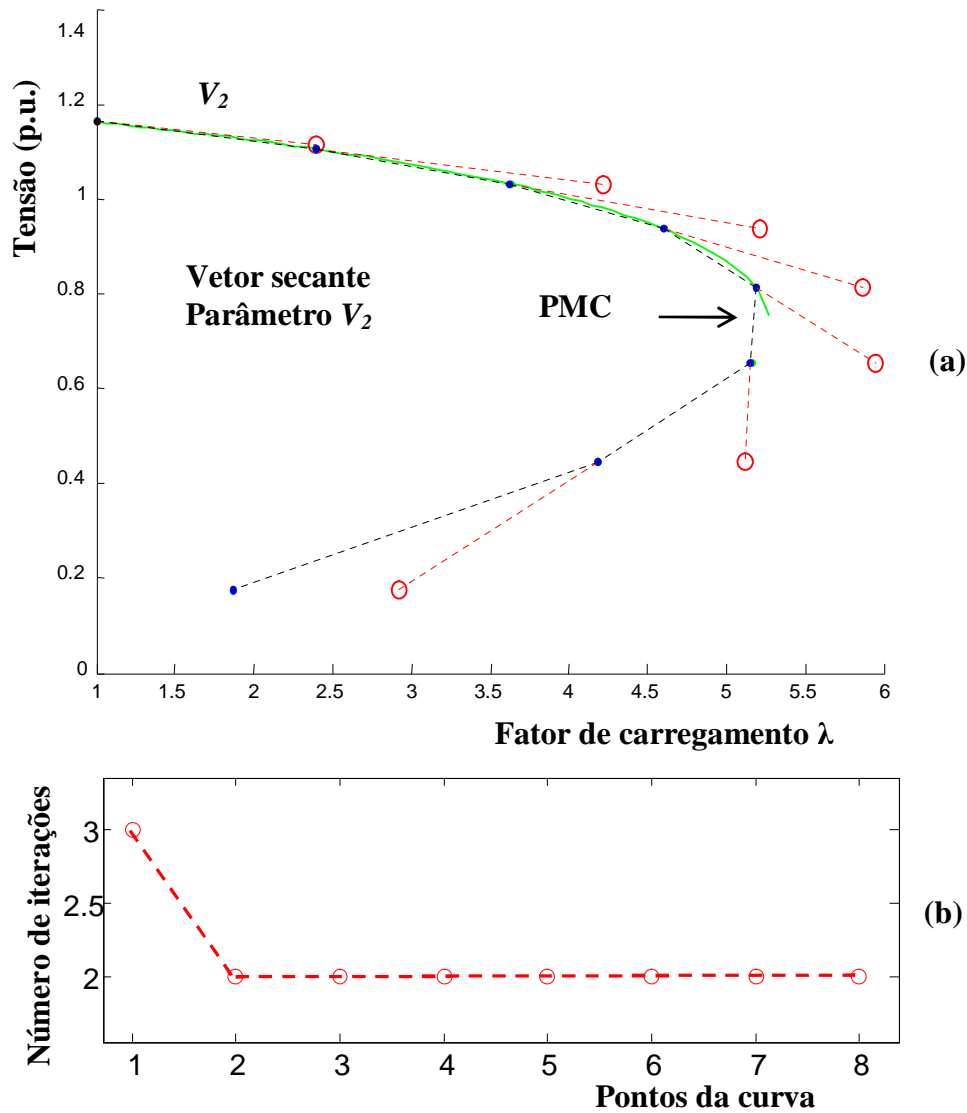


Figura 6: Desempenho da Metodologia Proposta para o sistema de 2 barras: (a) magnitude de tensão da barra 2 ( $V_2$ ) em função de  $\lambda$ , (b) número de iterações.

A Figura 7 apresenta o determinante da matriz  $J$  da Metodologia Proposta em função do carregamento  $\lambda$ , e nota-se que a Metodologia Proposta eliminou o problema de singularidade da matriz  $J$ .

A Figura 8 apresenta o traçado da curva P-V da mesma barra da figura anterior, só que agora foi iniciado com o parâmetro  $\lambda$  (pontos 2, 3 e 4 da Figura 8(a)) e se observar na região do PMC o parâmetro foi mudado para a tensão (pontos 5, 6 e 7). Isto ocorreu devido à equação (14), onde  $p$  é escolhido como sendo a componente que apresentar a máxima variação relativa. Na Figura 8(b) apresenta o número de iterações utilizadas com a Metodologia Proposta e com o Método de Newton Raphson, para os mesmos pontos (1, 2, 3, 4, 5, 6 e 7). Observa-se que o Método de Newton não obtém o

PMC nem a parte de baixo da curva, isto devido à singularidade da matriz  $J$ , retornando pelo mesmo local (parte de cima da curva P-V).

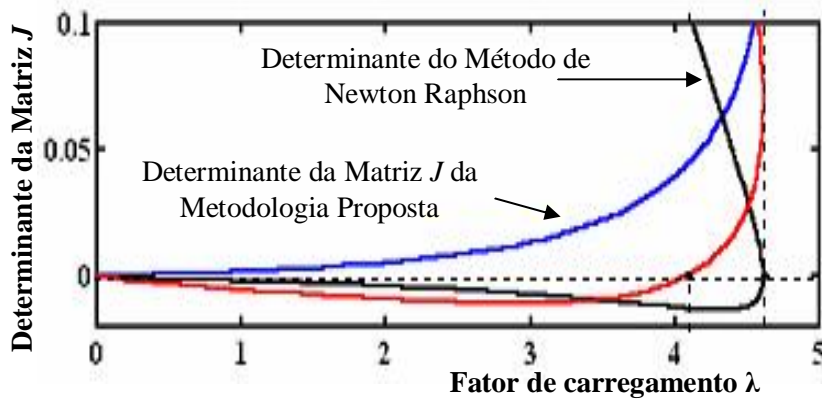


Figura 7: Determinante da Matriz  $J$  da Metodologia Proposta

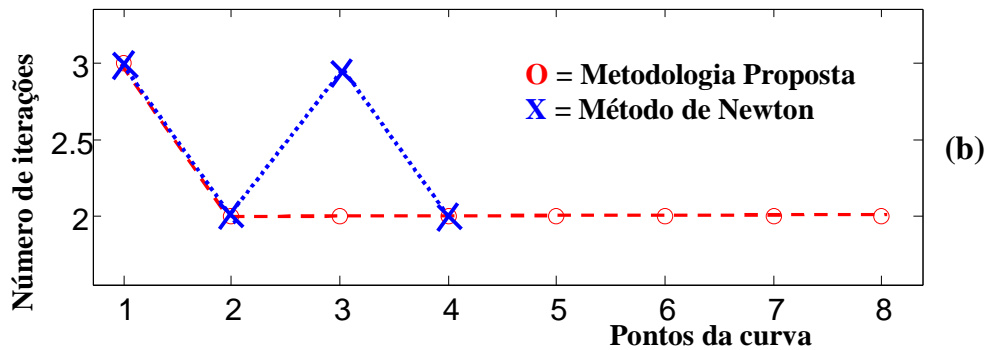
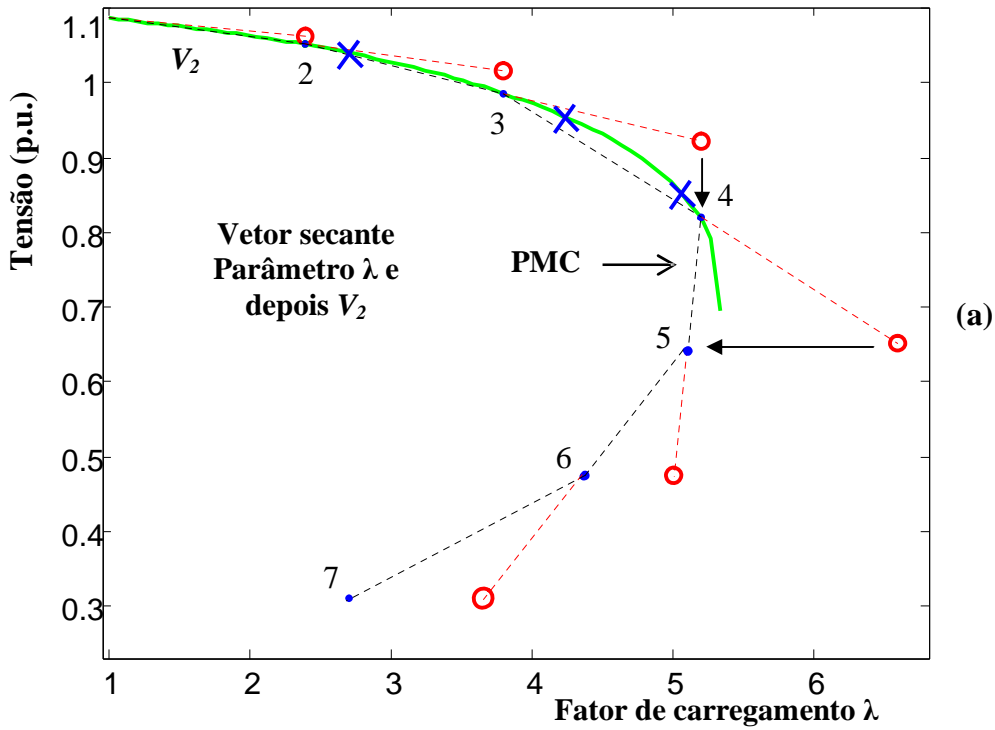


Figura 8: Desempenho da Metodologia Proposta e o Método de Newton para o sistema de 2 barras: (a) magnitude de tensão da barra 2 ( $V_2$ ) em função de  $\lambda$ , (b) número de iterações.

## Conclusões

O trabalho apresenta um esquema de parametrização que possibilita o traçado completo das curvas P-V baseado em simples modificações do método de Newton, enquanto se preserva suas vantagens características. O método proposto não só obtém êxito em encontrar, com a precisão desejada, as soluções na região do PMC e no próprio ponto, mas também permite a obtenção de soluções além deste (isto é, pontos da parte inferior da curva P-V) com um número baixo de iterações, onde o método de Newton convencional não conseguiria obter. Isto se tornou possível em virtude da remoção da singularidade da matriz Jacobiana no PMC. A remoção da singularidade foi obtida pelo Preditor Secante que utiliza as soluções atuais e anteriores para se estimar a próxima solução no plano formado pela magnitude da tensão nodal de uma barra  $k$  qualquer e o fator de carregamento  $\lambda$ .

## Referências Bibliográficas:

BARBOSA, J.L., CARMO, M.P. and ESCHENBURG, J.: **Stability of hypersurfaces of constant mean curvature in Riemannian manifolds**, Math. Z., vol. 197, 1, (1988), 123-138.

OSSERMAN, R.: **A survey of minimal surfaces**, Dover, New York, 2nd ed, (1986).

TRAIZET, M.: **Construction de surfaces minimales en recollant des surfaces de Scherk**, Ann. Inst. Fourier, vol. 46, (1996), 1385-1442.

AJJARAPU, V.; CHRISTY, C. The Continuation Power Flow: a Tool for Steady State Voltage Stability Analysis. **IEEE Trans. on Power Systems**, vol.7, 1, (1992), 416-423.

ALVES D. A.; SILVA, L. C. P.; CASTRO C. A.; COSTA V. F. Continuation load flow method parameterized by power losses. **IEEE Power Engineering Society Winter Meeting Conf.**, (2000), 1123-1128.

BONINI NETO, A. Proposição de uma Técnica de Parametrização Geométrica para o Fluxo de Carga Continuado Baseado nas Variáveis Tensão e Fator de Carregamento. **Dissertação de mestrado**. UNESP / DEE. Universidade Estadual Paulista, (2006).

CAÑIZARES, C. A.; ALVARADO, F. L.; DEMARCO, C. L.; DOBSON, I.; LONG, W. F. Point of Collapse Methods Applied to AC/DC Power Systems. **IEEE Trans. on Power Systems**, vol. 7, 2, (1992), 673-683.

CHIANG, H. D.; FLUECK, A.; SHAH, K.S.; BALU, N. CPFLOW: A Practical Tool for Tracing Power System Steady State Stationary Behavior Due to Load and Generation Variations. **IEEE Trans. on Power Systems**, vol.10, 2, (1995) 623-634.

IBA, K.; SUZUKI, H.; EGAWA, M.; WATANABE, T. Calculation of Critical Loading Condition with Nose Curve Using Homotopy Continuation Method. **IEEE Trans. on Power Systems**, vol.6, 2, (1991), 585-593.

KUNDUR, P. **Power System Stability and Control**. New York: McGraw-Hill, (1993), 1176.

LONG, B.; AJJARAPU, V. A critical review of the applications of homotopy and continuation methods to power system analysis. **28th NAPS Conf., M.I.T.**, (1996), 8-15.

MANSOUR, Y. Suggested Techniques for Voltage Stability Analysis. **IEEE Power Engineering Society**, (1993), 142.

MONTICELLI, A. J. Fluxo de carga aspectos gerais. In: **Fluxo de Carga em Redes de Energia Elétrica**. São Paulo: Edgard Blücher, (1983), 1-92.

SEYDEL, R. **From Equilibrium to Chaos: Practical Bifurcation and Stability Analysis**. 2ª ed. New York: Springer-Verlag, (1994), 407.

VAN CUTSEM, T.; VOURNAS, C. Voltage Stability of Electric Power System, **Kluwer Academic Publishers**, Boston, (1998), 376.

# SIMULAÇÃO NUMÉRICA DE ESCOAMENTOS VISCOELÁSTICOS ATRAVÉS DE UMA CONTRAÇÃO BRUSCA

Luciane Grossi Bombacini e Murilo Francisco Tomé.

Universidade Estadual de Ponta Grossa - UEPG, lg\_bombacini@yahoo.com.br;

Universidade de São Paulo - USP/ICMC -São Carlos, murilo@lcad.icmc.usp.br.

## **Resumo:**

Este trabalho tem como objetivo o desenvolvimento de um método numérico capaz de simular escoamentos viscoelásticos para o problema da contração brusca de razão 4:1 em geometria plana e axissimétrica. O escoamento de fluidos através de uma contração brusca está presente em inúmeras situações tais como: injeção de plástico, processos de extrusão e na extração de petróleo, entre outros. A simulação numérica de escoamentos de fluidos viscoelásticos através de uma contração é propensa a erros acentuados na avaliação da tensão e do campo de velocidade próximo às quinas do domínio geométrico e por serem escoamentos altamente elásticos estão sujeitos a instabilidades. Uma técnica numérica baseada em diferenças finitas é utilizada neste trabalho permitindo a análise do comportamento do fluido no problema da contração plana e axissimétrica.

**Palavras-chave:** diferenças finitas, escoamento viscoelástico, contração 4:1.

## **1. Introdução**

Escoamentos de fluidos viscoelásticos através de contrações bruscas têm sido intensamente investigados tanto experimental quanto numericamente devido a suas aplicações na indústria. O escoamento de um fluido viscoelástico através de uma contração brusca pode manifestar fenômenos complexos que desafiam os algoritmos numéricos. A grande dificuldade encontrada no tratamento deste problema é a presença de quinas, onde a tensão torna-se extremamente alta para fluidos de alta elasticidade resultando em instabilidade ou perda na convergência da solução.



No problema da contração brusca tem-se dado particular atenção aos vórtices que se formam próximos a região da contração. O tamanho e a forma dos vórtices que aparecem neste problema estão associados a: elasticidade do fluido, quantificada pelo número de Weissenberg; o tipo da contração (plana ou axissimétrica); a razão de contração e as propriedades reológicas do fluido. Ao tratar problemas para altos valores de Weissenberg, muitos algoritmos sofrem instabilidade numérica e como consequência a convergência pode não ser atingida. Este fato é atribuído a vários fatores: a presença de singularidades geométricas e condições de contorno para o escoamento, deficiência do método para tratar o problema e o domínio dos termos não lineares nas equações.

O problema da contração brusca 4:1 nas configurações plana e axissimétrica tem sido estudado por vários pesquisadores usando diferentes equações constitutivas e diferentes métodos de solução. Dentre os vários trabalhos na área, o desafio em empregar altos valores para Weissenberg ainda persiste, onde muitos autores consideram a existência de um valor limitante,  $We_{crítico}$  para o número de Weissenberg. Conjectura-se que para números de Weissenberg maiores que  $We_{crítico}$  o método numérico considerado pode não atingir a convergência. Ainda não há um consenso sobre o valor limite, pois depende da metodologia utilizada na simulação do problema.

O primeiro trabalho relevante para resolver o problema da contração plana 4:1 foi apresentado por MARCHAL et. al. (1987), o qual é considerado o pioneiro no tratamento de escoamentos com baixos números de Reynolds de um fluido Oldroyd-B. Utilizando uma técnica de elementos finitos acoplados encontrou-se  $We_{crítico} = 2.19$ . O aparecimento de "*lip vórtices*" foi detectado a partir de  $We = 0.95$ , os quais aumentaram em tamanho e intensidade com o valor de  $We$ .

PHILLIPS et. al. (1999) trataram o problema da contração plana usando um esquema semi-Lagrangiano empregando volumes finitos e em PHILLIPS et. al. (2002) estudaram a diferença no desenvolvimento dos vórtices gerados pelo escoamento de um fluido Oldroyd-B quando comparadas às geometrias de contração plana e axissimétrica, mostrando que a geometria do problema interfere no comportamento dos vórtices.

ABOUBACAR et. al. (2002) estudaram o comportamento dos vórtices na simulação numérica da contração 4:1 plana e axissimétrica para escoamentos em condições de "*creeping flow*" ( $Re = 0$ ), neste trabalho obtiveram  $We_{crítico} = 2.8$  no caso de geometria plana e  $We_{crítico} = 0.85$  com a geometria axissimétrica.

O objetivo deste trabalho é desenvolver um método numérico baseado na técnica de diferenças finitas para tratar escoamentos viscoelásticos do tipo Oldroyd-B e

empregá-lo na solução de problemas envolvendo a contração brusca de razão 4:1, no domínio geométrico planar e axissimétrico. Este trabalho foi desenvolvido com base no código computacional GENSMAC (TOME et al. (1994)), o qual resolve as equações de Navier-Stokes para variáveis primitivas sobre uma malha cartesiana uniforme.

## 2. Modelagem Matemática

As equações básicas que governam os escoamentos viscoelásticos de fluidos tipo Oldroyd-B (CROCHET et. al. (1984)) são as equação de conservação de quantidade de movimento:

$$\rho_0 \frac{D\mathbf{u}}{Dt} = -\nabla p + \nabla \cdot \boldsymbol{\sigma} + \rho_0 \mathbf{g}, \quad (1)$$

juntamente com as equações constitutivas:

$$\boldsymbol{\sigma} = -p\mathbf{I} + \boldsymbol{\tau}, \quad (2)$$

$$\boldsymbol{\tau} + \lambda_1 \overset{\nabla}{\boldsymbol{\tau}} = 2\mu_0 \left( \mathbf{D} + \lambda_2 \overset{\nabla}{\mathbf{D}} \right), \quad (3)$$

e a equação de conservação da massa

$$\nabla \cdot \mathbf{u} = 0, \quad (4)$$

em que  $\boldsymbol{\sigma}$  representa o tensor de tensões totais,  $p$  é a pressão,  $\mathbf{I}$  tensor unitário e  $\boldsymbol{\tau}$  é o tensor tensão extra e

$$D = \frac{1}{2} [(\nabla \mathbf{u}) + (\nabla \mathbf{u})^T] \quad (5)$$

é o tensor taxa de deformação.

Nas equações (1) - (3), tem-se que  $\mu_0$ ,  $\rho_0$  e  $u$  são respectivamente, viscosidade newtoniana, massa específica e velocidade,  $\lambda_1$  e  $\lambda_2$  são constantes temporais de relaxação e retardo do movimento. Nota-se que tomando  $\lambda_1 = \lambda_2$  o modelo reduz-se ao de um fluido newtoniano e tomando  $\lambda_2 = 0$  obtém-se o modelo de Maxwell. Na equação (3), o símbolo  $\overset{\nabla}{\boldsymbol{\tau}}$  representa a derivada contra-variante de  $\boldsymbol{\tau}$  (CROCHET et. al. (1984)).

Para resolver as equações (1) – (3), introduz-se a seguinte mudança de variáveis:

$$\boldsymbol{\tau} = 2\mu_0 \left( \frac{\lambda_2}{\lambda_1} \right) \mathbf{D} + \mathbf{S}, \quad (6)$$

onde  $\mathbf{S}$  representa a contribuição não newtoniana do tensor tensão extra.

Substituindo a equação (6) nas equações (1) e (3) e denotando  $L$ ,  $U$ , e  $\nu_0$  e  $\rho_0$  as unidades escalares de comprimento, velocidade, viscosidade e massa específica, respectivamente; as equações são adimensionalizadas e assumem a forma:

$$\frac{D\mathbf{u}}{Dt} = -\nabla p + \frac{1}{\text{Re}}(\beta \nabla \cdot \mathbf{D} + \nabla \cdot \mathbf{S}) + \frac{1}{Fr^2} \mathbf{g}, \quad (7)$$

$$\mathbf{S} + We \overset{\nabla}{\mathbf{S}} = 2(1 - \beta)\mathbf{D}. \quad (8)$$

onde  $\text{Re} = \frac{UL}{\nu_0}$ ,  $Fr = \frac{\sqrt{Lg}}{U}$  e  $We = \lambda_1 \frac{U}{L}$  são os números de Reynolds, Froude e

Weissenberg, respectivamente.  $\beta = \lambda_2 / \lambda_1$  (taxa que relaciona os parâmetros temporais).

As equações de conservação de massa, quantidade de movimento e constitutiva foram desenvolvidas em coordenadas bidimensionais para o caso da geometria plana GROSSI et. al. (2006) e em coordenadas cilíndricas para o caso axissimétrico (GROSSI (2003) e TOMÉ, et. al. (2007)).

### 3. Problema da contração brusca

O problema da contração brusca caracteriza-se pelo escoamento de um fluido através de um canal de diâmetro  $D$  que após uma contração brusca passa a escoar em um canal de diâmetro  $D_1 = D/4$ , obtendo a razão de contração 4:1. Diversos trabalhos na literatura tratam do problema do escoamento de fluidos viscoelásticos em contrações planares e axissimétricas. Esta geometria é interessante, pois embora seja simples a cinemática do escoamento comporta-se de forma bastante complexa.

As geometrias da contração 4:1 são apresentada abaixo. Na figura 1 tem-se a geometria da contração plana e na figura 2 a geometria da contração axissimétrica.

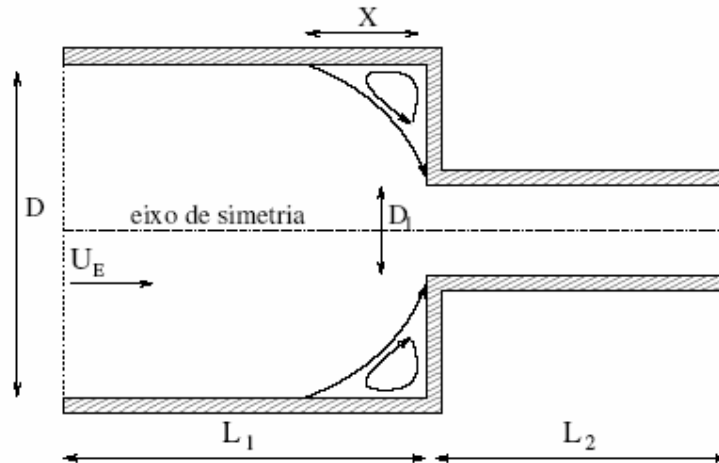


Figura 1: Configuração da contração plana

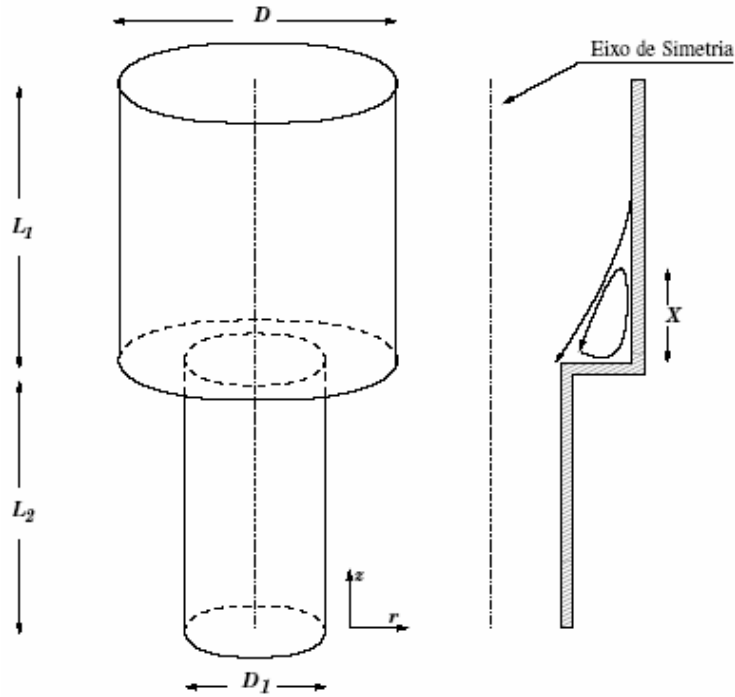


Figura 2: Configuração da contração axissimétrica

Nas duas geometrias aparecem vórtices situados no canto da contração e define-se o comprimento destes vórtices como sendo  $L_{vort} = 2X / D$ .

Um estudo detalhado das condições de contorno nas configurações plana e axissimétrica é tratado em GROSSI (2003), entretanto algumas informações são apresentadas a seguir.

### Contração plana :

Na entrada da cavidade da contração plana 4:1, as componentes da velocidade foram impostas como:

$$u(y) = -6y^2 + 6yD \quad e \quad v(y) = 0.$$

Para as componentes  $\mathbf{S}$  do tensor de tensão extra adotou-se a estratégia de CROCHET et. al. (1984) e MOMPEAN et. al. (1997), isto é:

$$S^{xx} = 0, \quad S^{xy} = 0 \quad e \quad S^{yy} = 0$$

Na saída de fluido a condição homogênea de Newmann foi imposta para as componentes da velocidade e para as componentes do tensor de tensão extra  $\mathbf{S}$  MOMPEAN et. al. (1997):

$$\frac{\partial u_n}{\partial n} = \frac{\partial u_t}{\partial n} = 0 \quad e \quad \frac{\partial S^{xx}}{\partial n} = \frac{\partial S^{xy}}{\partial n} = \frac{\partial S^{yy}}{\partial n} = 0$$

Na fronteira rígida a condição de não escorregamento para velocidade foi imposta ( $\mathbf{u} = 0$ ). Para o tensor  $\mathbf{S}$  foi feita uma mudança de variável e o tratamento detalhado das equações pode ser encontrado em GROSSI (2003) e GROSSI et. al. (2006).

### Contração axisimétrica :

Na entrada da cavidade da contração plana 4:1 as componentes da velocidade foram impostas como:

$$u_z(r) = \frac{2V_0(R^2 - r^2)}{R^2} \quad \text{e} \quad u_r(r) = 0.$$

Para as componentes  $\mathbf{S}$  do tensor de tensão extra adotou-se:

$$S^{rr} = S^{\theta\theta} = S^{zz} = S^{rz} = 0$$

Na saída de fluido, similar ao caso planar:

$$\frac{\partial u_n}{\partial n} = \frac{\partial u_t}{\partial n} = 0 \quad \text{e} \quad \frac{\partial S^{rr}}{\partial n} = \frac{\partial S^{\theta\theta}}{\partial n} = \frac{\partial S^{zz}}{\partial n} = \frac{\partial S^{rz}}{\partial n} = 0$$

Na fronteira rígida a condição de não escorregamento para velocidade foi imposta, isto é,  $\mathbf{u} = 0$ . Para o tensor  $\mathbf{S}$  foi feita uma mudança de variável, o tratamento detalhado das equações podem ser encontrados em GROSSI (2003).

Na fronteira de simetria (onde  $r = 0$ ) a condição de escorregamento para as componentes da velocidade foi aplicado. Esta condição garante que o fluido não transponha a fronteira de simetria, então:  $u(r, z) = 0$  e  $\frac{\partial v}{\partial r} = 0$ . Para o tensor  $\mathbf{S}$  o tratamento detalhado das equações podem ser encontrados em GROSSI (2003).

## 4. Método de Solução

Para resolver as equações de quantidade de movimento e as equações das componentes do tensor tensão extra empregou-se o seguinte procedimento. Suponha que em um dado tempo  $t_n$ , o campo de velocidade  $\mathbf{u}(\mathbf{x}, t_n)$  e o tensor extra tensão  $\mathbf{S}(\mathbf{x}, t_n)$  são conhecidos e as condições de contorno para o vetor velocidade  $\mathbf{u}$ , o tensor tensão extra  $\mathbf{S}$  e a pressão são fornecidos. Para calcular o campo de velocidade, o tensor tensão extra e a pressão no tempo atualizado  $t_{n+1} = t_n + \delta t$ , onde  $\delta t$  é o tamanho do passo no tempo, procede-se da seguinte forma:

1. Calcula-se o campo de velocidade intermediário,  $\tilde{\mathbf{u}}(\mathbf{x}, t_{n+1})$  da equação (7) aplicando as condições corretas de fronteira para  $\mathbf{u}(\mathbf{x}, t_n)$ .
2. Conhecido  $\tilde{\mathbf{u}}(\mathbf{x}, t_{n+1})$ , resolve-se a equação de Poisson  $\nabla^2 \psi(\mathbf{x}, t_{n+1}) = \nabla \cdot \tilde{\mathbf{u}}(\mathbf{x}, t_{n+1})$ , sujeita as suas condições fronteira (mais detalhes são encontrados em TOME et. al. (2007)).
3. Calcula-se o campo de velocidade final dado por  $\mathbf{u}(\mathbf{x}, t_{n+1}) = \tilde{\mathbf{u}}(\mathbf{x}, t_{n+1}) - \nabla \psi(\mathbf{x}, t_{n+1})$
4. Calcula-se o campo de pressão  $p(\mathbf{x}, t_{n+1}) = \tilde{p}(\mathbf{x}, t_{n+1}) + \frac{\psi(\mathbf{x}, t_{n+1})}{\delta t}$ .
5. Atualizam-se as componentes não newtonianas do tensor tensão extra na fronteira rígida e no eixo de simetria (GROSSI (2003)).
6. Calculam-se as componentes do tensor tensão extra da equação (8) empregando a forma conservativa para os termos convectivos (mais detalhes são encontrados em GROSSI (2003) e TOME et. al. (2007)).

## 5. Discretização das Equações Governantes

As equações que regem o escoamento de um fluido Oldroyd-B foram resolvidas aplicando a técnica de diferenças finitas. Uma malha diferenciada bidimensional com espaçamentos  $\delta x$  por  $\delta y$  foi utilizada na configuração plana e  $\delta r$  por  $\delta z$  na configuração axissimétrica. As componentes da velocidade localizam-se nas laterais e as componentes do tensor de tensões extra no centro da célula computacional (GROSSI (2003)).

Para resolver as equações das componentes da velocidade  $\tilde{u}$  e  $\tilde{v}$ , as derivadas temporais foram aproximadas explicitamente pelo método de Euler enquanto que os termos viscosos, o gradiente de pressão e as derivadas espaciais das componentes do tensor tensão extra foram aproximados por diferenças centrais (GROSSI (2003)). Para os termos convectivos utilizou-se o esquema "upwind" de alta ordem, VONOS (Variable-Order-Non-Oscillatory-Scheme), cujos detalhes podem ser encontrados em FERREIRA et. al. (2002).

## 6. Resultados Numéricos

Para mostrar que a técnica apresentada nesse trabalho simula escoamentos viscoelásticos de um fluido Oldroyd-B para o problema da contração 4:1 tanto no caso planar quanto no caso axissimétrico, serão apresentados a seguir os resultados de simulações numéricas para dois tipos de escoamentos: escoamentos inerciais com

$Re = 1.0$  e escoamentos a baixos números de Reynolds, onde  $Re = 0.01$ . Para os dois problemas foram utilizados diferentes números de Weissenberg. A razão  $\lambda_2 / \lambda_1 = 1/9$  foi mantida fixa em todas as simulações.

Os seguintes dados foram utilizados nas simulações envolvendo as configurações plana e axissimétrica, cujo domínio geométrico foi apresentado nas figuras 1 e 2.

Contração plana:

- comprimento das cavidades:  $L_1 = 15$  cm e  $L_2 = 15$  cm.
- diâmetros das cavidades:  $D = 8$  cm e  $D_1 = 2$  cm.
- velocidade na entrada da cavidade de maior diâmetro:  $U_E = 0.25$  cm/s.
- malha com  $80 \times 300$  células, (sendo  $\delta x = \delta y = 0.1$  cm)
- tolerância para resolver a equação de Poisson  $EPS = 10^{-10}$ .
- a força da gravidade é desprezada.
- parâmetros de escala:  $L = 1$ ,  $U = 1$ ,  $\nu_0 = 1$  para  $Re = 1$  e  $\nu_0 = 100$  para  $Re = 0.01$ .

Contração axissimétrica:

- comprimento dos tubos:  $L_1 = 15$  cm e  $L_2 = 15$  cm.
- raio dos tubos:  $R = 4$  cm e  $R_1 = 1$  cm.
- velocidade na entrada do tubo:  $U_E = 0.04$  cm/s (mesma vazão que no caso planar).
- considerando a metade do domínio computacional tem-se a malha com  $40 \times 300$  células, (sendo  $\delta x = \delta y = 0.1$  cm)
- tolerância para resolver a equação de Poisson  $EPS = 10^{-10}$ .
- a força da gravidade é desprezada.
- parâmetros de escala:  $L = 1$ ,  $U = 0.6$ ,  $\nu_0 = 0.6$  para  $Re = 1$  e  $\nu_0 = 60$  para  $Re = 0.01$ .

As figuras 3 e 4 apresentam as linhas de corrente provenientes da simulação do escoamento de um fluido Oldroyd-B com  $Re = 1.0$  através de uma contração plana e axissimétrica após atingir o estado estacionário.

Nota-se que para as simulações da contração plana, o tamanho do vórtice decresce com aumento do número de Weissenberg, enquanto que para a contração axissimétrica o tamanho do vórtice cresce com aumento de Weissenberg. Um comportamento análogo é descrito no artigo experimental de WALTER et. al. (2001).

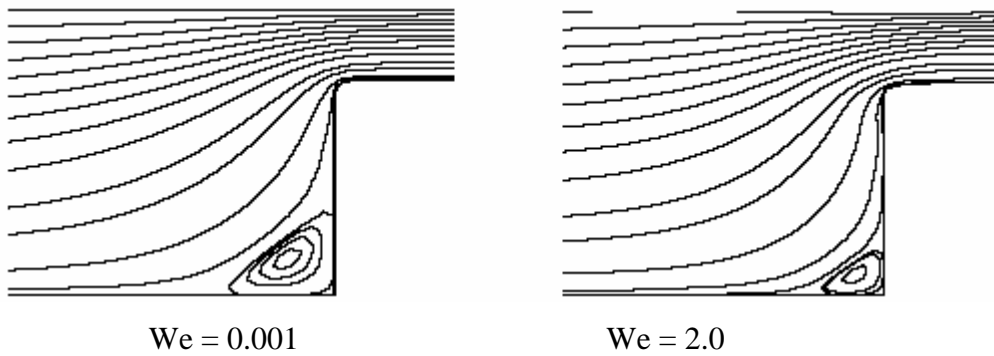


Figura 3: Simulação da contração plana. Visualização das funções de corrente no estado estacionário ( $Re = 1$ )

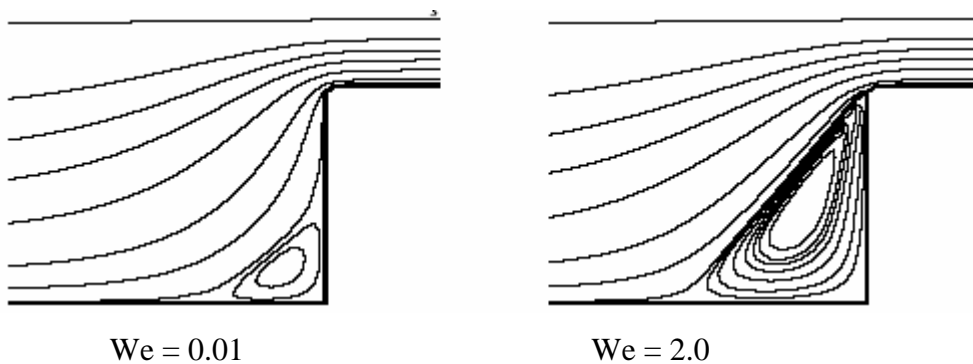


Figura 4: Simulação da contração axissimétrica. Visualização das funções de corrente no estado estacionário ( $Re = 1$ )

A figura 5 apresenta uma comparação entre os resultados obtidos nesse trabalho para a contração plana em que  $Re=1.0$ , com os resultados dos trabalhos de CAREW et. al. (1993), MATALLAH et. al. (1998), PHILLIPS et. al. (2002), e SATO et. al. (1994). No gráfico estão relacionados o número de Weissenberg e o tamanho do vórtice calculado por  $L_{vort} = 2X / D$ , onde X e D foram apresentadas nas figura 1 e 2.

Nota-se na figura 5, que os resultados obtidos pelo código GENSMAC estão em melhor concordância com os resultados obtidos por SATO et. al. (1994) e PHILLIPS et. al. (2002), especialmente para números de Weissenberg menores e iguais que um ( $We \leq 1.0$ ). Para  $We = 1.5$  e  $We = 2.0$  os resultados obtidos situam-se entre os valores de SATO et. al. (1994) e PHILLIPS et. al. (2002). Pode-se observar que os resultados de CAREW et. al. (1993), são maiores que os obtidos pelos outros autores. MATALLAH et. al. (1998) foram os únicos a conseguir convergência para  $We = 4.0$ .



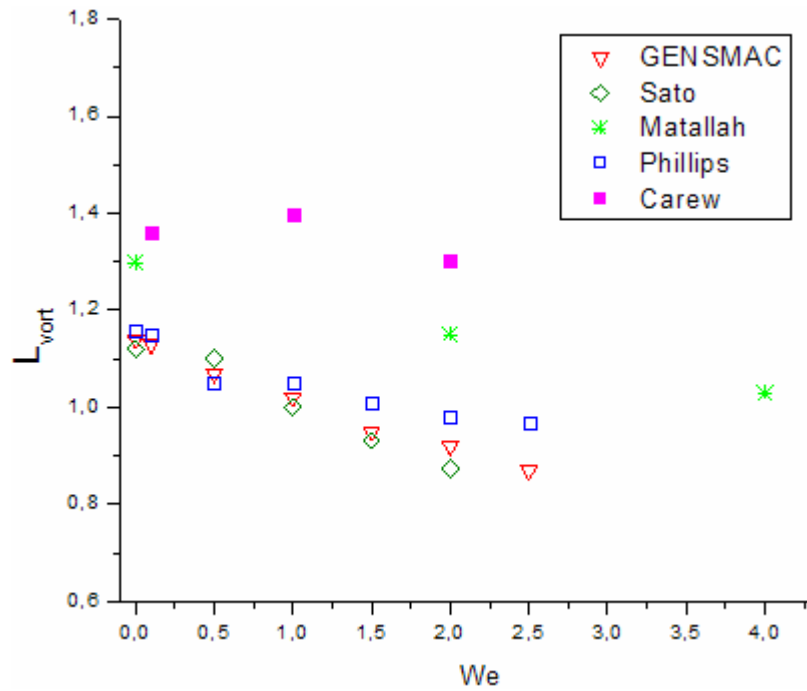


Figura 5: Comprimento do vórtice para  $Re=1$ . (caso da contração plana)

A figura 6 apresenta os resultados obtidos para a contração axissimétrica com  $Re=1.0$ , e estes são comparados com os valores obtidos no artigo de PHILLIPS et. al. (2002). Observa-se que os resultados obtidos pelo código GENSMAC estão mais próximos dos valores obtidos por PHILLIPS et. al. (2002) para  $We \leq 1.0$ .

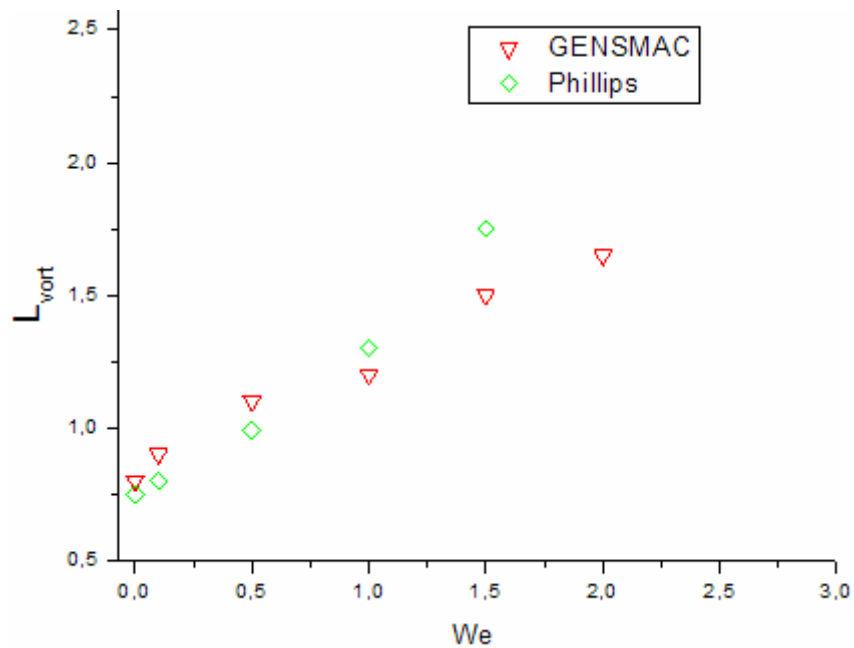


Figura 6: Comprimento do vórtice para  $Re=1$ . (caso da contração axissimétrica)

As figuras 7 e 8 apresentam a visualização dos vórtices resultantes das simulações da contração 4:1 plana e axissimétrica após atingir o estado estacionário. Os resultados das simulações da contração plana, Figura 7, mostram que o vórtice diminui de tamanho com aumento do número de Weissenberg. Para  $We = 2.0$  observa-se a formação de um "lip vortice" (pequeno vórtice na quina da contração). Muitos autores mencionam o aparecimento do "lip vortice" devido o aumento de  $We$  e sugerem que o aparecimento do "lip vortice" esteja relacionado com o método numérico empregado na resolução do problema e com o refinamento da malha.

Na figura 8 é possível visualizar os vórtices resultantes da simulação do escoamento na contração axissimétrica. Observa-se que o crescimento do vórtice com aumento dos valores de  $We$  tornou-se bem mais significativo que no caso inercial. Este comportamento também é observado no trabalho de PHILLIPS et. al. (2002).

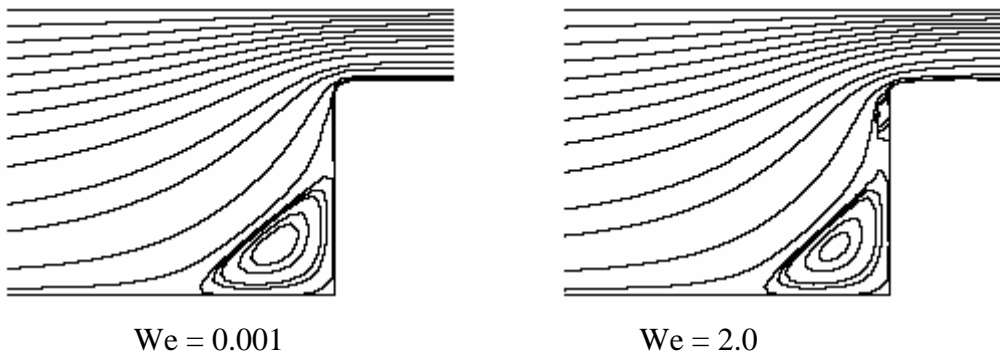


Figura 7: Simulação da contração plana. Visualização das funções de corrente no estado estacionário ( $Re = 0.01$ )

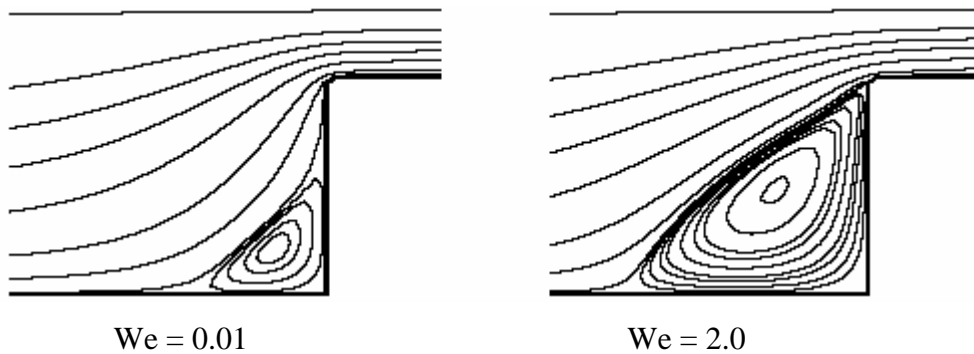


Figura 8: Simulação da contração axissimétrica. Visualização das funções de corrente no estado estacionário ( $Re = 0.01$ )

A figura 9 compara os resultados obtidos nesse trabalho para o caso planar com os apresentados nos trabalhos de PHILLIPS et. al. (1999), SATO et. al. (1994) e ABOUBACAR et. al. (2002). Nesses artigos, utilizou-se a razão  $\lambda_2 / \lambda_1 = 1/9$  e  $Re = 0$ , exceto no trabalho de SATO et. al. (1994), os quais utilizaram  $Re = 0.01$ . Na figura 9, pode-se observar que valores do  $L_{vort}$  obtidos pelo código GENSMAC estão bem próximos dos valores de PHILLIPS et. al. (1999). Para  $We = 0.5$  os resultados obtidos pelo GENSMAC praticamente coincidem com os de SATO et. al. (1994). Para  $We = 1.0$  encontra-se entre os resultados obtidos por PHILLIPS et. al. (1999) e ABOUBACAR et. al. (2002).

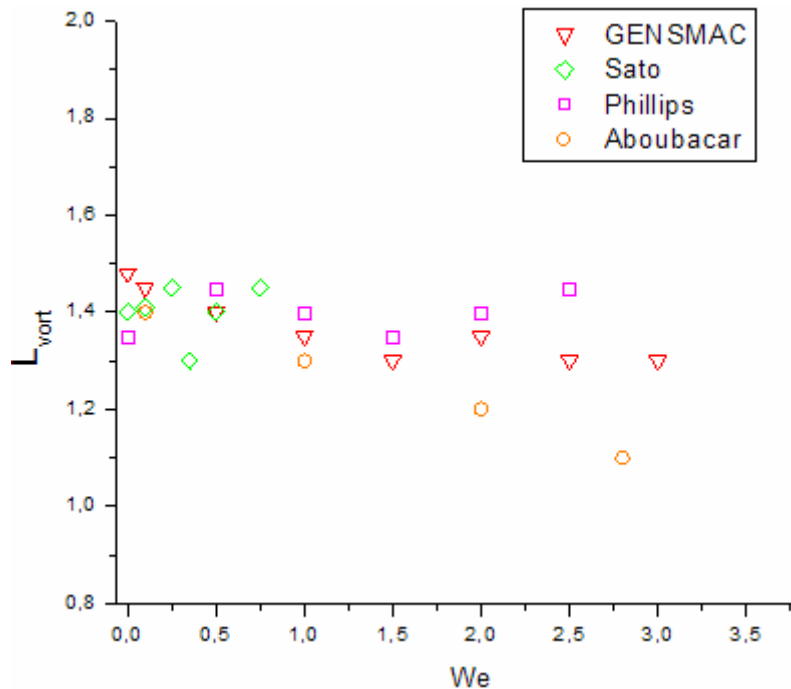


Figura 9: Comprimento do vórtice para  $Re=0.1$ . (caso da contração plana).

Na figura 10 são comparados os resultados de PHILLIPS et. al. (2002), ABOUBACAR et. al. (2002) com os obtidos neste trabalho, utilizando o código GENSMAC na simulação no caso da contração axissimétrica a baixos valores do número de Reynolds. PHILLIPS et. al. (2002) e ABOUBACAR et. al. (2002) utilizaram  $Re = 0$  em seus trabalhos. As simulações apresentadas aqui foram feitas utilizando  $Re = 0.01$ . Observa-se que os resultados obtidos pelo código GENSMAC estão mais próximos dos valores de PHILLIPS et. al. (2002) para  $We \leq 0.5$ . No caso axissimétrico ABOUBACAR et. al. (2002) obteve solução para valores de  $We \leq 0.85$ .

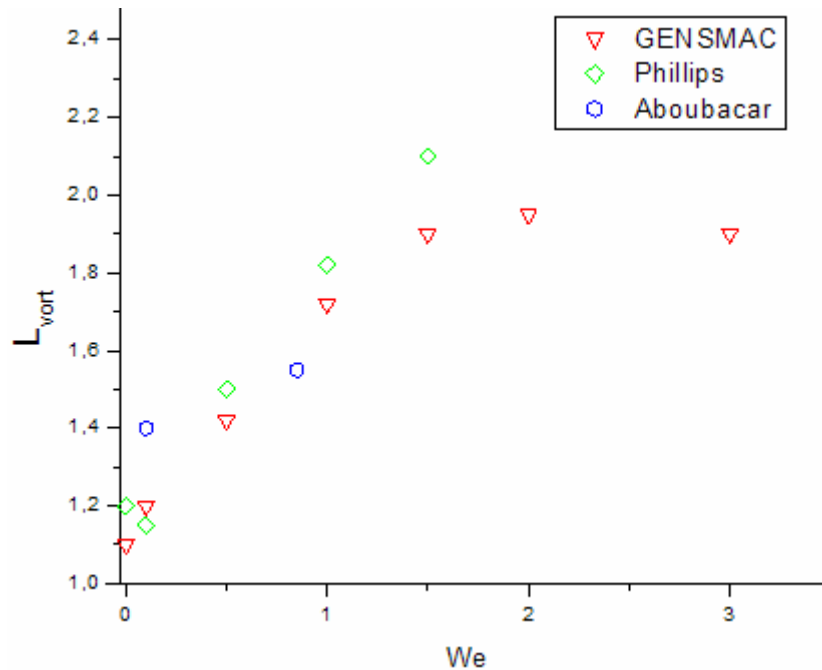


Figura 10: Comprimento do vórtice para  $Re=0.01$ . (caso da contração axissimétrica)

## 7. Conclusão

O presente trabalho descreve uma técnica numérica que permite analisar escoamentos viscoelásticos de fluidos Oldroyd-B através de uma contração brusca de razão 4:1. Os resultados numéricos obtidos para escoamentos inerciais e para escoamentos a baixos números de Reynolds "*creeping flows*" foram apresentados e comparados com os existentes na literatura tais como MATAALLAH et. al. (1998), PHILLIPS et. al. (2002), CAREW et. al. (1993), SATO et. al. (1994) e ABOUBACAR et. al. (2002).

Os resultados obtidos comprovam a forte influência do número de Weissenberg no tamanho do vórtice do canto. Constatou-se que o crescimento do vórtice no escoamento de um fluido viscoelástico em uma configuração axissimétrica é bem diferente quando o mesmo fluido é submetido a uma configuração plana. No caso da contração plana, o vórtice diminui seu tamanho com o aumento do valor de  $We$  enquanto que no caso axissimétrico, o vórtice apresenta um aumento no tamanho e intensidade. Os resultados obtidos apresentam boa concordância com os resultados apresentados na literatura. Futuramente, pretende-se aplicar esta técnica para estudar o aparecimento dos "*lip vórtice*" e estende-la ao tratamento de escoamentos viscoelásticos utilizando outros modelos constitutivos.

## 8. Referências Bibliográficas:

ABOUBACAR, M., MATAALLAH, H., TAMADDON, J. and WEBSTER, M. F.: **Numerical prediction of extensional flows in contraction geometries: hybrid finite volume/element method**, J. Non-Newtonian Fluid Mech., vol. 104, (2002), 125-164.

CAREW, E. O. A., TOWNSED, P., and WEBSTER, M. F.: **A Taylor-Petrov-Galerkin algorithm for viscoelastic flow**, J. Non-Newtonian Fluid Mech., vol. 50, (1993), 253-287.

CROCHET, M. J., and WALTERS, K.: **Numerical Simulation of Non-Newtonian Flow**, Elsevier, New York, 1984.

FERREIRA, V. G. ; M. F. Tomé ; FORTUNA, A. O. ; A. Castelo ; N. Mangiavacchi ; Cuminato, J.A. and MCKEE, S.: **A Comparative Study of High-Order Upwind Schemes on Axisymmetric Free Surface Flows**. Nonlinear Dynamics, Chaos, Control, and Their Applications to Engineering Sciences, Applications of Nonlinear Phenomena. Rio Claro: , 2002, vol. 06, p. 21-30.

GROSSI, L.: **Desenvolvimento de métodos numéricos para a simulação de escoamentos não-newtonianos e viscoelásticos com superfícies livres**, Tese de doutorado ICMC-USP, São Carlos, 2003.

GROSSI, L. ; TOMÉ, M. F. ; CASTELO, A. ; MANGIAVACCHI, N. ; CUMINATO, J. A. ; FERREIRA, V. G. and MCKEE, S.: **A Numerical Method for Solving Viscoelastic Contraction Flows**. In: Mohan C. Joshi; Amiya K. Pani; Sanjeev V. Sabnis. (Org.). Industrial Mathematics. New Delhi: Narosa.Publishing House, 2006, p. 205-220.

MARCHAL, J.M., and CROCHET, M. J.: **A new mixed finite element for calculating viscoelastic flow**, J. Non-Newtonian Fluid Mech., vol. 26, (1987), 77-114.

MATAALLAH, H., TOWNSEND, P. and WEBSTER, M. F.: **Recovery and stress-splitting schemes for viscoelastic flows**, J. Non-Newtonian Fluid Mech., vol. 75, (1998), 139-166.

MOMPEAN G. and DEVILLE, E M.: **Unsteady finite volume of Oldroyd-B fluid through a three-dimensional planar contraction**, J. Non-Newtonian Fluid Mech., vol. 72, (1997), 253-279.

PHILLIPS, T.N. and WILLIAMS, A.J.: **Comparioson of creeping and inertial flow of an Oldroyd B fluid though planar and axisymmetric contractions**, J. Non-Newtonian Fluid Mech., vol. 108, (2002), 25-47.

PHILLIPS, T.N. and WILLIAMS, A.J.: **Viscoelastic flow though a planar contraction using a semi-Lagrangian finite volume method**, J. Non-Newtonian Fluid Mech., vol. 87, (1999),215-246.

SATO, T. and RICHARDSON, S. M.: **Explicit numerical simulation of time-dependent viscoelastic flows problems by a finite element/finite volume method**, J. Non-Newtonian Fluid Mech., vol. 51, (1994), 249-275.

TOMÉ, M. F. ; GROSSI, L. ; CASTELO, A. ; CUMINATO, J. A. ; MCKEE, S. and WALTERS, K.: **Die-Swell, Splashing Drops and a Numerical Technique for Solving the Oldroyd-B Model for Axisymmetric Free Surface Flows**. Journal of Non-Newtonian Fluid Mechanics, v. 141, p. 148-166, 2007.

TOME, M. F. and MCKEE, S.: **GENSMAC: A Computational Marker-and-Cell Method for Free Surface Flows in General Domains**, J. of Computational Physics, vol. 110, (1994), 171-186.

TOME, M. F., MANGIAVACCHI, N., CUMINATO, J. A., CASTELO, A. and MCKEE, S.: **Marker-and-Cell technique for simulating unsteady viscoelastic free surface flows**, Notas do ICMC, vol. 63, (2001).

WALTER, K. and WEBSTER, M. F.: **The distinctive CFD challenges of computacional rheology**, ECOMAS, (2001).

# O PROBLEMA DA COBERTURA SOBRE ANÉIS E CONJUNTOS LIVRE DE PRODUTO

Irene Naomi Nakaoka, Otávio José Neto Tinoco Neves dos Santos.

Universidade Estadual de Maringá - UEM, innakaoka@uem.br;

Universidade Estadual de Mato Grosso do Sul - UEMS, ojneto@uems.br

**Resumo:** Dado um anel comutativo com unidade  $A$ , denotamos por  $c(A, n, R)$  como sendo a mínima cardinalidade dos conjuntos  $H$  de  $A^n$  que satisfazem a seguinte propriedade: qualquer elemento de  $A^n$  difere no máximo  $R$  coordenadas de um múltiplo de um elemento de  $H$ . Neste artigo os números  $c(\mathbb{Z}_m, 2, 0)$  são determinados, para todo inteiro  $m \geq 2$ . Também relacionamos as os números  $c(A, n, R)$  com conjuntos livre de produto, desta relação obtemos um limitante superior para  $c(\mathbb{Z}_q, 3, 1)$ , quando  $q$  é potência de um primo  $p$ . Além disso, obtemos um limitante para  $c(\mathbb{Z}_q \times A, 3, 1)$  em função do primo  $p$  e do número  $c(A, 3, 1)$ .

**Palavras-chave:** problema da cobertura, conjuntos livre de produto, ações de grupo

## 1. Introdução

Seja  $V_q^n$  o conjunto de todas as  $n$ -uplas de elementos de um conjunto finito de cardinalidade  $q$ , estas  $n$ -uplas são também chamadas de *vetores* ou *palavras*. A *distância de Hamming* entre duas palavras  $u = (u_1, \dots, u_n)$  e  $v = (v_1, \dots, v_n)$  de  $V_q^n$  é definida por  $d(u, v) = |\{i : u_i \neq v_i\}|$ . A *bola* de centro  $v$  e raio  $R$  corresponde ao conjunto  $B(v, R) = \{u \in V_q^n : d(u, v) \leq R\}$ . Um subconjunto  $C \subset V_q^n$  é uma  *$R$ -cobertura* de  $V_q^n$ , se para todo  $v \in V_q^n$  existe uma palavra  $c \in C$  tal que  $v \in B(c, R)$ . A mínima cardinalidade de uma  *$R$ -cobertura* de  $V_q^n$  é denotada por  $K_q(n, R)$ . Os números  $K_q(n, R)$  são bastante conhecidos, o problema extremal de determinar valores ou pelo menos limitantes para estes números é conhecido como problema da cobertura e é um importante tópico da teoria combinatória dos códigos. (Veja [1], para mais detalhes).

Recentemente, em [3], foi proposta a seguinte variante para o problema da cobertura. Seja  $\mathbb{F}_q$  um corpo finito com  $q$  elementos, definimos  $c(\mathbb{F}_q, n, R)$  como a mínima cardinalidade de um subconjunto  $H \subset \mathbb{F}_q^n$  com a propriedade que, para toda palavra  $v \in \mathbb{F}_q^n$ , existem  $\alpha \in \mathbb{F}_q$  e  $h \in H$  tais que  $v \in B(\alpha h, R)$ . Os melhores limitantes conhecidos para  $c(\mathbb{F}_q, n, R)$  são

$$\left\lceil \frac{q+1}{2} \right\rceil \leq c(\mathbb{F}_q, 3, 1) \leq \frac{3(q+4)}{4}, \quad (1)$$

que foram obtidos em [3] e [4] usando, entre outras ferramentas, ações de grupos e uma relação entre estas coberturas e conjuntos livre de produto.

Neste trabalho, estentemos a variante  $c$ , da função  $k$  como segue. Estaremos considerando sempre  $A$  como sendo um anel comutativo com unidade 1. Para um elemento  $h \in A^n$ , seja

$$E(h, R) = \bigcup_{\alpha \in A} B(\alpha h, R).$$

Se  $S \subset A^n$ , dizemos que o subconjunto  $H$  de  $A^n$  é uma  $R$ -cobertura curta de  $S$  quando  $S \subset \cup_{h \in H} E(h, R)$ . Definimos  $c(A, n, R)$  como sendo a mínima cardinalidade de uma  $R$ -cobertura curta de  $A^n$ .

O problema clássico da cobertura é um conceito puramente combinatório, depende apenas do número de elementos do espaço. Entretanto os números  $c(A, n, R)$  não dependem apenas da cardinalidade do anel  $A$ , mas também da sua estrutura algébrica, isto é, os números  $c(A_1, n, R)$  e  $c(A_2, n, R)$  podem ser diferentes mesmo quando  $|A_1| = |A_2|$ . Por exemplo, em [5] foi mostrado que  $c(\mathbb{Z}_4, 3, 1) = 4$  e  $c(\mathbb{Z}_2 \times \mathbb{Z}_2, 3, 1) = 3$ . No entanto, na seção 2, mostraremos que existe uma relação entre os números  $c(A, n, R)$  e uma classe especial dos números  $K_q(n, R)$  e como consequência desta relação obtemos um limitante inferior para  $c(A, 3, 1)$ .

Claramente,  $K_m(2, 0) = m^2$  para todo inteiro positivo  $m$ , mas determinar os valores de  $c(A, n, 0)$  não é assim tão trivial. Na seção 3 calculamos os valores de  $c(\mathbb{Z}_m, 2, 0)$ , para todo inteiro positivo  $m$ .



Na seção 4, motivados por [4], relacionamos o problema da cobertura curta sobre anéis com ações de grupos e conjuntos livre de produtos. Dessa relação obtemos, na seção 5, um limitante superior para  $c(\mathbb{Z}_q, 3, 1)$ , quando  $q$  é potência de primo. Também na seção 5, obtemos um limitante superior para  $c(\mathbb{Z}_p \times A, 3, 1)$  em função do primo  $p$  e do número  $c(A, 3, 1)$ . Como consequência destes resultados obtemos um limitante superior de  $c(\mathbb{Z}_p^n, 3, 1)$ , para todo primo  $p$ , e o valor exato de  $c(\mathbb{Z}_{2p}, 3, 1)$ , quando  $p$  é um primo ímpar.

## 2. Limitante Inferior

A seguir apresentaremos uma relação entre as funções  $K$  e  $c$ . Como é usual, a cardinalidade de um conjunto  $X$  será denotada por  $|X|$ . Uma  $R$ -cobertura curta de cardinalidade mínima será chamada de *R-cobertura curta mínima*.

**Proposição 1** [5].  $K_{|A|}(n, R) \leq (|A| - 1) c(A, n, R) + 1$ .

Kalbfleish e Stanton mostraram em [2] que  $K_q(3, 1) = \lceil q^2/2 \rceil$ . Aplicando a proposição anterior obtemos.

**Corolário 2** [5]. *Para todo anel comutativo com unidade  $A$ ,  $c(A, 3, 1) \geq \lceil (|A| + 1)/2 \rceil$*

O limitante inferior dado no Corolário 2 é bastante útil, por exemplo, dele segue que  $c(\mathbb{Z}_{2p}, 3, 1) \geq p + 1$ .

## 3. O Caso Bidimensional

Claramente  $c(A, 2, 1) = 1$ . Para o caso  $R = 0$ , observamos que uma 0-cobertura curta  $H$  de  $\mathbb{F}_q^2$  satisfaz  $\mathbb{F}_q H = \mathbb{F}_q^2$ , logo  $c(\mathbb{F}_q, 2, 0) = q + 1$ . Contudo, uma 0-cobertura curta mínima de  $\mathbb{Z}_q^2$  tem mais elementos, como mostra o próximo resultado.

**Proposição 3** [5]. *Se  $q = p^n$  é a potência de um primo, então  $c(\mathbb{Z}_q, 2, 0) = p^n + p^{n-1}$ .*

*Demonstração:* Seja  $H$  uma 1-cobertura curta mínima de  $\mathbb{Z}_q^2$  e sejam  $H_1 = \{(x_1, x_2) \in H : x_1 \in U(\mathbb{Z}_q)\}$ ,  $H_2 = \{(x_1, x_2) \in H \setminus H_1 : x_2 \in U(\mathbb{Z}_q)\}$  e  $H_3 = \{(x_1, x_2) \in H : x_1, x_2 \notin U(\mathbb{Z}_q)\}$ . Observamos que para todos  $a, b \in U(\mathbb{Z}_q)$  e  $1 \leq r, s < n$ ,  $(ap^r, bp^s) = p^r(a, bp^{s-r})$  se  $s \geq r$  e  $(ap^r, bp^s) = p^s(ap^{r-s}, b)$ , caso contrário. Além disso, se  $h, h' \in \mathbb{Z}_q^2$  e  $h = \alpha h'$  para algum  $\alpha \in \mathbb{Z}_q$ , então  $E(h, 0) \subset E(h', 0)$ . Portanto, podemos supor que  $H_3 = \emptyset$ . Vamos mostrar que  $|H_1| \geq p^n$  e  $|H_2| \geq p^{n-1}$ . Se  $|H_1| < p^n$ , existe  $x \in \mathbb{Z}_q$  tal que  $h_1^{-1}h_2 \neq x$ , para todos  $(h_1, h_2) \in H_1$ . Como  $H$  é uma 1-cobertura curta de  $\mathbb{Z}_q^2$ , existem  $(h_1, h_2) \in H_1$  e  $\lambda \in \mathbb{Z}_q$  tais que  $\lambda(h_1, h_2) = (1, x)$ , que implica  $h_1^{-1}h_2 = x$ ; esta contradição mostra que  $|H_1| \geq p^n$ . Analogamente obtemos que  $|H_2| \geq p^{n-1}$ . Consequentemente,  $c(\mathbb{Z}_q, 2, 0) \geq p^n + p^{n-1}$ . Para mostrarmos que vale a igualdade, provaremos que  $K = \{(1, x) \in \mathbb{Z}_q^2 : x \in \mathbb{Z}_q\} \cup \{(x, 1) \in \mathbb{Z}_q^2 : x \notin U(\mathbb{Z}_q)\}$  é uma 1-cobertura curta de  $\mathbb{Z}_q^2$ . Para  $u = (x, y) \in \mathbb{Z}_q^2$  temos as seguintes possibilidades:

- $x \in U(\mathbb{Z}_q)$ . Então  $k = (1, x^{-1}y) \in K$  e  $u = xk$ .
- $x \notin U(\mathbb{Z}_q)$  e  $y \in U(\mathbb{Z}_q)$ . Neste caso,  $k = (xy^{-1}, 1) \in K$  e  $u = yk$ .
- $x, y \notin U(\mathbb{Z}_q)$ . Podemos escrever  $x = ap^i$  e  $y = bp^j$  com  $a, b \in U(\mathbb{Z}_q)$  e  $i, j \geq 1$ . Se  $i > j$ , então  $k = (ab^{-1}p^{i-j}, 1) \in K$  e  $u = yk$ . Se  $i \leq j$ ,  $k = (1, a^{-1}bp^{j-i}) \in K$  e  $u = xk$ .

Portanto, em qualquer caso,  $u \in \cup_{k \in K} E(k, 0)$ , completando a demonstração.

**Teorema 4** [5]. *Se  $m = p_1^{r_1} p_2^{r_2} \cdots p_n^{r_n}$ , aqui  $p_1, p_2, \dots, p_n$  são primos distintos e  $r_i > 0$  para  $1 \leq i \leq n$ , então*

$$c(\mathbb{Z}_m, 2, 0) = \prod_{i=1}^n (p_i^{r_i} + p_i^{r_i-1}).$$

*Demonstração.* Segue como consequência da Proposição 3 e da Proposição 11 [5].

#### 4. Cobertura Curta, Ações e Conjuntos Livre de Produto

Como é usual,  $U(A)$  é o grupo das unidades de  $A$ . Considere o produto direto  $S_n \times U(A)$ , onde  $S_n$  denota o grupo simétrico de grau  $n$ . Para  $\varphi \in S_n$ ,  $\lambda \in U(A)$ , e

$v = (v_1, v_2, \dots, v_n) \in A^n$ , definimos a ação de  $S_n \times U(A)$  sobre o conjunto  $A^n$  colocando

$$v^{(\varphi, \lambda)} = (\lambda v_{(1)\varphi^{-1}}, \lambda v_{(2)\varphi^{-1}}, \dots, \lambda v_{(n)\varphi^{-1}}). \quad (2)$$

Para  $\varphi \in S_n$  e  $\lambda \in A$  escrevemos  $\varphi\lambda$  ao invés de  $(\varphi, \lambda)$ ; portanto  $v^{\varphi\lambda} = v^{(\varphi, \lambda)}$ ,  $v^\varphi = v^{(\varphi, 1)}$ , e  $v^\lambda = v^{(1, \lambda)}$ . Quando um grupo  $G$  age sobre um conjunto  $X$ , denotamos a órbita de  $v$  por  $v^G = \{v^g : g \in G\}$ . Se  $Y$  é um subconjunto de  $X$  e  $H$  é um subconjunto de  $G$ ,  $Y^H$  denota o conjunto  $\{y^h : y \in Y \text{ and } h \in H\}$ .

O próximo resultado estabelece um método para decidir quando um candidato  $H$  é realmente uma  $R$ -cobertura curta de  $A^n$

**Teorema 5** [5]. *Seja  $B$  um subconjunto de  $A^n$  invariante pela ação de  $S_n \times U(A)$ . Se  $L$  é um subconjunto de  $A^n$  invariante pela ação de  $S_n$  tal que cada órbita da ação de  $S_n \times U(A)$  sobre  $B$  contém um elemento  $v$  tal que  $d(v, \alpha h) \leq R$ , para algum  $\alpha \in A$  e algum  $h \in H$ , então  $L$  é uma  $R$ -cobertura curta de  $B$ .*

A seguir apresentaremos uma conexão, baseada numa construção feita em [4], entre coberturas curtas e conjuntos livre de produto. Esta conexão será útil para obtermos um limitante superior para  $c(\mathbb{Z}_q, 3, 1)$ , quando  $q$  é potência de primo. Nossa abordagem consiste em determinar uma 1-cobertura curta para o seguinte subconjunto de  $A^3$

$$\mathcal{U}(A) = \{(x, y, z) : x, y, z \in U(A) \text{ and } x, y, z \text{ são dois a dois distintos}\}.$$

O próximo passo, é encontrar uma 1-cobertura curta para o complemento de  $\mathcal{U}(A)$  em  $A^3$ .

Começamos observando que  $\mathcal{U}(A)$  é invariante pela ação do produto direto  $S_3 \times U(A)$ . Consequentemente, esta ação sobre  $A^3$  induz de maneira natural uma ação sobre  $\mathcal{U}(A)$ .

Um subconjunto  $X$  de um grupo multiplicativo  $G$  é chamado de livre de produto em  $G$ , se  $ab \notin X$  para todos  $a, b \in X$ .

Para uma família  $\mathcal{P} = \{(x_1, y_1), \dots, (x_k, y_k)\}$  de elementos de  $G \times G$ , colocamos

$$\Delta_{\mathcal{P}} = \{1\} \cup \bigcup_{i=1}^k \Delta_{(x_i, y_i)}$$

onde  $\Delta_{(x,y)} = \{x, y, x^{-1}, y^{-1}, xy^{-1}, yx^{-1}\}$ . Observamos que  $\Delta_{\mathcal{P}}$  é um conjunto inverso, isto é,  $\Delta_{\mathcal{P}} = (\Delta_{\mathcal{P}})^{-1} = \{x^{-1} \mid x \in \Delta_{\mathcal{P}}\}$ .

**Teorema 6** [5]. *Seja  $\mathcal{P} = \{(x_1, x_2), \dots, (x_k, y_k)\}$  uma coleção de pares de elementos de  $U(A)$ . Se o complemento de  $\Delta_{\mathcal{P}}$  em  $U(A)$  é livre de produto, então o conjunto*

$$H = \bigcup_{i=1}^k (1, x_i, y_i)^{S_3}$$

é uma 1-cobertura curta de  $\mathcal{U}(A)$ .

## 5. Limitante Superior para $c(\mathbb{Z}_m, 3, 1)$

Aplicaremos o Teorema 6 para obtermos uma 1-cobertura curta de  $\mathcal{U}(\mathbb{Z}_q)$ , quando  $q$  é potência de primo. Denotamos o complemento de  $S$  em  $U(\mathbb{Z}_q)$  por  $\bar{S}$ .

**Teorema 7** [5]. *Seja  $q = p^r$  a potência de um primo  $p$ , com  $q > 9$ .*

(i) *Se  $p$  é ímpar, então existe uma 1-cobertura curta de  $\mathcal{U}(\mathbb{Z}_q)$  com pelo menos  $9n/16 + 9/4$  elementos, aqui  $n = p^{r-1}(p-1)$*

(ii) *Se  $p = 2$  e  $r \geq 5$ , então existe uma 1-cobertura curta de  $\mathcal{U}(\mathbb{Z}_q)$  com  $9 \cdot 2^{r-5}$*

*Idéia da demonstração.* Para o caso  $p = 2$  e  $r \geq 5$ , observamos que o grupo multiplicativo  $U(\mathbb{Z}_q) = \{1, 3, 5, \dots, q-1\}$  é gerado por  $-1$  e  $5$  e isomorfo ao grupo  $\mathbb{Z}_2 \times \mathbb{Z}_{2^{r-2}}$ , como mostra o Teorema 5.44 [6]. Além disso, a ordem de  $5$  em  $U(\mathbb{Z}_q)$  é  $m = 2^{r-2}$ . Seja  $K$  o subgrupo de  $U(\mathbb{Z}_q)$  gerado por  $5$ . Usando a expansão binomial de  $5^i = (1+4)^i$  obtemos

$$K = \{1, 5, 5^2, \dots, 5^{m-1}\} = \{x \in U(\mathbb{Z}_q) \mid x \equiv 1 \pmod{4}\}.$$

É fácil verificar que  $\overline{K} = \{x \in U(\mathbb{Z}_q) \mid x \equiv 3 \pmod{4}\}$  é livre de produto em  $U(\mathbb{Z}_q)$ . Agora, basta determinar uma família  $\mathcal{P}$  tal que  $\overline{\Delta_{\mathcal{P}}} \subset \overline{K}$ . Se  $r = 5$  definimos  $\mathcal{P} = \{(5, 5^3), (5^4, 5^4)\}$  e se  $r > 5$ ,

$$\begin{aligned} \mathcal{P} &= \{(5, 5^{m/2-1}), (5^3, 5^{m/2-3}), \dots, (5^{m/4-1}, 5^{m/2-(m/4-1)})\} \\ &\cup \{(5^4, 5^8), (5^{12}, 5^{16}), \dots, (5^{4(m/8-1)}, 5^{4(m/8)})\}. \end{aligned}$$

Se  $r = 5$ , temos que  $\Delta_{\mathcal{P}} = K$ . Se  $r > 5$ , não é difícil checar que  $\{1, 5, 5^2, \dots, 5^{m/2}\} \subset \Delta_{\mathcal{P}}$  e como  $\Delta_{\mathcal{P}}$  é um conjunto inverso, segue que  $K \subset \Delta_{\mathcal{P}}$ . Assim, em qualquer um dos casos, temos que  $\overline{\Delta_{\mathcal{P}}} \subset \overline{K}$ , logo  $\overline{\Delta_{\mathcal{P}}}$  é livre de produto em  $U(\mathbb{Z}_q)$ . Portanto,  $\mathcal{P}$  induz uma 1-cobertura curta em  $\mathcal{U}(\mathbb{Z}_q)$ , contendo  $9 \cdot 2^{r-5}$  elementos.

Agora, suponhamos que  $p$  é um primo ímpar. Observamos que  $U(\mathbb{Z}_q)$  é um grupo cíclico de ordem  $n = p^{r-1}(p-1)$ . Seja  $\xi$  um gerador deste grupo. Como o conjunto  $\{\xi^1, \xi^3, \dots, \xi^{n-1}\}$  é livre de produto em  $U(\mathbb{Z}_q)$  é suficiente encontrarmos uma coleção  $\mathcal{P}$  tal que

$$\{1, \xi^2, \xi^4, \dots, \xi^{n-2}\} \subset \Delta_{\mathcal{P}}. \quad (3)$$

Dado um inteiro par  $k$ , seja

$$\mathcal{A}(k) = \{(\xi^2, \xi^{4k-2}), (\xi^6, \xi^{4k-6}), \dots, (\xi^{2k-2}, \xi^{2k+2})\}.$$

Se  $k/2$  é par, colocamos

$$\mathcal{B}(k) = \{(\xi^8, \xi^{16}), (\xi^{24}, \xi^{32}), \dots, (\xi^{8(k/2-1)}, \xi^{8(k/2)})\},$$

caso contrário, fazemos

$$\mathcal{B}(k) = \{(\xi^8, \xi^{16}), (\xi^{24}, \xi^{32}), \dots, (\xi^{8(k/2-2)}, \xi^{8(k/2-1)})\} \cup \{(\xi^{8(k/2)}, \xi^{8(k/2)})\}.$$

Seja  $k$  o quociente da divisão de  $n$  por 8. Como  $n$  é par, temos que  $n = 8k + r$ , com  $r = 0, 2, 4$  ou 6. Definimos

$$\mathcal{P} = \mathcal{A}(k) \cup \mathcal{B}(k), \text{ se } r = 0 \text{ ou } 2 \text{ e } k \text{ é par};$$

$$\mathcal{P} = \mathcal{A}(k-1) \cup \mathcal{B}(k-1) \cup \{(\xi^{4k-2}, \xi^{4k})\}, \text{ se } r = 0 \text{ ou } 2 \text{ e } k \text{ é ímpar};$$

$\mathcal{P} = \mathcal{A}(k) \cup \mathcal{B}(k) \cup \{(\xi^{4k+2}, \xi^{4k+2})\}$ , se  $r = 4$  ou  $6$  e  $k$  é par

$\mathcal{P} = \mathcal{A}(k-1) \cup \mathcal{B}(k-1) \cup \{(\xi^{4k-2}, \xi^{4k}), (\xi^{4k+2}, \xi^{4k+2})\}$ , se  $r = 4$  ou  $6$  e  $k$  é ímpar.

Não é difícil mostrar que, em qualquer caso,  $\mathcal{P}$  satisfaz a inclusão (3) e que, portanto,  $\overline{\Delta_{\mathcal{P}}}$  é livre de produto. Assim, pelo Teorema 6, em qualquer caso,  $\mathcal{P}$  induz uma 1-cobertura curta de  $\mathcal{U}(\mathbb{Z}_q)$ . Uma simples contagem mostra que cada cobertura contém no máximo  $9n/16 + 9/4$  elementos.

A seguir, completaremos a 1-cobertura curta de  $\mathcal{U}(\mathbb{Z}_q)$ , obtida no teorema anterior, a uma 1-cobertura curta de  $\mathbb{Z}_q^3$  obtendo um limitante superior para  $c(\mathbb{Z}_q, 3, 1)$ .

**Teorema 8** [5]. *Seja  $q = p^r$  a potência de um primo  $p$ .*

(i) *Se  $p \neq 2$  e  $q > 9$ , então  $c(\mathbb{Z}_q, 3, 1) \leq \frac{9}{16}q \left(1 + \frac{23}{9p}\right) + \frac{13}{4}$ .*

(ii) *Se  $p = 2$  e  $r \geq 5$ , então  $c(\mathbb{Z}_q, 3, 1) \leq \frac{41}{32}q + 1$ .*

*Idéia da demonstração.* Definimos o conjunto  $K$ , colocando  $K = \emptyset$ , se  $r = 1$ , caso contrário, colocamos  $K = \{(1, ip, 1), (ip, 1, ip) : i = 1, 2, \dots, p^{n-1} - 1\}$ . Se  $H$  é uma 1-cobertura curta de  $\mathcal{U}(\mathbb{Z}_q)$ , então  $L = H \cup (1, 1, 0)^{S_3} \cup K$  é uma 1-cobertura curta de  $\mathbb{Z}_q^3$  e o resultado segue usando o teorema anterior e uma simples contagem.

Segue do Teorema 8 que se  $p$  é um primo ímpar, então  $c(\mathbb{Z}_p, 3, 1) \leq 9p/16 + 75/16$ . Para  $p \geq 23$ , este resultado melhora o limitante superior dado em (1).

**Teorema 9** [5]. *Se  $A$  é um anel comutativo com unidade e  $p$  é um primo, então*

$$c(A \times \mathbb{Z}_p, 3, 1) \leq (p + 1)c(A, 3, 1).$$

*Idéia da demonstração.* Seja  $H$  uma 1-cobertura curta mínima de  $A^3$ . Para cada  $h = (a, b, c)$  em  $H$  consideremos o conjunto

$$W_h = T_h \cup \{(a, 1), (b, l), (c, -(l+1)) : l = 1, 2, \dots, p-2\},$$

onde  $T_h$  denota o conjunto

$$\{((a, 0), (b, 1), (c, -1)), ((a, -1), (b, 0), (c, 1)), ((a, 1), (b, -1), (c, 0))\}.$$

Notemos que  $|W_h| = p + 1$ . O resultado é obtido mostrando que  $K = \cup_{h \in H} W_h$  é uma 1-cobertura curta  $(A \times \mathbb{Z}_p)^3$ .

**Corolário 10** [5]. *Se  $p$  é primo um primo ímpar, então  $c(\mathbb{Z}_{2p}, 3, 1) = p + 1$ .*

*Demonstração.* Segue do Teorema 9 e da Proposição 2.

**Corolário 11** [5]. *Seja  $p$  um primo. Para todo  $n \geq 1$ , temos que*

(i)  $c(\mathbb{Z}_p^n, 3, 1) \leq (p + 1)^{n-1}c(\mathbb{Z}_p, 3, 1)$ , se  $p$  é ímpar.

(ii)  $c(\mathbb{Z}_2^n, 3, 1) \leq 3^{n-1}$ .

*Demonstração.* Segue do Teorema 9, usando o fato que  $\{(1, 1, 1)\}$  é uma 1-cobertura curta de  $\mathbb{Z}_2^3$ .

## Agradecimento

Agradecemos a E. L. Monte Carmelo pelas sugestões durante o desenvolvimento desta pesquisa.

## Referências Bibliográficas:

- [1] COHEN, G., HONKALA, I., LITSYN, S. and LOBSTEIN, A. **Covering Codes**, North-Holland Publishing, Amsterdam, (1997).
- [2] KALBFLEISH, J.G. and STANTON, R.G., **A combinatorial problem in matching**, J. London Math. Soc., VOL. 44, (1969), 60-64.
- [3] MONTE CARMELO, E.L., NAKAOKA, I.N. and GERÔNIMO, J.R., **A covering problem on finite spaces and rook domains**, Int. J. Appl. Math, vol. 20, (2007), 875-886.

- [4] MONTE CARMELO, E.L and NAKAOKA, I.N., **Short coverings in tridimensional spaces arising from sum-free sets**, Europ. J. Combin., vol. 29, (2008), 227-233.
  
- [5] NAKAOKA, I.N. and SANTOS, O.J.N.T.N, **A covering problem over ring and product-free sets**, Int. J. Appl. Math, aceito para publicação.
  
- [6] ROTMAN, J.J., **An Introduction to the Theory of Groups**, Graduate Texts in Mathematics, vol. 148, Springer-Verlag, New York, (1995).



# COMPARAÇÃO COMPUTACIONAL ENTRE OS MÉTODOS NUMÉRICOS DE EULER, HEUN E RUNGE-KUTTA PARA EQUAÇÕES DIFERENCIAIS ORDINÁRIAS DE 1ª ORDEM.

Marcos Freitas de Moraes.

Universidade Estadual do Oeste do Paraná - UNIOESTE, profmarcosmoraes@gmail.com.

**Palavras-chave:** Métodos Numéricos, Runge-Kutta, Equações Diferenciais.

## Resumo:

A aplicação dos Métodos Numéricos para aproximar soluções de Equações Diferenciais Ordinárias (EDO) de 1ª. ordem vem sendo cada dia mais utilizado em problemas de valor inicial (PVI), pois facilitam os cálculos e não é necessário calcular as suas derivadas. Nesta pesquisa optou-se por uma apresentação e comparação entre três métodos numéricos aplicáveis a uma EDO. de 1ª. ordem: o Método de Euler, o Método de Heun e o Método de Runge-Kutta de 4ª. ordem. Primeiramente foram efetuados cálculos manuais e em seguida desenvolvido programas no "software" livre Scilab 3.0, um para cada método para efetuar a comparação computacional. Por reduzir o tamanho do passo pela metade e o erro se reduzir a um fator de cerca de  $2^4 = 16$ , o método de Runge-Kutta de 4ª. ordem se comportou como o mais preciso e eficiente nos resultados, tanto nos cálculos manuais quanto nos computacionais, mas isso não quer dizer que em casos específicos os outros dois métodos não sejam mais práticos e eficazes.

## Introdução:

Uma equação diferencial ordinária (EDO) linear não tem necessariamente uma solução, e mesmo que possua, nem sempre podemos achá-la explícita ou implicitamente. Se existe uma solução de uma equação diferencial, ela representa um lugar de pontos (pontos unidos por uma curva suave) no plano cartesiano. A razão mais forte de introduzirmos métodos numéricos para aproximar soluções de problemas de valor inicial

(PVI) é a dificuldade de se encontrar, analiticamente, as soluções da equação. Em muitos casos, a teoria nos garante existência e unicidade de solução, mas não sabemos qual é a expressão analítica desta solução.

Os métodos comparados nesta pesquisa baseiam-se em:

$$\text{Dado Problema de Valor Inicial (PVI): } y' = (f, y) \rightarrow y(x_0) = y_0. \quad (1)$$

## Objetivo

O objetivo desta pesquisa é apresentar e comparar os métodos numéricos de Euler, Heun e Runge-Kutta, aproveitar as suas qualidades que permitem determinar valores  $y_i \approx y(t_i)$  por meio de relações de recorrência de modo que o valor de  $y_{i+1}$  venha expresso em função de  $y_i, y_{i-1}, \dots, y_0$ , sendo  $y_0 = y(a) = \alpha$ , mostrar a eficiência de cada método e definir qual deles possui melhor resultado computacionalmente.

## Materiais e Métodos

### O Método de Euler

O método, também chamado de método das tangentes, constitui-se numa técnica para aproximar soluções de equações diferenciais. Suponha que se queira aproximar a solução do problema de valor inicial (1).

Supondo um valor uniforme (constante) de  $h$ , pode-se obter uma sucessão de pontos  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$  que devem estar próximos dos pontos  $(x_1, y(x_1)), (x_2, y(x_2)), \dots, (x_n, y(x_n))$  da curva solução  $y$ .

Utilizando  $(x_1, y_1)$  pode-se obter  $y_2$ , que é a ordenada do ponto solução em uma nova tangente.

$$\text{Tem-se } \frac{y_2 - y_1}{x_2 - x_1} = y'_1 \text{ ou } y_2 = y_1 + h \cdot y'_1.$$

$$\text{Genericamente tem-se: } y_{n+1} = y_n + h \cdot y'_n = y_n + h \cdot f(x_n, y_n).$$

### O Método de Heun ou Euler Melhorado

Suponhamos novamente que se queira aproximar a solução do problema de valor inicial (1). Dado os coeficientes angulares das retas  $\alpha_0 = f(x_0, y_0)$  e  $\alpha_1 = f(x_1, y_1^*)$ , tomamos a média desses coeficientes, obtendo outro coeficiente angular ao longo da reta

por  $(x_0, y_0)$  ao invés de uma nova inclinação com ordenada  $y_1^*$ , como visto no método de Euler anterior, até atingir  $x_1$ . Isto constitui numa melhora em relação ao método de Euler, sendo mais preciso.

$$\text{Então: } y_{n+1}^* = y_n + h \cdot f(x_n, y_n),$$

$$y_{n+1} = y_n + h \cdot \frac{f(x_n, y_n) + f(x_n + h, y_n + h \cdot f(x_n, y_n))}{2}$$

e como podemos chamar

$$K_1 = h \cdot f(x_n, y_n);$$

$$K_2 = h \cdot f(x_n + h/2, y_n + K_1/2).$$

### Método de Runge-Kutta de 4ª. Ordem

A fórmula clássica do método de Runge-Kutta de 4ª. ordem, o qual foi utilizado na pesquisa para a solução do PVI (1) é:

$$y_{n+1} = y_n + \frac{1}{6}(K_1 + 2K_2 + 2K_3 + K_4)$$

onde

$$K_1 = h \cdot f(x_n, y_n)$$

$$K_2 = h \cdot f(x_n + h/2, y_n + K_1/2)$$

$$K_3 = h \cdot f(x_n + h/2, y_n + K_2/2)$$

$$K_4 = h \cdot f(x_n + h, y_n + K_3)$$

Podemos interpretar a soma  $(K_1 + 2K_2 + 2K_3 + K_4)/6$  como um coeficiente angular médio. Nota-se que  $K_1$  é o coeficiente angular no extremo esquerdo do intervalo,  $K_2$  é o coeficiente angular no ponto médio utilizando-se a fórmula de Euler para ir de  $x_n$  a  $x_n + h/2$ ,  $K_3$  é uma segunda aproximação ao coeficiente angular no ponto médio e, finalmente,  $K_4$  é o coeficiente angular em  $x_{n+h}$  usando-se a fórmula de Euler e o coeficiente angular  $K_3$  para ir de  $x_n$  a  $x_{n+h}$ .

### Discussão

#### Comparação entre os Métodos

A seguir é apresentado um exemplo contendo uma tabela com os respectivos resultados encontrados para os três métodos comparados com o valor correto de um PVI. Nesta comparação optou-se por iniciar no ponto  $x_1 = 1,00$  até  $x_5 = 1,50$ .

Exemplo 1: Aplicando os métodos com  $h = 0,1$ , obter uma aproximação de  $y(1,5)$  para a solução de:  $y' = 2xy$ , com  $y(0) = 1$ . Solução efetuada pelo programa livre **Scilab 3.0**.

Tabela 1: Comparação entre os Métodos Numéricos

$x_n$	Método de Euler	Método de Heun	Runge-Kutta	Valor Correto
1,00	1,0000	1,0000	1,0000	1,0000
1,10	1,2000	1,2320	1,2337	1,2337
1,20	1,4640	1,5479	1,5527	1,5527
1,30	1,8154	1,9832	1,9937	1,9937
1,40	2,2874	2,5908	2,6116	2,6117
1,50	2,9278	3,4509	3,4902	3,4904

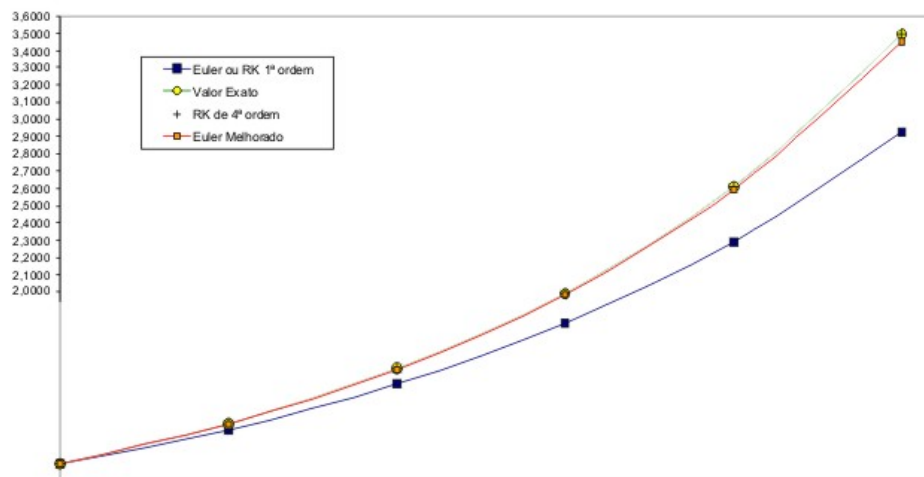


Figura 1: Comparação entre os Métodos

Ao observar-se os resultados obtidos, constata-se que surge no primeiro passo um erro absoluto de 0,0337 para o Método de Euler, 0,0017 para o Método de Heun e de

0,0000 para o Runge-Kutta. É claro que para quatro casas decimais, caso houvesse mais casas já haveria uma margem de erro também para o Runge-Kutta de 4ª ordem. Para os passos seguintes, o Runge-Kutta de 4ª ordem se mantém estável, enquanto os outros métodos tendem a aumentar o erro. Somente no passo 1,50 que haverá diferença no resultado do Runge-Kutta.

### **Conclusão**

Por esta pesquisa pode-se concluir claramente que o método de Runge-Kutta de 4ª. ordem aplicado a um PVI como o apresentado em (1) é muito mais eficiente na aproximação que os outros dois métodos. Para uma melhor exatidão do método de Runge-Kutta, deve-se manter muitas casas decimais na solução numérica, para ver o efeito da redução de passo  $h$  à metade e o erro ser dividido por um fator de cerca de  $2^4 = 16$ . Ao aceitar os resultados de um método numérico, o usuário deve se preocupar sempre com a estabilidade do método. Isto tem grande importância, porque os números utilizados em uma calculadora ou em um computador não são exatos. Todos os métodos de passo único são estáveis, desde que se utilize um passo suficientemente pequeno.

### **Referências Bibliográficas:**

RUGGIERO, Márcia A. G.; LOPES, Vera Lúcia da R.: **Cálculo numérico: aspectos teóricos e computacionais.**, Makron Books, 2ª Ed., São Paulo, (1996).

MUNEN, M. A.; FOULIS, D. J.: **Cálculo**, LTC, Vol. 2, (1982).

BUTCHER, J. C.: **The Numerical Analysis of Ordinary Equation**, Chichester, (1987).

VALENÇA, M. R.: **Métodos Numéricos**, I.N.I.C., Braga, (1988).

# CODES SATISFYING THE CHAIN CONDITION OVER ROSENBLOOM-TSFASMAN SPACES<sup>1</sup>

Luciano Panek, Emerson Lazzarotto, Fernando Mucio Bando

Universidade Estadual do Oeste do Paraná - UNIOESTE - lucpanek@gmail.com;

lazzarotto@unioeste.br; bando@unioeste.br

## Abstract

Let  $M_{n \times m}(\mathbb{F}_q)$  be the linear space of all  $n \times m$  matrices over the finite field  $\mathbb{F}_q$ , equipped with the new weight introduced by Rosenbloom and Tsfasman. In this paper we extend the concept of generalized Wei weights for Rosenbloom-Tsfasman weights and show that all linear codes satisfy the chain condition in  $M_{n \times 1}(\mathbb{F}_q)$ .

**Key Words:** Rosenbloom-Tsfasman weight, generalized Hamming weights, chain condition.

## Introduction

Let  $M_{n \times m}(\mathbb{F}_q)$  be the linear space of all  $n \times m$  matrices over the finite field  $\mathbb{F}_q$ . In 1997 Rosenbloom and Tsfasman they equipped the space  $M_{n \times m}(\mathbb{F}_q)$  with new weight ([1]), called *Rosenbloom-Tsfasman weight*  $w_\rho$ , defined as follows: if  $(a_{ij}) \in M_{n \times m}(\mathbb{F}_q)$ , then

$$w_\rho((a_{ij})) = \sum_{j=1}^m \max \{i : a_{ij} \neq 0\},$$

assuming  $\max \emptyset = 0$ . Moreover, if  $n = 1$  the Rosenbloom-Tsfasman weight  $w_\rho$  over space  $M_{1 \times m}(\mathbb{F}_q)$  are the Hamming weight  $w_H$  of classical coding theory. The Rosenbloom-

---

<sup>1</sup>Accepted for publication in the Journal of Pure and Applied Mathematics.

Tsfasman weights constitute an important family of weights which can be applied to concrete communication systems (see [1], [2]).

Motivated by several applications in cryptography, Victor Wei introduced in 1991 the concept of generalized Hamming weights ([3]). We extend here the concept of generalized Wei weights to Rosenbloom-Tsfasman weights. If  $D$  is a linear subspace of the linear code  $C$  we write  $D \leq C$ . When  $D$  is a proper subspace of  $C$  we write  $D < C$ . The *generalized Rosenbloom-Tsfasman weight*  $\|\cdot\|_\rho$ , we write *generalized  $\rho$  weight*, of a  $r$ -dimensional subspace  $D \leq M_{n \times m}(\mathbb{F}_q)$  is defined as

$$\|D\|_\rho = \sum_{j=1}^m \max \{i : a_{ij} \neq 0, a_{ij} \text{ element of the } j\text{-th column of } (a_{ij}) \in D\}.$$

The  $r$ -th minimum  $\rho$  weight of a  $k$ -dimensional code  $C \leq M_{n \times m}(\mathbb{F}_q)$  is

$$d_r(C) = \min \left\{ \|D\|_\rho : D \leq C, \dim(D) = r \right\}.$$

A code  $C$  is called an  $[n \cdot m; k; d_1(C), \dots, d_k(C)]_q$ -code. The sequence  $(d_1(C), \dots, d_k(C))$  is called  $\rho$  weights hierarchy of  $C$ . If  $n = 1$ , then the  $r$ -th minimum  $\rho$  weight is usual  $r$ -th minimum Hamming weight of  $M_{1 \times m}(\mathbb{F}_q) \simeq \mathbb{F}_q^m$ .

As in [3], we have the monotonicity of the minimum  $\rho$  weights.

**Proposition 1** For any  $[n \cdot m; k; d_1(C), \dots, d_k(C)]_q$ -code  $C \leq M_{n \times m}(\mathbb{F}_q)$  we have that

$$1 \leq d_1(C) < d_2(C) < \dots < d_k(C) \leq n \cdot m.$$

**Proof.** We observed initially that  $d_{r-1}(C) \leq d_r(C)$ . In fact, let  $D_{r-1}$  and  $D_r$  subcodes of  $C$  with dimensions  $r-1$  and  $r$  respectively such that  $\|D_{r-1}\|_\rho = d_{r-1}(C)$  and  $\|D_r\|_\rho = d_r(C)$ . If  $\|D_{r-1}\|_\rho > \|D_r\|_\rho$ , then for any subcode  $D'_{r-1} < D_r$  of dimension  $r-1$  we have that  $\|D'_{r-1}\|_\rho \leq \|D_r\|_\rho < \|D_{r-1}\|_\rho = d_{r-1}(C)$ . But this contradicts the minimality of  $d_{r-1}(C)$ . We claim that the inequality  $d_{r-1}(C) \leq d_r(C)$  is strict. Let  $D$  a subcode of  $C$  with dimension  $r$  such that  $\|D\|_\rho = d_r(C)$ . Then there exists  $s$  such that

$$\{i : a_{is} \neq 0, a_{is} \text{ element of the } s\text{-th column of } (a_{ij}) \in D\} \neq \emptyset.$$

If  $i'$  is the maximal element of the above set, then  $D_{i'} := \{(x_{ij}) \in D : x_{i's} = 0\}$  is a subcode of  $C$  with dimension  $r - 1$  such that

$$d_{r-1}(C) \leq \|D_{i'}\|_\rho \leq \|D\|_\rho - 1 = d_r(C) - 1.$$

□

Since  $d_{r+1}(C) \geq d_r(C) + 1$  and  $d_k(C) \leq n \cdot m$  we immediately get the generalized *Singleton bound*:

**Corollary 1** For an  $[n \cdot m; k; d_1(C), \dots, d_k(C)]_q$ -code  $C \leq M_{n \times m}(\mathbb{F}_q)$ ,

$$r \leq d_r(C) \leq n \cdot m - k + r.$$

We investigated in this work the possibility of the existence of new codes satisfying the chain condition with the generalized  $\rho$  weights. A  $k$ -dimensional code  $C \leq M_{n \times m}(\mathbb{F}_q)$  satisfies the *chain condition* if there exists a sequence of linear subspaces

$$D_1 < D_2 < \dots < D_{k-1} < D_k = C,$$

with  $\|D_r\|_\rho = d_r(C)$  and  $\dim(D_r) = r$  for all  $r \in \{1, 2, \dots, k\}$ . In the case that  $n = 1$  ( $w_\rho = w_H$ ) the Hamming codes, dual Hamming codes, Reed-Muller codes for all orders, maximum-separable-distance codes and Golay codes satisfy the chain condition (see [4]).

In this work we will show that any code  $C$  over the Rosenbloom-Tsfasman space  $M_{n \times 1}(\mathbb{F}_q) \simeq \mathbb{F}_q^n$  satisfies the chain condition. Moreover, the sequence of linear subspaces  $D_1 < D_2 < \dots < D_{k-1} < D_k = C$  that achieve the minimum Rosenbloom-Tsfasman weights in  $M_{n \times 1}(\mathbb{F}_q)$  is unique. It follows that if  $\|D_r\|_\rho = d_r(C)$  for all  $r \in \{1, 2, \dots, k\}$ , then  $D_1 < D_2 < \dots < D_{k-1} < D_k = C$ .

### Codes Satisfying the Chain Condition

Initially we give an example of a code satisfying the chain condition with  $\rho$  weight that does not satisfy the chain condition with the usual Hamming weight.



**Example 1** The  $[9; 3; 3, 6, 9]_2$ -code

$$C = \text{span} \left\{ (111000000)^T, (000111100)^T, (000001111)^T \right\}$$

does not satisfy the chain condition with the generalized Hamming weights. Now with the  $\rho$  weight the  $[9; 3, 3, 7, 9]_2$ -code  $C$  satisfies the chain condition:

$$\text{span} \left\{ (111000000)^T \right\} \subset \text{span} \left\{ \begin{array}{l} (111000000)^T \\ (000111100)^T \end{array} \right\} \subset C.$$

Now we will show that every code  $C$  satisfies the chain condition in the Rosenbloom-Tsfasman space  $M_{n \times 1}(\mathbb{F}_q)$ .

**Theorem 1** Let  $C$  be a code in  $M_{n \times 1}(\mathbb{F}_q) \simeq \mathbb{F}_q^n$ , endowed with the  $\rho$  weight. Then  $C$  satisfies the chain condition.

**Proof.** We observed that  $1 \leq d_1(C) < d_2(C) < \dots < d_k(C) \leq n$ . Since

$$\|D\|_\rho = \max \{i : x_i \neq 0, (x_1, \dots, x_m) \in D\},$$

so that for every  $j \in \{1, 2, \dots, k\}$  there is  $v_j \in C$  such that  $w_\rho(v_j) = d_j(C)$ . The set  $\{v_1, v_2, \dots, v_k\}$  is clearly linearly independent and consequently

$$\dim(\text{span}\{v_1, v_2, \dots, v_j\}) = j$$

and

$$\text{span}\{v_1\} < \text{span}\{v_1, v_2\} < \dots < \text{span}\{v_1, v_2, \dots, v_k\} = C.$$

Since  $\|\text{span}\{v_1, v_2, \dots, v_j\}\|_\rho = d_j(C)$  for every  $j \in \{1, 2, \dots, k\}$ , we find that  $C$  satisfies the chain condition.  $\square$

There are not many the sequences of subspaces that achieve the generalized minimum poset-weights, as display the next result.

**Theorem 2** *If  $C \leq M_{n \times 1}(\mathbb{F}_q) \simeq \mathbb{F}_q^n$  satisfies the chain condition, then there is a unique sequence of subspaces that achieve the generalized minimum  $\rho$  weights.*

**Proof.** Let  $k = \dim(C)$ ,  $(d_1(C), d_2(C), \dots, d_k(C))$  the  $\rho$  weights hierarchy of  $C$  and  $\{e_1, e_2, \dots, e_n\}$  the canonical base of  $\mathbb{F}_q^n$ . Let  $D_1 \leq C$  an 1-dimensional subcode of  $C$  such that  $\|D_1\|_\rho = d_1(C)$ . We will prove that  $D_1$  is unique. In fact, let  $D'_1 \leq C$  be an 1-dimensional subcode of  $C$  such that  $\|D'_1\|_\rho = d_1(C)$  and  $D'_1 \cap D_1 = \{\mathbf{0}\}$ . Then there are  $u \in D_1$  and  $v \in D'_1$  such that

$$u = \alpha_1 e_1 + \dots + \alpha_{d_1(C)-1} e_{d_1(C)-1} + e_{d_1(C)},$$

$$v = \beta_1 e_1 + \dots + \beta_{d_1(C)-1} e_{d_1(C)-1} + e_{d_1(C)}$$

with  $\alpha_j \neq \beta_j$  for some  $j \in \{1, 2, \dots, d_1(C) - 1\}$ . If

$$l = \max \{j \in \{1, 2, \dots, d_1(C) - 1\} : \alpha_j \neq \beta_j\},$$

it follows that  $u + (q-1)v$  is a non zero vector of  $C$  such that  $\|u + (q-1)v\|_\rho = l < d_1(C)$ . But this contradicts the minimality condition of the 1-th minimum Rosenbloom-Tsfasman weight of the code  $C$ . We conclude that  $D_1$  is the unique subcode of  $C$  that achieves the 1-th minimum Rosenbloom-Tsfasman weight of  $C$ .

The result follows now by induction on  $\dim(D_r) = r$ . Let  $D_1 < D_2 < \dots < D_{t-1} < C$ , with  $t-1 < k$ , be the sequence of linear subspaces that achieve the  $r$ -th minimum Rosenbloom-Tsfasman weights of the code  $C$  with  $r \in \{1, 2, \dots, t-1\}$ . Suppose that  $D_t$  and  $D'_t$  are  $t$ -dimensional subcodes of  $C$  containing  $D_{t-1}$  such that  $D_t \neq D'_t$  and  $\|D_t\|_\rho = \|D'_t\|_\rho = d_t(C)$ . Then there exists  $w \in D_t$  and  $z \in D'_t$  such that

$$w = \gamma_1 e_1 + \dots + \gamma_{d_t(C)-1} e_{d_t(C)-1} + e_{d_t(C)},$$

$$z = \eta_1 e_1 + \dots + \eta_{d_t(C)-1} e_{d_t(C)-1} + e_{d_t(C)},$$

with  $\gamma_j \neq \eta_j$  for some  $j \in \{1, 2, \dots, d_t(C) - 1\}$ . If

$$m = \max \{j \in \{1, 2, \dots, d_t(C) - 1\} : \gamma_j \neq \eta_j\},$$

then  $x = w + (q - 1)z$  is a non zero vector of  $C$  such that  $\|x\|_\rho = m < d_t(C)$  and  $x \notin D_{t-1}$ . We denote by  $\text{span}X$  the linear subspace of  $\mathbb{F}_q^n$  spanned by the set  $X \subset \mathbb{F}_q^n$ , and then, for every linearly independent subset  $\{y_1, \dots, y_{t-1}\} \subset D_{t-1}$ , we find that  $\text{span}\{y_1, \dots, y_{t-1}, x\}$  is a  $t$ -dimensional subspace of  $C$  such that  $\|\text{span}\{y_1, \dots, y_{t-1}, x\}\|_\rho \leq d_t(C) - 1$ , contradicting the minimality of  $d_t(C)$ .

By induction, the sequence of linear subspaces  $D_1 < D_2 < \dots < D_{k-1} < C$  that achieve the  $r$ -th minimum Rosenbloom-Tsfasman weights of code  $C$  is unique.  $\square$

For each subset  $P' \subseteq \{(1, 1), (2, 1), \dots, (n, m)\}$  we denoted by  $[P']$  the subspace of  $M_{n \times m}(\mathbb{F}_q)$  generated by the base  $\{e_{rs}\}_{(r,s) \in P'}$ ,  $e_{rs}$  the canonical vector of  $M_{n \times m}(\mathbb{F}_q)$ . So, if  $P_i = \{(1, i), (2, i), \dots, (n, i)\}$  we have that  $M_{n \times m}(\mathbb{F}_q)$  is a direct sum

$$[P_1] \oplus \dots \oplus [P_m].$$

As  $[P_i]$  is isometric to the  $M_{n \times 1}(\mathbb{F}_q)$  for every  $i \in \{1, 2, \dots, m\}$ , the Theorems 1 and 2 they assure that:

**Corollary 2** *Let  $C$  be a linear code over the space  $M_{n \times m}(\mathbb{F}_q)$ , endowed with the  $\rho$  weight, such that  $C \leq [P_i]$  for some  $i \in \{1, 2, \dots, m\}$ . Then there exists only one sequence of linear subspaces  $D_1 < D_2 < \dots < D_{k-1} < C$  with the property that  $\|D_r\|_\rho = d_r(C)$  for all  $r \in \{1, 2, \dots, k\}$ . Consequently, if  $D_1, D_2, \dots, D_{k-1}, D_k = C$  is a sequence of subspaces of  $C$  such that  $\|D_j\|_\rho = d_j(C)$  and  $\dim(D_j) = j$  for all  $j \in \{1, 2, \dots, k\}$ , then  $D_1 < D_2 < \dots < D_{k-1} < C$ .*

Now we presented a lower bound for the number of codes satisfying the chain condition.

**Proposition 2** *The number of codes satisfying the chain condition in the Rosenbloom-Tsfasman space  $M_{n \times m}(\mathbb{F}_q)$  is larger than or equal to the*

$$\sum_{i=1}^m \sum_{j=1}^n \prod_{k=1}^j \frac{(q^n - q^{k-1})}{(q^j - q^{k-1})}.$$

**Proof.** If  $P_i = \{(1, i), (2, i), \dots, (n, i)\}$  we have that  $M_{n \times m}(\mathbb{F}_q)$  is a direct sum

$$[P_1] \oplus \dots \oplus [P_m]$$

where each  $[P_i]$  is isometric to the  $M_{n \times 1}(\mathbb{F}_q)$ . The Corollary 2 assures that all code  $C \leq [P_i]$  satisfies the chain condition. Therefore the number of  $j$ -dimensional codes of  $[P_i]$  satisfying the chain condition equals

$$\prod_{k=1}^j \frac{(q^n - q^{k-1})}{(q^j - q^{k-1})}$$

for each  $i \in \{1, 2, \dots, m\}$ . This completes the proof of the Proposition.  $\square$

## References

- [1] M.Y. Rosenbloom and M.A. Tsfasman, **Codes for the  $m$ -metric**, Probl. Inf. Transm. **33** (1997) 45-52.
- [2] M.M. Skriganov, **Coding theory and uniform distributions**, St. Petesburg Math. J. **13**, n. 2, pp. 301-337, 2002.
- [3] V.K. Wei, **Generalized Hamming Weights for Linear Code**, IEEE Trans. Inform. Theory **37**, n. 5, pp. 1412-1418, September 1991.
- [4] V.K. Wei and K. Yang, **On the Generalized Hamming Weights for Product Code**, IEEE Trans. Inform. Theory **39**, n. 5, pp. 1709-1713, September 1993.

# O CÓDIGO BINÁRIO DE HAMMING

Pedro Henrique Galvão dos Santos, Luciano Panek

Colégio Estadual Professor Flávio Warken

Universidade Estadual do Oeste do Paraná - UNIOESTE

Foz do Iguaçu

E-mail: lucpanek@gmail.com

## Resumo:

Neste trabalho estudamos os procedimentos de detecção e correção de erros baseados no código binário de Hamming de dimensão quatro. Para tanto apresetamos uma breve introdução sobre corpos finitos de cardinalidade prima e códigos lineares.

**Palavras-chave:** códigos lineares; distância de Hamming; códigos binários de Hamming

## Introdução

Em um sistema de comunicação (por exemplo, celular, internet ou transmissão de imagem via satélite) os dados são associados a sequências de bits e transmitidos em forma de ondas (com amplitude variada, dependendo do canal). As ondas ficam sujeitas aos diversos tipos de interferências físicas (como campos magnéticos) e no processo inverso (quando demodulamos os dados) existe a possibilidade de ocorrência de erros (troca de bits). A teoria dos códigos corretores de erros, teoria responsável pelos mecanismos matemáticos que recuperam (sobre certas condições) os dados transmitidos a partir dos dados recebidos (contendo erros), teve início em 1950 com Richard Hamming.

Em 1950 Richard Hamming apresentou a primeira família de códigos corretores de erros, os chamados *códigos binários de Hamming*. Os códigos binários de Hamming têm a capacidade de detectar dois erros e corrigir um erro. Neste trabalho (de iniciação

científica) descrevemos alguns dos mecanismos de detecção e correção de erros baseados nos códigos binários de Hamming. Todos os nossos argumentos serão construídos sobre o código binário de Hamming de dimensão quatro (a extensão para o caso genérico é bastante natural e pode ser encontrada em qualquer bom livro de teoria dos códigos (ver [HEFEZ])). Como veremos, a estrutura algébrica dos códigos binários de Hamming é a estrutura vetorial sobre um corpo com dois elementos. Por este motivo começamos nosso trabalho com uma breve introdução (somente com os resultados necessários para o desenvolvimento do texto) sobre corpos finitos de cardinalidade prima e espaços vetoriais. Nossas fontes são os livros [ALVES], [BOLDRINI] e [DOMINGUES]. Omitiremos as demonstrações de alguns resultados. Estas demonstrações podem ser encontradas nos materiais citados.

## Corpos Finitos

Seja  $K$  um conjunto não vazio munido com duas operações  $+$  e  $\cdot$  chamadas de *soma* e *produto* (ou adição e multiplicação) respectivamente. Diremos que  $(K, +, \cdot)$  é um *corpo* se: as operações de adição e multiplicação são associativas e comutativas; existe o elemento neutro em relação a adição (elemento nulo); todo elemento de  $K$  é simetrizável em relação a adição (existência do elemento oposto); existe o elemento neutro em relação ao produto (existência da unidade); todo elemento não nulo de  $K$  é simetrizável em relação a multiplicação (existência do elemento inverso); o produto é distributivo em relação à adição.

Um corpo  $(K, +, \cdot)$  é dito *finito* se  $K$  é finito. É conhecido da Teoria de Galois que se um corpo  $K$  é finito então sua cardinalidade é uma potência de primo. Mais ainda, para todo primo  $p$  e para todo inteiro positivo não nulo  $m$  existe um corpo finito  $K$  de cardinalidade  $p^m$ . Neste trabalho nos restringiremos aos corpos finitos de cardinalidade  $p$ .

Seja  $\mathbb{Z}$  o conjunto dos números inteiros e para cada inteiro  $m > 0$  considere a

relação de congruência módulo  $m$ :  $a \equiv b \pmod{m}$  se, e somente se,  $a - b = km$  para algum  $k \in \mathbb{Z}$ . Valem as seguintes propriedades:

**Proposição 1** *Sejam  $a, b, c, d \in \mathbb{Z}$ . Então:*

(i)  $a \equiv a \pmod{m}$ ;

(ii) se  $a \equiv b \pmod{m}$ , então  $b \equiv a \pmod{m}$ ;

(iii) se  $a \equiv b \pmod{m}$  e  $b \equiv c \pmod{m}$ , então  $a \equiv c \pmod{m}$ ;

(iv) se  $a \equiv b \pmod{m}$  e  $c \equiv d \pmod{m}$ , então  $(a + c) \equiv (b + d) \pmod{m}$ ;

(v) se  $a \equiv b \pmod{m}$  e  $c \equiv d \pmod{m}$ , então  $a \cdot c \equiv b \cdot d \pmod{m}$ .

Segue das propriedades (i), (ii) e (iii) que  $\equiv$  é uma relação de equivalência sobre  $\mathbb{Z}$  e portanto induz uma partição em  $\mathbb{Z}$ : seja

$$[a] = \{b \in \mathbb{Z} : a \equiv b \pmod{m}\}$$

a classe de equivalência de  $a$  segunda a relação de congruência módulo  $n$ ; então

$$\mathbb{Z} = [0] \cup [1] \cup \dots \cup [m - 1]$$

com  $[r] \cap [s] = \emptyset$  se  $r \neq s$ ,  $0 \leq r, s \leq m - 1$ .

As propriedades (iv) e (v) nos mostram que as operação

$$[a] + [b] = [a + b]$$

e

$$[a] \cdot [b] = [a \cdot b]$$

sobre  $\mathbb{Z}_m = \{[0], [1], \dots, [m - 1]\}$  não dependem dos representantes das classes de equivalência. O elemento neutro e a unidade de  $\mathbb{Z}_m$  são respectivamente as classes  $[0]$  e  $[1]$ . O conjunto  $(\mathbb{Z}_m, +, \cdot)$  satisfaz todas as propriedade de corpo exceto a propriedade

que todo elemento não nulo é inversível. É bem conhecido que  $[a] \in \mathbb{Z}_m$  é inversível se, e somente se,  $\text{mdc}(a, m) = 1$ . Segue deste fato que  $(\mathbb{Z}_m, +, \cdot)$  é um corpo se, e somente se,  $p$  é um número primo.

Daqui em diante omitiremos os colchetes quando nos referirmos aos elementos do conjunto  $\mathbb{Z}_m$ .

**Exemplo 1** Considere as tabelas de adição e produto de  $\mathbb{Z}_2$ ,  $\mathbb{Z}_3$  e  $\mathbb{Z}_4$  respectivamente:

+	0	1
0	0	1
1	1	0

·	0	1
0	0	0
1	0	1

+	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

·	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

+	0	1	2	3
0	0	1	2	3
1	1	2	3	0
2	2	3	0	1
3	3	0	1	2

·	0	1	2	3
0	0	0	0	0
1	0	1	2	3
2	0	2	0	1
3	0	3	1	2

$\mathbb{Z}_2$  e  $\mathbb{Z}_3$  são corpos finitos. Já  $\mathbb{Z}_4$  não é um corpo finito: 2 não é inversível.

## Espaços Vetoriais sobre Corpos Finitos

Restringiremos as clássicas definições da teoria de espaços vetoriais ao caso particular dos corpos finitos  $\mathbb{Z}_p$ .

**Definição 1** Seja  $V$  um conjunto não vazio munido com duas operações  $+$  :  $V \times V \rightarrow V$  e  $\cdot$  :  $\mathbb{Z}_p \times V \rightarrow V$  tal que:



1.  $(x + y) + z = x + (y + z)$  para todo  $x, y, z \in V$ ;
2.  $x + y = y + x$  para todo  $x, y \in V$ ;
3. Existe  $\mathbf{0} \in V$  tal que  $\mathbf{0} + x = x$  para todo  $x \in V$ ;
4. Para cada  $x \in V$  existe  $-x \in V$  tal que  $x + (-x) = \mathbf{0}$ ;
5.  $\alpha(x + y) = \alpha x + \alpha y$  para todo  $x, y \in V$  para todo  $\alpha \in \mathbb{Z}_p$ ;
6.  $(\alpha + \beta)x = \alpha x + \beta x$  para todo  $x \in V$  para todo  $\alpha, \beta \in \mathbb{Z}_p$ ;
7.  $\alpha(\beta x) = (\alpha\beta)x$  para todo  $x \in V$  para todo  $\alpha, \beta \in \mathbb{Z}_p$ ;
8.  $1x = x$  para todo  $x \in V$ ;

Então  $(V, +, \cdot)$  é dito um espaço vetorial sobre  $\mathbb{Z}_p$ .

Os elementos de um espaço vetorial  $(V, +, \cdot)$  são chamados de *vetores*. Já os elementos do corpo finito  $\mathbb{Z}_p$  são chamados de *escalares*. Os vetores  $x_1, \dots, x_n \in V$  são ditos *linearmente independentes* se

$$\alpha_1 x_1 + \dots + \alpha_n x_n = \mathbf{0}$$

implica  $\alpha_1 = \dots = \alpha_n = 0$ . Se existem  $\alpha_1, \dots, \alpha_n \in \mathbb{Z}_p$  não todos nulos tal que  $\sum \alpha_i x_i = \mathbf{0}$  diremos que  $x_1, \dots, x_n$  são *linearmente dependentes*. Um subconjunto não vazio  $U$  de um espaço vetorial  $V$  é dito um *subespaço vetorial* de  $V$  se  $x + y \in U$  para todo  $x, y \in U$  e  $\alpha x \in U$  para todo  $\alpha \in \mathbb{Z}_p$  e  $x \in U$ . O conjunto

$$[x_1, \dots, x_n] = \{y \in V : y = \alpha_1 x_1 + \dots + \alpha_n x_n \text{ com } \alpha_1, \dots, \alpha_n \in \mathbb{Z}_p\}$$

é um subespaço vetorial de  $V$  e é chamado de *subespaço gerado* por  $x_1, \dots, x_n$ . Se  $V = [x_1, \dots, x_n]$  e  $x_1, \dots, x_n$  são linearmente independentes então  $\{x_1, \dots, x_n\}$  é dito uma *base* de  $V$ . Nestas condições diremos que  $V$  é um espaço vetorial de *dimensão*

finita. O número de vetores da base é a chamada *dimensão* de  $V$  e denotada por  $\dim(V)$ . Assim, se  $\{x_1, \dots, x_n\}$  é uma base de  $V$  e  $y \in V$ , existem únicos  $\alpha_1, \dots, \alpha_n \in \mathbb{Z}_p$  tal que  $y = \alpha_1 x_1 + \dots + \alpha_n x_n$ . Os escalares  $\alpha_1, \dots, \alpha_n$  são as *coordenadas* de  $y$ .

**Exemplo 2** Seja  $\mathbb{Z}_p^n$  o conjunto de todas as  $n$ -uplas com entradas em  $\mathbb{Z}_p$ . O conjunto  $\mathbb{Z}_p^n$  munido com as operações

$$(\alpha_1, \dots, \alpha_n) + (\beta_1, \dots, \beta_n) = (\alpha_1 + \beta_1, \dots, \alpha_n + \beta_n)$$

e

$$\gamma(\alpha_1, \dots, \alpha_n) = (\gamma\alpha_1, \dots, \gamma\alpha_n), \gamma \in \mathbb{Z}_p,$$

tem a estrutura de espaço vetorial. Como

$$\{e_1 = (1, 0, \dots, 0), e_2 = (0, 1, \dots, 0), \dots, e_n = (0, 0, \dots, 1)\}$$

é uma base para  $\mathbb{Z}_p^n$ , concluímos que  $\dim(\mathbb{Z}_p^n) = n$ . Segue daí que a cardinalidade de  $\mathbb{Z}_p^n$  é  $p^n$ .

## Códigos Lineares

Nosso objetivo agora é construir um mecanismo eficiente capaz de detectar e corrigir erros. Daqui em diante  $\mathbb{Z}_p^n$  será o nosso espaço ambiente. Um subconjunto  $C$  de  $\mathbb{Z}_p^n$  será chamado de  $(n, |C|)$  *código*, sendo  $|X|$  a cardinalidade do conjunto finito  $X$ . Os elementos de um código serão as chamadas *palavras-código*.

**Definição 2** Dados  $x = (x_1, \dots, x_n)$  e  $y = (y_1, \dots, y_n)$  em  $\mathbb{Z}_p^n$ , definimos a *distância de Hamming*  $d(x, y)$  entre  $x$  e  $y$  como sendo o número de coordenadas distintas entre  $x$  e  $y$ :

$$d(x, y) = |\{i : x_i \neq y_i\}|.$$

A distância de Hamming é uma métrica em  $\mathbb{Z}_p^n$ :  $d(x, y) \geq 0$ ;  $d(x, y) = 0$  se, e somente se,  $x = y$ ;  $d(x, y) = d(y, x)$ ;  $d(x, y) \leq d(x, z) + d(z, y)$ . Definimos a *bola* de centro  $x$  e raio  $r$  como sendo o conjunto

$$B_r(x) = \{y \in \mathbb{Z}_p^n : d(x, y) \leq r\}.$$

O número de coordenadas não nulas de  $x = (x_1, \dots, x_n) \in \mathbb{Z}_p^n$  é o chamado *peso de Hamming* de  $x$  e denotado por  $w(x)$ . Note que  $w(x) = d(\mathbf{0}, x)$  e  $d(x, y) = w(x - y)$ .

Seja  $C$  um código em  $\mathbb{Z}_p^n$ . Suponha que  $x$  é a mensagem recebida e que  $x$  não é uma palavra-código. Decodificamos  $x$  como sendo a palavra código “mais próxima” de  $x$ , ou seja, a palavra-código  $c$  tal que

$$d(x, c) = \min \{d(x, c') : c' \in C\}.$$

Este processo de decodificação é conhecido como *decodificador por máxima verossimilhança*. Um problema de primeira grandeza em teoria dos códigos é determinar algoritmos de decodificação eficientes.

O quanto podemos errar para termos a certeza de que o decodificador por máxima verossimilhança está retornando a mensagem original? Esta grandeza é conhecida como a *capacidade de correção de erros* do código  $C$  e é definida como sendo o número

$$t = \max \{r : B_r(c) \cap B_r(c') = \emptyset, c \neq c' \in C\}.$$

A *distância mínima de Hamming* de um código  $C$  é

$$d(C) = \min \{d(c, c') : c \neq c' \in C\}.$$

**Teorema 1** *Seja  $t$  a capacidade de correção de erros de  $C$ . Então*

$$t = \left\lfloor \frac{d(C) - 1}{2} \right\rfloor.$$

Temos dois problemas para resolver (para minimizar). O cálculo da distância mínima e a detecção de erros. Os códigos utilizados na prática são gigantes e por isso a

comparação dois a dois (tanto para a detecção de erros quanto para o cálculo da distância mínima) é pouco viável. Estes problemas são minimizados se consideramos a estrutura vetorial.

**Definição 3** *Um subespaço vetorial  $C$  do espaço  $\mathbb{Z}_p^n$  é chamado de código linear. Se  $\dim(C) = k$  diremos que  $C$  é um  $[n, k]$  código. Um  $[n, k]$  código  $C$  é um  $(n, q^k)$  código.*

No caso dos códigos lineares o cálculo da distância mínima se reduz ao cálculo do peso mínimo:

**Teorema 2** *Seja  $C$  um código linear. Então*

$$d(C) = \min \{w(c) : \mathbf{0} \neq c \in C\}.$$

Apresentaremos agora, para um código específico, soluções satisfatórias para os dois problemas mencionados acima.

### Código Binário de Hamming

Seja

$$C = \{(a, b, c, d, b + c + d, a + c + d, a + b + d) : a, b, c, d \in \mathbb{Z}_2\}.$$

Temos que

$$\begin{aligned} (a, b, c, d, b + c + d, a + c + d, a + b + d) &= a(1, 0, 0, 0, 0, 1, 1) + \\ &+ b(0, 1, 0, 0, 1, 0, 1) + c(0, 0, 1, 0, 1, 1, 0) + d(0, 0, 0, 1, 1, 1, 1). \end{aligned}$$

Como

$$\alpha = \{(1, 0, 0, 0, 0, 1, 1), (0, 1, 0, 0, 1, 0, 1), (0, 0, 1, 0, 1, 1, 0), (0, 0, 0, 1, 1, 1, 1)\}$$

é linearmente independente em  $\mathbb{Z}_2^7$  concluímos que  $C$  é um  $[7, 4]$  código linear gerado por  $\alpha$  ( $\alpha$  é uma base de  $C$ ).

O código  $C$  acima é conhecido como *código binário de Hamming* (ou simplesmente código de Hamming). O código de Hamming é denotado classicamente por  $\mathcal{H}_3$ . Nosso objetivo é calcular a distância mínima (e consequentemente a capacidade de correção de erros) de  $\mathcal{H}_3$ . Vamos fazer isto sem explicitar as  $2^4$  palavras-código de  $\mathcal{H}_3$ .

Como  $\alpha$  gera  $\mathcal{H}_3$ , pondo

$$G = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix},$$

temos que

$$\mathcal{H}_3 = \{(a, b, c, d) \cdot G : a, b, c, d \in \mathbb{Z}_2\}.$$

Considere agora

$$H = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{pmatrix}.$$

Temos que  $H \cdot G^T = \mathbf{0}$  (a matriz nula). Isto implica que  $H \cdot c^T = \mathbf{0}$  para todo  $c \in \mathcal{H}_3$ . Como o posto (número máximo de linhas ou colunas linearmente independentes) de  $H$  é igual a 3, concluímos que a dimensão do espaço solução do sistema homogêneo  $H \cdot x^T = \mathbf{0}$  é  $7 - 3 = 4$ , que é igual a dimensão de  $\mathcal{H}_3$ . Isto implica que

$$\mathcal{H}_3 = \{x \in \mathbb{Z}_2^7 : H \cdot x^T = \mathbf{0}\}.$$

As matrizes  $G$  e  $H$  descritas acima são conhecidas respectivamente como matrizes *geradora* e de *paridade* de  $\mathcal{H}_3$ .

Seja  $c \in \mathcal{H}_3$  tal que  $w(c) = s$ . Seja  $\{i_1, i_2, \dots, i_s\}$  as entradas não nulas de  $c$ . Como  $H \cdot c^T = \mathbf{0}$ , concluímos que as colunas  $h_{i_1}, h_{i_2}, \dots, h_{i_s}$  de  $H$  são linearmente dependentes. Vale a recíproca: se  $h_{i_1}, h_{i_2}, \dots, h_{i_s}$  são colunas linearmente dependentes de  $H$ , então existe  $c \in \mathcal{H}_3$  tal que  $w(c) = s$ . Conclusão:

**Teorema 3**  $d(\mathcal{H}_3) = s$  se, e somente se, todo conjunto com  $s - 1$  colunas de  $H$  é linearmente independente e existem  $s$  colunas linearmente dependentes.

Como todo conjunto com 2 colunas de  $H$  é linearmente independente e existem 3 colunas linearmente dependentes ( $\{h_2, h_3, h_4\}$  por exemplo):

**Teorema 4**  $d(\mathcal{H}_3) = 3$ .

Consequentemente:

**Corolário 1** A capacidade de correção de erros de  $\mathcal{H}_3$  é igual a 1.

O código de Hamming contém  $2^4$  palavras-código. Sendo assim podemos utilizar  $\mathcal{H}_3$  para codificar  $2^4$  informações. Como a capacidade de correção de erros de  $\mathcal{H}_3$  é 1,  $\mathcal{H}_3$  deve ser implementado em um canal de comunicação com as seguintes características: a probabilidade de ocorrer mais de um erro é menor do que a probabilidade de ocorrer um erro.

Suponha que nosso canal de comunicação tem as características descritas acima e recebemos a mensagem

$$x = (1, 1, 0, 0, 1, 1, 1).$$

Como  $H \cdot x^T = (0, 0, 1) \neq (0, 0, 0)$ ,  $x \notin \mathcal{H}_3$ . Sabemos que uma das coordenadas foi trocada (baseado nas características do canal). Qual? Poderíamos responder a questão comparando  $x$  com cada palavras-código de  $\mathcal{H}_3$ . No entanto, observando que  $x = c + e_i$  para algum  $c \in \mathcal{H}_3$  com  $e_i$  sendo o  $i$ -ésimo vetor canônico (a única coordenada não nula de  $e_i$  é a  $i$ -ésima coordenada), temos

$$\begin{aligned} H \cdot x^T &= H \cdot (c + e_i)^T = H \cdot c^T + H \cdot e_i^T = \mathbf{0} + H \cdot e_i^T \\ &= i\text{-ésima coluna de } H. \end{aligned}$$

Assim o erro em  $x$  se localiza na última coordenada pois  $(0, 0, 1)$  é a última coluna de  $H$ : o vetor  $x$  é decodificado como sendo a palavra-código

$$c = (1, 1, 0, 0, 1, 1, 0).$$

## Sobre os Autores

Pedro Henrique Galvão dos Santos é aluno do 1o. ano do ensino médio do Colégio Estadual Professor Flávio Warken, em Foz do Iguaçu, e tem participado nos últimos anos da Olimpíada Brasileira de Matemática das Escolas Públicas - OBMEP, tendo conquistado em 2005 a medalha de ouro (e contemplado com a bolsa de iniciação científica junior do CNPq), em 2006 uma menção honrosa e em 2007 a medalha de prata (e novamente contemplado com a bolsa de iniciação científica junior do CNPq).

Luciano Panek é professor assistente de matemática da Universidade Estadual do Oeste do Paraná - UNIOESTE, campus Foz do Iguaçu, e atua como professor orientador dos alunos premiados pela OBMEP da região de Foz do Iguaçu.

## Referências Bibliográficas:

- ALVES, M.M.S.: **Códigos Corretores de Erros**, XI ERMAC, (2007), UFPR, Curitiba-PR.
- BOLDRINI, J.L. et al.: **Álgebra Linear**, Harbra, 3a. Edição, (1986).
- DOMINGUES, H.H., IEZZI, G.: **Álgebra Moderna**, Atual, (1982).
- HEFEZ, A., VILLELA, M.L.T.: **Códigos Corretores de Erros**, Série de Computação e Matemática, IMPA, (2002).

# CÓDIGOS LINEARES: CODIFICAÇÃO E DECODIFICAÇÃO DE CÓDIGOS COM UM ERRO

Santos Richard Wieller Sanguino Bejarano<sup>1</sup>, Maycon Gonçalves Carneiro<sup>2</sup>.

Universidade Tecnológica Federal do Paraná - UTFPR - Campus Pato Branco,

srichardwsb@utfpr.edu.br; mayconadv@gmail.com

## Resumo:

Estudado por Matemáticos e Engenheiros, os códigos corretores de erros são extremamente usados na transmissão ou armazenamento de informações. Na prática, a classe de códigos corretores de erros mais usada, é a chamada classe dos códigos lineares. Utilizando linguagens de programação, podemos desenvolver algoritmos que codifiquem e decodifiquem mensagens de forma mais rápida e prática. Este trabalho visa explicar brevemente os processos de codificação e decodificação, dos códigos lineares com apenas um erro, por meio de um exemplo feito em linguagem de programação Pascal. Mostramos assim, uma aplicação prática da matemática aplicada e computacional.

**Palavras-chave:** Códigos Lineares, codificação, decodificação, pascal.

## Introdução

Segundo Hefez(2002), um código corretor de erros é um modo organizado de adicionar-se algum dado à uma mensagem que se queira transmitir ou armazenar, que possibilite, ao recuperar a informação, detectar e corrigir erros, ou seja, diferente de criptografia, onde se introduz o erro, para que não se consiga decifrar a mensagem.

---

<sup>1</sup>Professor orientador do curso de Licenciatura em Matemática

<sup>2</sup>Projeto desenvolvido no PIBIC com apoio financeiro da Fundação Araucária



Admitindo que o público alvo do trabalho possua conhecimento das ferramentas algébricas envolvidas com a Teoria e das definições e resultados básicos, passamos então à alguns resultados mais específicos para o desenvolvimento do nosso exemplo utilizando Pascal.

### Códigos Lineares

Seja  $C \subset F^n$  um código linear, define-se  $C^\perp = \{v \in F^n : \langle v, u \rangle = 0, \forall u \in C\}$ , onde  $\langle v, u \rangle$  denota o produto interno em  $F^n$ .

Temos que  $C^\perp$  é um subespaço vetorial de  $F^n$ , ortogonal a  $C$  pois,  $x \in C^\perp \Leftrightarrow Gx^t = 0$ , onde  $G$  é a matriz geradora de  $C$  associada a alguma base. Tal subespaço,  $C^\perp$ , é também um código linear que será chamado de código dual de  $C$ .

Sendo  $C \subset F^n$  um código de dimensão  $k$ , com matriz geradora  $G = (Id_k \mid A_{k \times (n-k)})$ , na forma padrão, então  $\dim C^\perp = n - k$  e  $H = (-A^t \mid Id_{n-k})$  é uma matriz geradora de  $C^\perp$ , o que se verifica pois se  $v \in C^\perp \Leftrightarrow Gv^t = 0$  então

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & | & a_{1,k+1} & \dots & a_{1,n} \\ 0 & 1 & 0 & 0 & 0 & | & a_{2,k+1} & \dots & a_{2,n} \\ 0 & 0 & \ddots & \vdots & \vdots & | & a_{3,k+1} & \dots & a_{3,n} \\ \vdots & \vdots & \vdots & 1 & 0 & | & \vdots & & \vdots \\ 0 & 0 & \dots & 0 & 1 & | & a_{k,k+1} & \dots & a_{k,n} \end{pmatrix}_{k \times n} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{pmatrix}_{n \times 1} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_{k \times 1} \quad \text{daí}$$

$$\begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_k \end{pmatrix} = -A \begin{pmatrix} v_{k+1} \\ v_{k+2} \\ \vdots \\ v_n \end{pmatrix}$$

Logo, como  $F$  possui  $q$  elementos, existem  $q^{n-(k+1)+1} = q^{n-k}$  possibilidades para  $v$ , ou seja,  $C^\perp$  possui  $q^{n-k}$  elementos, então sua dimensão é  $n - k$ .

Temos ainda que  $H$  gera um subespaço vetorial de dimensão  $n - k$ , pois as linhas de  $H$  são linearmente independentes devido ao bloco  $Id_{n-k}$ . Notemos que, sendo  $H_i, 1 \leq$

$i \leq n - k$  é a  $i$ -ésima linha de  $H$ ,  $H_i = (-a_{1i}, -a_{2i}, \dots, -a_{ki}, 0, 0, \dots, 1, 0, \dots, 0)$  e a  $j$ -ésima linha de  $G$ , denotada  $G_j$ ,  $1 \leq i \leq k$ , será  $G_j = (0, 0, \dots, 1, 0, \dots, 0, 0, a_{j1}, a_{j2}, \dots, a_{j,n-k})$  então  $\langle H_i, G_j \rangle = -a_{ji} + a_{ij} = 0$ , ou seja, todas as linhas de  $H$  são ortogonais às linhas de  $G$ . Assim, o espaço gerado pelas linhas de  $H$  está contido em  $C^\perp$ , e os dois subespaços têm a mesma dimensão, assim eles coincidem, logo  $H$  é uma matriz geradora de  $C^\perp$ .

Vemos que se  $C$  é um código linear em  $F^n$  de dimensão  $k$ , então  $C^\perp$  é um código de dimensão  $n - k$ , e que se  $G$  é a matriz geradora de  $C$ , então  $H$  de ordem  $(n - k) \times n$  com coeficientes em  $F$  e linhas linearmente independentes, é uma matriz geradora de  $C^\perp$  se, e somente se,  $GH^t = 0$ .

Tal matriz,  $H = (-A^t \mid Id_{n-k})$ , geradora de  $C^\perp$  é chamada de matriz teste de paridade.

Assim, dado um código  $C$  e  $H$  a matriz teste de paridade desse código, para verificarmos se um determinado vetor  $v \in F^n$  pertence ou não à  $C$ , basta verificar se é nulo o vetor  $Hv^t$ . Tal vetor,  $Hv^t$ , é chamado de *síndrome* de  $v$ .

Vimos portanto, o processo de **codificação** dos códigos lineares. Passemos agora ao processo de **decodificação**. Tal processo consta da detecção e correção de erros de um determinado código.

Primeiramente precisamos definir um vetor erro  $e$ , como sendo a diferença entre o vetor recebido  $v$  e o vetor transmitido  $c$ , isto é  $e = v - c$  logo, se obtivermos  $Hv^t = 0$ , assumindo que no processo de transmissão **apenas** um erro ocorra, e que nosso código possua  $d \geq 3$  então, o vetor recebido  $v$  pertence ao código, o qual será igual a um único vetor deste. Então  $e = (0, 0, \dots, 0)$  e  $v = c$ . Caso  $Hv^t \neq 0$  então, ocorreram erros durante a transmissão, logo teremos um vetor erro  $e = v - c$ , onde o peso desse vetor  $e$ , corresponde ao número de erros cometidos entre a transmissão e a recepção.

Como  $Hc^t = 0$ , então temos  $He^t = H(v^t - c^t) = Hv^t - Hc^t = Hv^t$ , assim a palavra recebida e o vetor erro têm mesma síndrome.

O problema então, é como determinar esse vetor erro  $e$  a partir de  $Hv^t$ . Vejamos

a determinação desse vetor  $e$  quando  $w(e) \leq 1$ .

Para tanto, suponhamos que  $C$  tenha distância mínima  $d \geq 3$ , pois como  $C$  pode corrigir até  $t = \lfloor \frac{d-1}{2} \rfloor$  erros, assim com  $d \geq 3$ ,  $C$  corrige pelo menos um erro. Suponhamos ainda que  $w(e) \leq 1$ , ou seja, o canal introduziu no máximo um erro.

Se  $He^t = 0$ , então  $v \in C$  e tomamos  $c = v$ .

Suponhamos  $He^t \neq 0$ , logo  $w(e) = 1$  e, portanto o vetor erro  $e$ , tem apenas uma coordenada não nula. Consideremos que  $e = (0, \dots, 0, \alpha, 0, \dots, 0)$  com  $\alpha \neq 0$  na  $i$ -ésima posição. Assim,  $He^t = \alpha h^i$ , onde  $h^i$  é a  $i$ -ésima coluna de  $H$ . Portanto tomamos tal vetor  $e$  como sendo o vetor com todas as componentes nulas exceto a  $i$ -ésima componente que é  $\alpha$ .

### Implementação Computacional

Estabelecemos então, o algoritmo de decodificação em códigos corretores de um erro.

Seja  $H$  a matriz teste de paridade do código  $C$  e seja  $v$  um vetor recebido. (Suponha  $d \geq 3$ )

1. Calcule  $Hv^t$ .
2. Se  $Hv^t = 0$ , aceite  $v$  como sendo a palavra transmitida.
3. Se  $Hv^t = s^t \neq 0$  compare  $s^t$  com as colunas de  $H$ .
4. Se existirem  $i$  e  $\alpha$  tais que  $s^t = \alpha h^i$ , para  $\alpha \in F$ , então  $e$  é a  $n$ -upla com  $\alpha$  na posição  $i$  e zero nas outras posições. Corrija  $v$  pondeo  $c = v - e$ .
5. Se não existirem  $i$  e  $\alpha$  tais que  $s^t = \alpha h^i$ , então mais de um erro foi cometido.

Vejamos agora um exemplo, com a aplicação da codificação e decodificação de um código corretor de um erro, com a ajuda do programa Pascal.

Para isso, considere que tenhamos o seguinte código fonte:

$$\left[ \begin{array}{ccccc} A = 10000 & B = 01000 & C = 00100 & D = 00010 & E = 00001 \\ F = 11000 & G = 10100 & H = 10010 & I = 10001 & J = 01100 \\ L = 01010 & M = 01001 & N = 00110 & O = 00101 & P = 00011 \\ Q = 11100 & R = 10110 & S = 10101 & T = 11010 & U = 11001 \\ V = 01110 & X = 00111 & Z = 11110 & ESPAÇO = 00000 & \end{array} \right]$$

e que queiramos codificar e transmitir a seguinte mensagem: XII ERMAC .

Para isso, considere a matriz

$$G = \begin{pmatrix} 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

geradora de  $C$ , a qual possui  $d = 3$ , assim  $C$  pode corrigir um erro.

Sendo  $c$  uma palavra do nosso código fonte, para codificá-la fazemos  $c.G$ . Para isso, usamos o seguinte programa em linguagem pascal:

```
Program codific;
var a,b,c,i,j,h,k:integer;
g,x,v: array [1..100,1..100] of integer;
Begin
clrscr;
writeln('informe o número de elementos do corpo k = F..');
readln(a);
writeln('informe a dimensão do espaço de saída');
readln(b);
writeln('informe a dimensão do espaço de chegada');
readln(c);
writeln('informe as entradas da matriz de transformação:');
```

```

for i:= 1 to b do
for j:= 1 to c do
Begin
write('Entre com o valor G['i, ',',j, ']: ');
readln(g[i,j]);
End;
for h:=1 to 2 do
begin
writeln("");
writeln('informe o vetor a ser codificado:');
for i:= 1 to b do
begin
write('entre com o valor v[1, ',',i,']:');
readln(v[1,i]);
end;
for i:=1 to 1 do
begin
for j:=1 to c do
begin
x[i,j]:=0;
for k:= 1 to b do x[i,j]:=((x[i,j]+v[i,k]*g[k,j])mod a);
write(x[i,j]);
end;
end;
end;
End.

```

Neste programa o usuário informa o corpo a ser trabalhado, a dimensão e o comprimento da matriz  $G$ , então o programa pede para o usuário informar a palavra a ser codificada, fazendo então a codificação. Com esse programa podemos codificar qualquer código linear através de uma matriz geradora.

Assim, nossa mensagem codificada será: 100001101 - 111110100 - 111110100 - 000000000 - 001100100 - 011111001 - 101000110 - 110010000 - 111000001.

Agora, suponha que tenhamos recebido a seguinte mensagem: 101010110 - 110010000 - 000011010 - 001100100 - 101000110 - 110010000 - 000011110 - 111110100 - 111000001 - 110010000. Onde no máximo um erro foi introduzido pelo canal.

Para decodificá-la, precisamos encontrar a síndrome da mesma para saber se houve erro ou não, e onde ocorreu o erro. Para isso precisamos primeiramente da matriz teste de paridade de  $C$ , ou seja,  $H$  que é obtida de  $G$ , sendo:

$$H = \begin{pmatrix} 1 & 0 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Assim, o programa calculará a síndrome do vetor, identificando a presença de erro ou não. Esse programa é "composto" de duas partes, visto que num primeiro momento ele calcula a síndrome de uma palavra de qualquer código linear com apenas um erro. Já na segunda parte, ele foi desenvolvido especificamente para este exemplo, retornando a palavra codificada, visto que no algoritmo de decodificação apenas calcula-se a síndrome e apresenta-se o vetor transmitido sem o erro, mas não a palavra decodificada. Assim, segue o programa específico para este exemplo:

```

Program decodific;
var a,b,c,i,j,h,k:integer;
g,x,v: array [1..100,1..100] of integer;
Begin

```

```

clrscr;
writeln('informe o número de elementos do corpo k = F..');
readln(a);
writeln('informe a dimensão da matriz teste de paridade H:');
readln(b);
writeln('informe o comprimento da matriz teste de paridade H:');
readln(c);
writeln('informe as entradas da matriz teste de paridade H:');
for i:= 1 to b do
for j:= 1 to c do
Begin
write('Entre com o valor G[',i, ', ',j, ']: ');
readln(g[i,j]);
End;
for h:=1 to 10 do
begin
writeln("");
writeln('informe o vetor a ser decodificado:');
for j:= 1 to c do
begin
write('entre com o valor vt[',1, ', ',j, ']:');
readln(v[1,j]);
end;
write('a síndrome é:');
for i:=1 to 1 do
begin
for j:=1 to b do

```

```

begin
x[i,j]:=0;
for k:= 1 to c do x[i,j]:=((x[i,j]+v[i,k]*g[j,k])mod a);
write(x[i,j]);
end;
end;
if ((x[1,1]=0) and (x[1,2]=0) and (x[1,3]=0) and (x[1,4]=0)) or
((x[1,1]=1) and (x[1,2]=1) and (x[1,3]=1) and (x[1,4]=1)) or
((x[1,1]=1) and (x[1,2]=0) and (x[1,3]=1) and (x[1,4]=1)) or
((x[1,1]=0) and (x[1,2]=1) and (x[1,3]=1) and (x[1,4]=1)) or
((x[1,1]=1) and (x[1,2]=1) and (x[1,3]=0) and (x[1,4]=1)) then
begin
writeln(' e a palavra enviada foi: ',v[1,5],v[1,8],v[1,9],v[1,6],v[1,7],'.');
end;
if (x[1,1]=1) and (x[1,2]=0) and (x[1,3]=1) and (x[1,4]=0)then
begin
writeln(' e a palavra enviada foi: ',(((v[1,5])+1)mod 2),v[1,8],v[1,9],v[1,6],v[1,7],'.');
end;
if (x[1,1]=1) and (x[1,2]=0) and (x[1,3]=0) and (x[1,4]=0)then
begin
writeln(' e a palavra enviada foi: ',v[1,5],v[1,8],v[1,9],(((v[1,6])+1)mod 2),v[1,7],'.');
end;
if (x[1,1]=0) and (x[1,2]=1) and (x[1,3]=0) and (x[1,4]=0)then
begin
writeln(' e a palavra enviada foi: ',v[1,5],v[1,8],v[1,9],v[1,6],(((v[1,7])+1)mod 2),'.');
end;
if (x[1,1]=0) and (x[1,2]=0) and (x[1,3]=1) and (x[1,4]=0)then

```



```

begin
writeln(' e a palavra enviada foi: ',v[1,5],(((v[1,8])+1)mod 2),v[1,9],v[1,6],v[1,7],'.');
end;
if (x[1,1]=0) and (x[1,2]=0) and (x[1,3]=0) and (x[1,4]=1)then
begin
writeln(' e a palavra enviada foi: ',v[1,5],v[1,8],(((v[1,9])+1)mod 2),v[1,6],v[1,7],'.');
end; if ((x[1,1]=1) and (x[1,2]=1) and (x[1,3]=1) and (x[1,4]=0)) or
((x[1,1]=0) and (x[1,2]=1) and (x[1,3]=0) and (x[1,4]=1)) or
((x[1,1]=0) and (x[1,2]=1) and (x[1,3]=1) and (x[1,4]=0)) or
((x[1,1]=0) and (x[1,2]=0) and (x[1,3]=1) and (x[1,4]=1)) or
((x[1,1]=1) and (x[1,2]=1) and (x[1,3]=0) and (x[1,4]=0)) or
((x[1,1]=1) and (x[1,2]=0) and (x[1,3]=0) and (x[1,4]=1)) then
begin
writeln(' foi cometido mais que um erro. ');
end;
end;
End.

```

Assim ao passarmos nossa mensagem recebida pelo programa o mesmo nos retorna as seguintes informações:

```

101010110- a síndrome é:1010 e a palavra enviada foi:01001
110010000- a síndrome é:0000 e a palavra enviada foi:10000
000011010- a síndrome é:0000 e a palavra enviada foi:11010
001100100- a síndrome é:0000 e a palavra enviada foi:00001
101000110- a síndrome é:0000 e a palavra enviada foi:01001
110010000- a síndrome é:0000 e a palavra enviada foi:10000
000011110- a síndrome é:0100 e a palavra enviada foi:11010
111110100- a síndrome é:0000 e a palavra enviada foi:10001

```

111000001- a síndrome é:0000 e a palavra enviada foi:00100

110010000- a síndrome é:0000 e a palavra enviada foi:10000.

Logo, notamos que ocorreram erros na transmissão da primeira e da sétima palavra, cujas síndromes foram diferentes de 0 então o programa corrigiu o erro e nos retornou a palavra correta que foi enviada.

Estes são exemplos simples que nos mostram a grande utilidade da matemática aplicada e computacional, para resolvermos problemas com mais facilidade. Como foi mencionado, esse programa possuiu uma segunda parte específica para esse problema, retornando as síndromes bem como as palavras corretas, mas o mesmo pode ser alterado conforme o problema proposto.

### **Agradecimentos**

Agradecemos ao Professor Carlos A. Ribeiro Martins e aos colegas do grupo de iniciação científica pela colaboração no desenvolvimento do presente trabalho.

### **Referências Bibliográficas:**

HEFEZ, A., VILLELA, M.L.T.: **Códigos Corretores de Erros**: Serie Computação e Matemática. IMPA: Rio de Janeiro. 2002.

HEFEZ, A.: **Curso de Álgebra**: Coleção Matemática Universitária. IMPA: Rio de Janeiro, vol.1, 2ed.,1993.

SILVA, J.G.Filho: **Informação, codificação e segurança de dados**, Universidade de Brasília. Disponível em: <http://www.ene.unb.br/juliana/cursos/teoriainf/codsec07.pdf>  
Acesso em: 02 mai. 2008.

SOUZA, A.O.; CÂMARA, M.A.: **Códigos Corretores de Erros Lineares**, Monografia. Uberlândia, 2006. 36p. Especialização (Matemática) - Faculdade de Matemática, Universidade Federal de Uberlândia - (UFU).

CARROL, D.W.: **Progromação em Turbo Pascal**. McGraw-Hill: São Paulo, 1988.

# UMA APLICAÇÃO DA TEORIA FUZZY NA BIOMATEMÁTICA: DIAGNÓSTICO DE DOENÇAS INFANTIS

Franciele Buss Frescki, Priscila Pigatto.

Universidade Estadual do Oeste do Paraná - UNIOESTE, f\_\_ran@hotmail.com; Universidade  
Estadual do Oeste do Paraná - UNIOESTE, pricipigatto@hotmail.com

**Palavras-chave:** Modelos Fuzzy, Diagnóstico Médico, Biomatemática.

## Resumo

Este trabalho apresenta de forma breve a Teoria dos conjuntos Fuzzy, com algumas definições pertinentes. Após, apresentamos uma das diversas aplicações desta teoria, com a modelagem de um diagnóstico médico de doenças infantis.

Bassanezi (1999), afirma que matemática aplicada é fundamentalmente interdisciplinar e seu objetivo consiste em trazer aplicabilidade a alguma estrutura matemática fora do seu campo de estudo imediato; a modelagem, por sua vez, é uma ferramenta indispensável da Matemática Aplicada.

A modelagem de um fenômeno pode ser realizada com o auxílio de diferentes teorias da Matemática; pode-se lançar mão, por exemplo, de EDPs, EDOs, teoria fuzzy, teoria do caos e bifurcações, teoria dos fractais, entre outras. Para ilustrar uma situação que poderia ser modelada pela lógica fuzzy, imagine uma população de fumantes onde se quer descobrir como estará esta população num futuro. Supõe-se que em algum instante  $t_1$ , um fumante está sujeito a uma taxa de mortalidade. Para avaliar cada indivíduo, pode-se pensar nele como fumante ou não fumante, atribuindo valores 1 e 0 respectivamente. Mas se imaginarmos que uma pessoa pode ser classificada pela quantidade de cigarros que fuma num dia, este quadro muda, e poderiam ser atribuídos valores diferentes de 1 e 0, a depender do número de cigarros fumados por dia por indivíduo. Este número entre 0 e 1, diz-se que está medindo o grau de pertinência do indivíduo ao conjunto dos fumantes. Por exemplo, considerando uma pessoa que fuma 3 carteiras de cigarro ao dia como um fumante (grau de pertinência 1), e uma pessoa que não fuma (grau de pertinência 0), qual o grau de pertinência ao conjunto de fumantes de uma pessoa que fuma 1 carteira de cigarros por dia? Esta questão de ser ou não ser

fumante é uma característica dos conjuntos fuzzy, ou seja, os conjuntos fuzzy carregam uma característica de subjetividade.

Na modelagem matemática de processos variacionais, a propriedade fundamental, utilizando sistemas de equações determinísticas, é a exatidão alcançada nas previsões de fenômeno. Logicamente, tais previsões ou inferências são dependentes de informações precisas que são inseridas nos valores médios dos parâmetros utilizados. Por outro lado, nos modelos estocásticos, as soluções médias dos modelos são obtidas a posteriori quando se tem alguma distribuição estatística de dados referentes ao fenômeno analisado.

A palavra “fuzzy”, em inglês, significa incerto, impreciso, subjetivo, nebuloso. Um outro exemplo de relação fuzzy pode ser dado considerando-se uma amizade. Na teoria clássica, considera-se uma pessoa pode ser amigo ou não amigo de outra. Mas uma relação de amizade fuzzy entre duas pessoas considera o grau de amizade entre elas. Sendo assim dois ou mais indivíduos podem se relacionar com diferentes graus de amizade, desde 1,0 (são certamente amigos) até 0,0 ( não são amigos). Nesse intervalo, pode-se ter diversos graus de amizade.

### **Preliminares matemáticas**

Formalmente, uma relação fuzzy  $R$  entre duas variáveis  $x$  que pertence a  $X$  e  $y$  que pertence a  $Y$ , é definida por uma função que mapeia o par ordenado  $(x,y)$  no espaço  $X \times Y$  para o seu grau na relação, ou seja,  $R : X \times Y$  que vai para o intervalo  $[0,1]$ .

Seja  $U$  um conjunto.

**Definição:** um subconjunto fuzzy  $F$  de  $U$  é caracterizado por uma função  $\mu(x) : U \rightarrow [0,1]$ , chamada função da pertinência do conjunto fuzzy  $F$ .

O valor  $\mu(x) \in [0,1]$  indica com que grau que o elemento  $x$  de  $U$  está no conjunto fuzzy  $F$ , com  $\mu(x) = 0$  e  $\mu(x) = 1$  indicando, respectivamente, a não pertinência e a pertinência completa de  $x$  ao conjunto fuzzy.

Do ponto de vista formal, a definição de um subconjunto fuzzy foi obtida simplesmente ampliando-se o contra-domínio da função característica, que é o conjunto  $\{0,1\}$ , para o intervalo  $[0,1]$ .

A Teoria Fuzzy surgiu em 1965, com uma publicação de Lotfi A. Zadeh como uma nova teoria de conjuntos. Zadeh criou uma teoria em que não há discontinuidades, ou seja, um elemento pode “pertencer em parte” a um conjunto.

Com a percepção de que, para modelagem de fenômenos industriais, biológicos ou químicos, as técnicas convencionais tornariam os problemas demasiadamente complexos, Zadeh utilizou os sistemas fuzzy. Isto significa que com outras ferramentas os modelos ficariam talvez não insolúveis, mas necessitariam de muitas variáveis. Assim, a lógica fuzzy vem para resolver problemas em modelos que envolvam subjetividade na classificação de variáveis.

O princípio fundamental da Teoria Fuzzy, princípio da dualidade, estabelece que dois eventos antagônicos podem existir simultaneamente. Isto é, um elemento pode pertencer, em um certo grau, a um conjunto e, em um outro grau, a um outro conjunto.

A composição em relações fuzzy é um conceito importante na modelagem de um diagnóstico médico, e será introduzida:

Considere  $R$  e  $S$  duas relações fuzzy binárias em  $U_1 \times U_2 \times \dots \times U_n$ , respectivamente. Então a composição  $RoS$  é uma relação fuzzy binária em  $U_1 \times U_3$ , com função de pertinência dada por:

$$u_{R \circ S}(x_1, x_3) = \max_{x_2 \in U_2} [\min(u_R(x_1, x_2), u_S(x_2, x_3))]$$

Sejam dois conjuntos  $U_1$  e  $U_2$ . Sejam  $F(U_1)$  e  $F(U_2)$  as classes dos conjuntos fuzzy de  $U_1$  e  $U_2$ , respectivamente, e  $R$  uma relação binária sobre  $U_1 \times U_2$ . Então a relação  $R$  define um funcional de  $F(U_1)$  em  $F(U_2)$  que a cada elemento  $A_1$  que pertence a  $F(U_1)$  faz corresponder o elemento  $A_2$  que pertence a  $F(U_2)$ . A função de pertinência é dada por:

$$u_{A_2}(x_2) = u_{R(A_1)}(x_2) = \max_{x_1 \in U_1} [\min(u_{A_1}(x_1), u_R(x_1, x_2))]$$

Utilizando estas e outras definições, (que serão omitidas aqui, mas que podem ser encontradas em Villela, Santos e Jafelice (2007)), foi possível construir um modelo para o diagnóstico de doenças infantis, isto após a coleta e avaliação de dados.

### **Uma aplicação médica**

A aplicação da Teoria Fuzzy na Biomatemática apresentada neste trabalho refere-se ao diagnóstico médico de doenças infantis. Nesta aplicação, Villela, Santos e Jafelice (2007) propõem um modelo que imite o diagnóstico de um médico sobre seus pacientes a partir dos sintomas apresentados. O objetivo foi relacionar os sintomas de pacientes com as possíveis doenças, que seriam catapora, meningite, caxumba e coqueluche. Para a composição do modelo foram analisados os dados de dois médicos.

Assim, espera-se com esta modelagem por um sistema fuzzy, que se possa ajudar o médico em seu diagnóstico e, conseqüentemente, o ajude a tomar decisões e optar por exames laboratoriais mais detalhados.

Considere os seguintes conjuntos:

- A = conjuntos dos pacientes do primeiro médico;
- B = conjuntos dos pacientes do segundo médico;
- S = conjunto dos sintomas;
- D = conjunto das doenças.

Foram analisadas as informações de dois médicos, os quais informaram sobre sete pacientes  $p_1, p_2, p_3, p_4, p_5, p_6, p_7$ , com sintomas (s) e doenças (d):

$s_1$  = calafrios

$s_{12}$  = infecção das glândulas salivares

$s_2$  = cansaço

$s_{13}$  = mal estar

$s_3$  = cefaléia

$s_{14}$  = perda de apetite

$s_4$  = coceira

$s_{15}$  = pintas vermelhas no corpo

$s_5$  = confusão mental

$s_{16}$  = rigidez na nuca

$s_6$  = coriza

$s_{17}$  = tosse seca

$s_7$  = dor ao mastigar ou engolir

$s_{18}$  = vômito

$s_8$  = dor de garganta

$d_1$  = catapora

$s_9$  = dor muscular

$d_2$  = caxumba

$s_{10}$  = febre

$d_3$  = coqueluche

$s_{11}$  = fraqueza

$d_4$  = meningite

Montou-se três tabelas com os dados fornecidos pelos médicos. Numa tabela associou-se sintomas a doenças, com respectivo grau de pertinência da relação de cada sintoma com cada doença (por exemplo, para a doença  $d_1$ , que é catapora, atribui-se o grau de pertinência 1 para a relação com o sintoma  $s_1$ , que são pintas vermelhas no corpo; enquanto que atribui-se grau de pertinência 0,2 do mesmo sintoma  $s_1$  em relação à meningite,  $d_4$ ), e em outras duas tabelas (uma para cada médico), associou-se os graus de pertinência com que os sintomas se relacionam com cada um dos sete pacientes. A resposta da composição foi também um conjunto fuzzy, fornecendo a distribuição de possibilidades do paciente no conjunto de doenças. Assim, Vilela, Santos e Jafelice verificaram que a doença com maior possibilidade de ocorrer em cada paciente estava condizente com as moléstias diagnosticadas pelo especialista ou através de consulta à prontuários.

Além disso, encontrou-se a solução de dois sistemas de equações diferenciais relacionados com as epidemias de doenças. Considerando os conjuntos  $P = P(t)$ : pessoas sadias, mas suscetíveis à doença, podendo ser infectadas quando em contato com pessoas doentes;  $M = M(t)$ : pessoas portadoras da doença (infecciosas);  $V = V(t)$ : indivíduos imunes que já contraíram a doença e se recuperaram, ou estão isoladas ou morreram, tem-se um sistema de equações diferenciais ordinárias que descreve o modelo:

$$\begin{aligned}\frac{dP}{dt} &= -\beta MP \\ \frac{dM}{dt} &= \beta MP - \alpha M \\ \frac{dV}{dt} &= \alpha M\end{aligned}$$

onde  $\beta > 0$  é o coeficiente de infecciosidade e  $\alpha > 0$  é a taxa de remoção.

Desta forma, a utilização da teoria dos conjuntos fuzzy na medicina pode auxiliar o médico no diagnóstico, colaborando em suas tomadas de decisões, optando por exames laboratoriais mais detalhados. Espera-se, com o desenvolvimento do estudo em modelagem de fenômenos determinísticos, auxiliar na prática do médico, e com tempo e desenvolvimento de pesquisas, aperfeiçoar modelos para que possam ser incorporados no exercício desta profissão, servindo de ferramenta para a tomada de decisões.

#### **Referências Bibliográficas:**

- BARROS, L.C. **Teoria Fuzzy x Biomatemática**. Campinas: IMECC-UNICAMP, 2002. 52 p. Disponível em <http://www.ime.usp.br/~tonelli/verao-fuzzy/laecio/minicurso3.pdf>. Acesso em 21 jun 2008.
- BASSANEZI, R.C. Modelagem Matemática: Uma disciplina emergente nos programas de formação de professores. In: **Biomatemática**, 1999. p. 9 – 22.
- VILLELA, M.F.S; SANTOS, P.B; JAFELICE, R.S.M. Diagnóstico médico fuzzy de doenças infantis. In: **FAMAT em Revista**. n. 9. p. 329-346, 2007. [http://www.lps.usp.br/neo/fuzzy/fuzzy\\_historico.htm#TrabalhosAplicações](http://www.lps.usp.br/neo/fuzzy/fuzzy_historico.htm#TrabalhosAplicações). Acesso em 21 jun 2008.

# On the Norm of the Fourier-Laplace Projection

Alexander Kushpel.

Universidade Estadual do Campinas, IMECC-UNICAMP, ak99@ime.unicamp.br

## Resumo:

Sharp asymptotic estimates for the norms of Fourier-Laplace projections on  $S^d$ , the  $d$ -dimensional sphere, are established. These results improve weak asymptotic estimates found by Gronwall in 1914 in the case of  $S^2$  [2] and later, using the same method, extended in [3] to the case of  $S^d$ ,  $d \geq 3$ .

**Palavras-chave:** Sphere, Projection, Polynomial.

## Apresentação:

### 1. Introduction

Let  $R^d$  be a  $d$ -dimensional Euclidean space with the usual scalar product  $\langle x, y \rangle = \sum_{k=1}^d x_k y_k$  and for  $d \geq 2$ , let  $S^d$  be the  $d$ -dimensional unit sphere in  $R^{d+1}$ , i.e.

$$S^d = \left\{ x \in R^{d+1} \mid \sum_{k=1}^{d+1} |x_k|^2 = 1 \right\}.$$

Let  $\phi$  be a function on  $S^d$  with a finite  $L_p(S^d)$ -norm, the latter being defined as

$$\|\phi\|_p = \|\phi\|_{L_p(S^d)} = \begin{cases} (\int_{S^d} |\phi(x)|^p d\mu(x))^{1/p}, & 1 \leq p < \infty, \\ \text{ess sup}\{|\phi(x)|, x \in S^d\} & p = \infty. \end{cases}$$

If  $d\mu$  is the normalized rotation-invariant measure on the sphere, then the space  $L_2(S^2)$  has the decomposition

$$L_2(S^d) = \bigoplus_{k=0}^{\infty} H_k,$$



with  $H_k$  denoting the space of homogeneous polynomials of degree  $k$ . If we let  $\gamma_k = -k(k + d - 1)$ , we find that  $H_k = \{f \mid \Delta f = \gamma_k f\}$ , where  $\Delta$  is the Laplace-Beltrami operator on the sphere. Thus  $H_k$  is the eigenspace corresponding to the eigenvalue  $\gamma_k$  of  $\Delta$ . A function  $Z_\eta$  is zonal with respect to a pole  $\eta \in S^d$  if it is invariant under the action of all rotations  $\sigma$  of  $S^d$  that fix  $\eta$ . That is, if  $\sigma \in \mathbf{SO}(d + 1)$  with  $\sigma\eta = \eta$ , then  $Z_\eta(x) = Z_\eta(\sigma x)$  for any  $x \in S^d$ . If  $Z_\eta$  is zonal with respect to  $\eta$ , then there exists a function  $\tilde{Z}$  such that  $Z_\eta(x) = \tilde{Z}(\langle x, \eta \rangle)$  for any  $x \in S^d$ . Whenever the role of pole  $\eta$  is not essential, we shall write  $Z(x)$  instead of  $Z_\eta(x)$ . For  $x = \sigma\eta$ , we define the spherical convolution

$$(\phi * Z)(x) = \int_{S^d} \phi(y) \tilde{Z}(\langle x, y \rangle) d\mu(y) = \int_{S^d} \phi(y) Z_\eta(\sigma^{-1}y) d\mu(\eta).$$

There is a unique, real zonal element  $Z_\eta^{(k)}(x) \in \tilde{Z}^{(k)}(\langle x, \eta \rangle) \in H_k$ , which is a kernel of an orthogonal projection onto  $H_k$ . Elements of  $H_k$  are called spherical harmonics of degree  $k$ . It is known that

$$d_k = \dim H_k = \frac{(2k + d - 1)\Gamma(k + d - 1)}{\Gamma(d)\Gamma(k + 1)} \asymp k^{d-1}, \quad n \rightarrow \infty.$$

If  $\{Y_1^{(k)}, \dots, Y_{d_k}^{(k)}\}$  is any orthonormal basis of  $H_k$ , then the following addition formula is known

$$Z_\eta^{(k)}(x) = \sum_{l=1}^{d_k} \overline{Y_l^{(k)}(\eta)} Y_l^{(k)}(x).$$

Moreover,  $\|Z_\eta\|_2^2 = \sum_{l=1}^{d_k} |Y_l^{(k)}(\eta)|^2 = d_k$  for any  $\eta \in S^d$ . The zonal harmonics have a simple expression in terms of the ultraspherical (Gegenbauer) polynomials  $P_k^\lambda$ , which can be defined in terms of the generating function

$$\frac{1}{(1 - 2\rho t + \rho^2)} = \sum_{k=0}^{\infty} \rho^k P_k^\lambda(\xi), \quad 0 \leq |\rho| < 1, \quad |\xi| \leq 1, \quad \lambda > 0.$$

It is known that

$$Z_\eta^{(k)}(x) = \tilde{Z}^{(k)}(\langle x, \eta \rangle) = \frac{2k + d - 1}{d - 1} P_k^{(d-1)/2}(\cos \theta), \quad (1)$$

where  $\cos \theta = \langle x, \eta \rangle$ . Let  $\phi \in L_1(S^d)$ . Then  $\phi$  has a formal Fourier expansion

$$\phi(x) \sim c_0(\phi) + \sum_{k=1}^{\infty} \sum_{m=1}^{d_k} c_{k,m}(\phi) Y_m^{(k)}(x),$$

where

$$c_{k,m}(\phi) = \int_{S^d} \phi \overline{Y_m^{(k)}} d\mu.$$

The Fourier-Laplace projections (or Fourier sums)  $S_n(\phi, x)$  of  $\phi$  are defined as

$$S_n(\phi, x) = c_0(\phi) + \sum_{k=1}^n \sum_{m=1}^{d_k} c_{k,m}(\phi) Y_m^{(k)}(x).$$

See [1,5] for more information concerning the harmonic analysis on  $S^d$ . In what follows, we shall study an asymptotic behavior of the norms of operator  $S_n : L_\infty(S^d) \rightarrow L_\infty(S^d)$ , as  $n \rightarrow \infty$ .

## 2. Estimates of norms of Fourier-Laplace projections

Let

$$E_n(f) = \inf \{ \|f - t_n\|_\infty \mid t_n \in \oplus_{k=0}^n H_k \}$$

be the best approximation of a function  $f$  by the subspace of polynomials of order  $\leq n$ ,  $\oplus_{k=0}^n H_k$ . Then, by the Lebesgue inequality we get

$$\|f - S_n(f, x)\| \leq (1 + \|S_n\|_{L_\infty(M^d) \rightarrow L_\infty(M^d)}) E_n(f),$$

where  $\|S_n\|_{L_\infty(M^d) \rightarrow L_\infty(M^d)} = \sup \{ \|S_n(f, x)\|_{L_\infty(M^d)} \mid f \in L_\infty(M^d) \}$ . It means that  $S_n(f, x)$  converges uniformly to  $f$  if

$$E_n(f) = \bar{o} \left( \|S_n\|_{L_\infty(S^d) \rightarrow L_\infty(S^d)} \right)^{-1}, \quad n \rightarrow \infty.$$

In the case of  $S^2$  weak asymptotic estimates of  $\|S_n\|_{L_\infty(S^2) \rightarrow L_\infty(S^2)}$  have been found by Gronwall [2] in 1914. Using similar method, Ragozin [3] extended these estimates to

the  $S^d$ , the  $d$ -dimensional sphere. Namely, it was shown that for any  $n$  there are such positive constants  $C_1$  and  $C_2$  that

$$C_1 n^{(d-1)/2} \leq \|S_n\|_{\infty(S^d) \rightarrow L_\infty(S^d)} \leq C_2 n^{(d-1)/2}.$$

The main result of this article is the following statement.

**Theorem 1.** *For the norms of Fourier-Laplace projections the following is true*

$$\|S_n\|_{L_\infty(S^d) \rightarrow L_\infty(S^d)} = \mathcal{K} n^{(d-1)/2} + O(n^{(d-3)/2}), \quad n \rightarrow \infty$$

where

$$\mathcal{K} = \frac{1}{2^{d/2+3/2}\pi^{3/2}\Gamma(d/2)} \int_0^\pi \cos\left(\frac{\theta}{2}\right) (\sin\theta)^{(d-3)/2} d\theta.$$

**Remark.** *In particular, if  $d = 2$ , then*

$$\|S_n\|_{L_\infty(S^2) \rightarrow L_\infty(S^2)} = \frac{\sqrt{n}}{8\pi^{3/2}} \int_0^\pi \sqrt{\cotg\left(\frac{\theta}{2}\right)} d\theta + O(n^{-1/2}), \quad n \rightarrow \infty.$$

If  $d = 3$  we have

$$\|S_n\|_{L_\infty(S^3) \rightarrow L_\infty(S^3)} = \frac{n}{6\pi^2} + O(1), \quad n \rightarrow \infty.$$

If  $d = 5$ , then

$$\|S_n\|_{L_\infty(S^5) \rightarrow L_\infty(S^5)} = \frac{n^2}{16\pi^2} + O(n), \quad n \rightarrow \infty.$$

**Proof.** The Jacoby polynomials  $P_k^{(\alpha,\beta)}(x)$  are orthogonal on  $[-1, 1]$  with the weight function  $\omega(x) = (1-x)^\alpha(1+x)^\beta$ ,  $\alpha > -1$ ,  $\beta > -1$  and normalized as

$$P_k^{(\alpha,\beta)}(1) = \frac{\Gamma(k+\alpha+1)}{\Gamma(\alpha+1)\Gamma(k+1)} \asymp k^\alpha, \quad k \rightarrow \infty.$$

If  $Z_\eta^{(k)} \in H_k$ ,  $k = 0, 1, \dots$ , then

$$\tilde{Z}^{(k)}(t) = C_k(S^d)P_k^{(\alpha,\beta)}(t),$$

where  $\alpha = \beta = (d - 2)/2$  and

$$\begin{aligned} C_k(S^d) &= \frac{(2k + \alpha + \beta + 1)\Gamma(\beta + 1)\Gamma(k + \alpha + \beta + 1)}{\Gamma(\alpha + \beta + 2)\Gamma(k + \beta + 1)} \\ &= \frac{(2k + d - 1)\Gamma(d/2)\Gamma(k + d - 1)}{\Gamma(d)\Gamma(k + d/2)} \asymp k^{d/2}, \quad k \rightarrow \infty. \end{aligned}$$

Now we have,

$$\begin{aligned} S_n(\phi, x) &= c_0(\phi) + \sum_{k=1}^n \sum_{m=1}^{d_k} c_{k,m}(\phi) Y_m^{(k)}(x) \\ &= \int_{S^d} \phi(y) \overline{Y_1^{(0)}(y)} d\mu(y) + \sum_{k=1}^n \sum_{m=1}^{d_k} \left( \int_{S^d} \phi(y) \overline{Y_m^{(k)}(y)} d\mu(y) \right) Y_m^{(k)}(x) \\ &= \int_{S^d} \sum_{k=0}^n \left( \sum_{m=1}^{d_k} \overline{Y_m^{(k)}(y)} Y_m^{(k)}(x) \right) \phi(y) d\mu(y) \\ &= \int_{S^d} \sum_{k=0}^n Z_x^{(k)}(y) \phi(y) d\mu(y) \\ &= \int_{S^d} K_n(x, y) \phi(y) d\mu(y), \end{aligned} \tag{2}$$

where

$$K_n(x, y) = \sum_{k=0}^n Z_x^{(k)}(y). \tag{3}$$

In terms of Jacoby polynomials (1) can be rewritten as

$$Z_\eta^{(k)}(x) = \sum_{m=1}^{d_k} Y_m^{(k)}(x) \overline{Y_m^{(k)}(\eta)} = C_k(S^d) P_k^{((d-2)/2, (d-2)/2)}(\langle x, \eta \rangle). \tag{4}$$

Putting  $x = \eta$  and integrating both sides we find

$$d_k = \sum_{m=1}^{d_k} \int_{S^d} |Y_m^{(k)}(x)|^2 d\mu(x) = C_k(S^d) P_k^{((d-2)/2, (d-2)/2)}(1). \tag{5}$$

Taking the square of both sides in (1) and then integrating with respect to  $x$  we get

$$\sum_{m=1}^{d_k} |Y_m^{(k)}(y)|^2 = (C_k(S^d))^2 \int_{S^d} \left( P_k^{((d-2)/2, (d-2)/2)}(\langle x, y \rangle) \right)^2 d\mu(x) \tag{6}$$

and since  $d\mu$  is shift invariant then

$$\int_{S^d} \left( P_k^{((d-2)/2, (d-2)/2)}(\langle x, y \rangle) \right)^2 d\mu(x) = c^{-1} \|P_k^{((d-2)/2, (d-2)/2)}(t)\|_2^2,$$

where

$$\|P_k^{((d-2)/2, (d-2)/2)}(t)\|_2^2 = \int_{-1}^1 (P_k^{((d-2)/2, (d-2)/2)}(t))^2 (1-t^2)^{(d-2)/2} dt$$

and

$$c = \int_{-1}^1 (1-t^2)^{(d-2)/2} dt = 2^{d-1} \frac{(\Gamma(d/2))^2}{\Gamma(d)} \quad (7)$$

due to the normalization condition  $\int_{S^d} d\mu = 1$ . To get the value of the constant  $c$  we apply a well known formulae for the Euler integral of the first kind

$$B(p, q) = \int_0^1 \xi^{p-1} (1-\xi)^{q-1} d\xi = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}, \quad p > 0, \quad q > 0.$$

Comparing (4) and (5) - (7) we obtain

$$Z_\eta^{(k)}(x) = \frac{c P_k^{((d-2)/2, (d-2)/2)}(1)}{\|P_k^{((d-2)/2, (d-2)/2)}(t)\|_2^2} P_k^{((d-2)/2, (d-2)/2)}(\langle x, \eta \rangle).$$

and therefore,

$$K_n(x, y) = \sum_{k=0}^n Z_y^{(k)}(x) = \sum_{k=0}^n \frac{c P_k^{((d-2)/2, (d-2)/2)}(1)}{\|P_k^{((d-2)/2, (d-2)/2)}(t)\|_2^2} P_k^{((d-2)/2, (d-2)/2)}(\langle x, y \rangle)$$

by (3). Let us denote

$$K_n^{((d-2)/2, (d-2)/2)}(\gamma, \delta) = \sum_{k=0}^n \frac{P_k^{((d-2)/2, (d-2)/2)}(\gamma) P_k^{((d-2)/2, (d-2)/2)}(\delta)}{\|P_k^{((d-2)/2, (d-2)/2)}(t)\|_2^2},$$

then [4] p.71,

$$K_n^{((d-2)/2, (d-2)/2)}(\gamma, 1) = 2^{-d-3} \frac{\Gamma(n+d)}{\Gamma(d/2)\Gamma(n+d/2)} P_n^{(d/2, (d-2)/2)}(\gamma)$$

and

$$K_n(x, \mathbf{e}) = c 2^{-d-3} \frac{\Gamma(n+d)}{\Gamma(d/2)\Gamma(n+d/2)} P_n^{(d/2, (d-2)/2)}(\langle x, \mathbf{e} \rangle),$$

where  $\mathbf{e}$  is the north pole on  $S^d$ , Using (7) we get

$$K_n(x, \mathbf{e}) = \frac{\Gamma(d/2)}{16\Gamma(d)} \frac{\Gamma(n+d)}{\Gamma(n+d/2)} P_n^{(d/2, (d-2)/2)}(\langle x, \mathbf{e} \rangle). \quad (8)$$

Clearly, by (8)

$$\begin{aligned} \|S_n\|_{L_\infty(S^d) \rightarrow L_\infty(S^d)} &= \sup_{\|\phi\|_{L_\infty(S^d)} \leq 1} \|S_n(\phi, x)\|_{L_\infty(S^d)} \\ &= \sup \left\{ \int_{S^d} |K_n(x, y)| d\mu(y) : x \in S^d \right\} \\ &= \int_{S^d} |K_n(x, \mathbf{e})| d\mu(y) \\ &= \frac{\Gamma(d/2)}{16\Gamma(d)} \frac{\Gamma(n+d)}{\Gamma(n+d/2)} \int_{S^d} |P_n^{(d/2, (d-2)/2)}(\langle x, \mathbf{e} \rangle)| d\mu(x) \\ &= \frac{\Gamma(d/2)}{16\Gamma(d)} \frac{\Gamma(n+d)}{\Gamma(n+d/2)} c^{-1} \int_{-1}^1 |P_n^{(d/2, (d-2)/2)}(t)| (1-t^2)^{(d-2)/2} dt \\ &= \frac{1}{2^{d+3}\Gamma(d/2)} \frac{\Gamma(n+d)}{\Gamma(n+d/2)} \int_{-1}^1 |P_n^{(d/2, (d-2)/2)}(t)| (1-t^2)^{(d-2)/2} dt, \\ &= \frac{1}{2^{d+3}\Gamma(d/2)} \frac{\Gamma(n+d)}{\Gamma(n+d/2)} \int_0^\pi |P_n^{(d/2, (d-2)/2)}(\cos \theta)| (\sin \theta)^{d-1} d\theta \end{aligned} \quad (9)$$

since  $K_n$  is a zonal function and  $d\mu$  is shift invariant. It is known [4] p.196 that for  $0 < \theta < \pi$ ,

$$P_n^{(d/2, (d-2)/2)}(\cos \theta) = \frac{1}{\sqrt{n}} k(\theta) \cos(N\theta + \gamma) + O(n^{-3/2}), \quad n \rightarrow \infty,$$

where

$$\begin{aligned} k(\theta) &= \pi^{-1/2} \left( \sin \frac{\theta}{2} \right)^{-d/2-1/2} \left( \cos \frac{\theta}{2} \right)^{-d/2+1/2} \\ &= \pi^{-1/2} 2^{(d+1)/2} (\sin \theta)^{-d/2-1/2} \cos \left( \frac{\theta}{2} \right) \end{aligned}$$

and

$$N = n + \frac{d-1}{2}, \quad \gamma = -\frac{d-1}{4}\pi.$$

From (9) it follows now that

$$\|S_n\|_{L_\infty(S^d) \rightarrow L_\infty(S^d)} = \kappa_n \frac{\Gamma(n+d)}{\sqrt{n} \Gamma(n+d/2)} + O(n^{(d-4)/2}), \quad n \rightarrow \infty, \quad (10)$$

where

$$\kappa_n = \frac{1}{\sqrt{2^{d+5}\pi} \Gamma(d/2)} \int_0^\pi \left| \cos \left( \left( n + \frac{d-1}{2} \right) \theta - \frac{d-1}{4} \pi \right) \right| \cos \left( \frac{\theta}{2} \right) (\sin \theta)^{(d-3)/2} d\theta,$$

Finally, it is easy to check that

$$\frac{\Gamma(n+d)}{\sqrt{n} \Gamma(n+d/2)} = n^{(d-1)/2} (1 + O(1/n)), \quad n \rightarrow \infty$$

and

$$\begin{aligned} & \int_0^\pi \left| \cos \left( \left( n + \frac{d-1}{2} \right) \theta - \frac{d-1}{4} \pi \right) \right| \cos \left( \frac{\theta}{2} \right) (\sin \theta)^{(d-3)/2} d\theta \\ &= \frac{2}{\pi} \int_0^\pi \cos \left( \frac{\theta}{2} \right) (\sin \theta)^{(d-3)/2} d\theta + O \left( \frac{1}{n} \right), \quad n \rightarrow \infty, \end{aligned}$$

so that

$$\begin{aligned} & \|S_n\|_{L_\infty(S^d) \rightarrow L_\infty(S^d)} \\ &= n^{(d-1)/2} \left( \frac{1}{2^{d/2+3/2}\pi^{3/2}\Gamma(d/2)} \int_0^\pi \cos \left( \frac{\theta}{2} \right) (\sin \theta)^{(d-3)/2} d\theta \right) (1 + O(1/n)), \quad n \rightarrow \infty. \end{aligned}$$

### Referências Bibliográficas:

1. ERDÉLYI, A. (1953), **Higher Transcendental Functions**, Vol. 2, McGraw-Hill, New York, (1953).
2. GRONWALL, T.H., **On the degree of convergence of Laplace series**, Trans. Amer. Math. Soc., vol. 15, 1, (1914), 1-30.
3. RAGOZIN, D.L., **Uniform convergence of spherical harmonic expansions**, Math. Ann., vol. 195, (1972), 87-94.
4. SZEGÖ, G., **Orthogonal Polynomials**, AMS, New York, (1939).
5. VLLENKIN, N.J. (1965), **Special functions and theory of representation of groups**, Nauka, Moscow, (1965).

# UM ESTUDO NUMÉRICO DO PROBLEMA DE QUADRADOS MÍNIMOS NÃO LINEARES COM RESTRIÇÕES

Mario Cesar Zambaldi, Juliano de Bem Francisco.

Universidade Federal de Santa Catarina - UFSC,

zambaldi@mtm.ufsc.br; juliano@mtm.ufsc.br

## Resumo:

Neste trabalho fazemos um estudo numérico para o problema de quadrados mínimos não linear com Restrições também não lineares de igualdade. Embora seja um problema de otimização, exploramos suas particularidades e características específicas considerando principalmente os sistemas não lineares subjacentes. Executamos alguns testes com um modelo das equações da rede elétrica.

**Palavras-chave:** Otimização Numérica, Quadrados Mínimos Não Lineares, Sistemas Não Lineares.

## Introdução

Um dos problemas centrais em Otimização é a resolução de sistemas de equações não lineares. Inúmeras aplicações nas ciências aplicadas e engenharias consideraram modelos matemáticos deste tipo. Em alguns casos, o sistema pode não ter solução por várias razões. Mesmo assim, encontrar uma solução alternativa que minimiza o erro e que tenha um sentido físico é o problema que se procura resolver. Isto caracteriza o Problema de Quadrados Mínimos não Lineares.

Mesmo podendo ser considerado um problema de minimização de uma função dada pela soma dos quadrados dos resíduos do sistema, o problema de quadrados mínimos



possui características específicas que devem ser exploradas. É possível ainda que o sistema deva obedecer certas restrições imposta pelo modelo, restrições esta que podem ser não lineares.

Neste trabalho vamos considerar a resolução de um modelo cuja formulação consiste num Problema de Quadrados Mínimos não Lineares que apresentam restrições. A ênfase será dada aos aspectos numéricos do mesmo assim como sua relação com o sistema de equações não lineares que o resultou. Fazemos o estudo considerando um problema de engenharia em redes de potência para restauração da solução das equações de uma rede elétrica.

## Métodos de Quadrados Mínimos não Lineares Restrito

O problema de quadrados mínimos não lineares pode ser escrito da forma:

$$\begin{aligned} \text{Min} \quad & \frac{1}{2} \|f(x)\|_2^2 \\ \text{s.a} \quad & g(x) = 0 \end{aligned} \tag{1}$$

onde  $f(x) : \mathbb{R}^n \rightarrow \mathbb{R}^m$  e  $g(x) : \mathbb{R}^n \rightarrow \mathbb{R}^k$  são funções não lineares.

Considerando-o como um problema típico de otimização escrevemos a função Lagrangeana:

$$L(x, \lambda) = \frac{1}{2} f(x)^T f(x) + \lambda^T g(x) \tag{2}$$

cuja condição de primeira ordem é

$$\nabla L(x^*, \lambda^*) = 0.$$

Definindo

$$R(x, \lambda) \doteq \nabla L(x, \lambda) = \begin{bmatrix} \nabla_x L(x, \lambda) \\ \nabla_\lambda L(x, \lambda) \end{bmatrix}$$

tomamos,

$$\begin{aligned} R_1(x, \lambda) &\doteq \nabla_x L(x, \lambda) = F(x)^T f(x) + G(x)^T \lambda \\ R_2(x, \lambda) &\doteq \nabla_\lambda = g(x) \end{aligned} \quad (3)$$

onde  $F(x) = \frac{\partial f(x)}{\partial x}$  e  $G(x) = \frac{\partial g(x)}{\partial x}$ .

Portanto,  $(x^*, \lambda^*)$  deve satisfazer :

$$F(x^*)^T f(x^*) + G(x^*)^T \lambda^* = 0 \quad (4)$$

$$g(x^*) = 0 \quad (5)$$

Neste momento temos então um sistema de equações não lineares dado pelo gradiente do Lagrangeano. Podemos então, utilizar o método de Newton para obter a solução  $(x^*, \lambda^*)$ .

O Jacobiano de  $R(x, \lambda)$  é dado por:

$$W(x, \lambda) = \begin{bmatrix} R_{1x}(x, \lambda) & R_{1\lambda}(x, \lambda) \\ R_{2x}(x, \lambda) & R_{2\lambda}(x, \lambda) \end{bmatrix} \quad (6)$$

sendo

$$\begin{aligned} R_{1x}(x, \lambda) &= J(x)^T J(x) + \sum_{i=1}^m f_i(x) \nabla^2 f_i(x) + \sum_{i=1}^k \lambda_i \nabla^2 g_i(x); \\ R_{1\lambda} &= G(x)^T; \\ R_{2x} &= G(x); \\ R_{2\lambda} &= O \in \mathbb{R}^{k \times k}. \end{aligned} \quad (7)$$

Na  $k$ -ésima iteração de Newton tem-se que resolver o seguinte sistema linear:

$$\begin{bmatrix} R_{1x}(x_k, \lambda_k) & R_{1\lambda}(x_k, \lambda_k) \\ R_{2x}(x_k, \lambda_k) & R_{2\lambda}(x_k, \lambda_k) \end{bmatrix} \begin{bmatrix} \Delta x_k \\ \Delta \lambda_k \end{bmatrix} = - \begin{bmatrix} R_1(x_k, \lambda_k) \\ R_2(x_k, \lambda_k) \end{bmatrix} \quad (8)$$

e assim obtem-se:

$$x_{k+1} = x_k + \alpha_k \Delta x_k \quad (9)$$

$$\lambda_{k+1} = \lambda_k + \alpha_k \Delta \lambda_k \quad (10)$$

onde  $\alpha_k$  é a correção do passo de Newton obtido pela busca linear na direção obtida por (8).

A busca linear é um dos possíveis métodos de globalização do método de Newton. Neste trabalho, parâmetro  $\alpha$  é obtido pela minimização unidimensional da seguinte função de mérito:

$$M(x, \lambda) \doteq \frac{1}{2} R(x, \lambda)^T R(x, \lambda).$$

Quando os termos de segunda ordem de (8) são nulos ou suficientemente pequenos, pode-se empregar o método de Gauss-Newton [7]. Entretanto este não é o caso do presente trabalho, onde os termos do Hessiano relativos à função  $f$  são relevantes.

## O Problema das Equações da Rede Elétrica

Considere um sistema elétrico com  $n_b$  barras e matriz de admitância dada por  $Y = G + jB$ , em que  $j = \sqrt{-1}$ . Os elementos desta matriz são constantes matemáticas e apresentam sua própria representatividade intrínseca no sistema elétrico. Vamos denotar as potências ativa e reativa de demanda, e as potências ativa e reativa geradas, respectivamente, por  $P_i^d$ ,  $Q_i^d$ ,  $P_i^g$  e  $Q_i^g$ , as quais são fornecidas pelo problema. Portanto, definindo  $c_{ik} = \cos(\theta_i - \theta_k)$  e  $s_{ik} = \sin(\theta_i - \theta_k)$ , as equações do problema de fluxo de carga em coordenadas polares podem ser representadas pelo seguinte conjunto de

equações não-lineares:

$$\begin{aligned} P_i(V, \theta) &= [G_{ii}V_i^2 + V_i \sum_{k \in \Omega_i} V_k (G_{ik}c_{ik} + B_{ik}s_{ik})] - P_i^g + P_i^d \\ Q_i(V, \theta) &= [-B_{ii}V_i^2 + V_i \sum_{k \in \Omega_i} V_k (G_{ik}s_{ik} + B_{ik}c_{ik})] - Q_i^g + Q_i^d, \end{aligned}$$

em que  $i = 1, \dots, n_b$ . As variáveis do problema são  $V \in \mathbb{R}^{n_b}$  e  $\theta \in \mathbb{R}^{n_b}$ , que representam, respectivamente, a magnitude de tensão e o ângulo de tensão nodal. O conjunto de índices  $\Omega_i$  define a topologia do sistema, isto é, a conexão da barra  $i$  com as demais, que por sua vez interfere diretamente na estrutura esparsa das matrizes associadas ao problema. Maiores detalhes desta formulação podem ser obtidas em [6].

Em sistemas elétricos com carregamento excessivo várias equações do sistema acima não se realizam, dando origem a um problema de mínimos quadrados não-linear. Em virtude de solicitações do sistema elétrico, algumas equações não devem apresentar resíduos, sendo, portanto, incorporadas como restrições de igualdade no problema. Nesta formulação, atribuindo  $x = (V_1, \dots, V_{n_b}, \theta_1, \dots, \theta_{n_b}) \in \mathbb{R}^{2n_b}$  às variáveis, o problema central deste trabalho pode ser modelado como um problema de mínimos quadrados não-linear restrito,

$$\begin{aligned} \min \quad & \frac{1}{2} \|F(x)\|^2 \\ \text{s.a} \quad & h(x) = 0, \end{aligned} \tag{11}$$

em que  $F : \mathbb{R}^n \rightarrow \mathbb{R}^p$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$  e  $p + m = n$ . Se o sistema elétrico operar dentro dos limites de demanda, o problema se resume a encontrar  $x^*$  tal que

$$\begin{bmatrix} F(x^*) \\ h(x^*) \end{bmatrix} = 0.$$

## Solução Numérica

Os testes numéricos foram realizados com o MATLAB 7.0 [8]. Os resultados para os sistemas 30 e 57 barras são apresentados nas tabelas 1 e 2 respectivamente, que mostram a evolução da convergência do gradiente do Lagrangeano em cada iteração.

Tabela 1: Sistema de 30 barras

iteração	grad. do Lagrangeano
1	1.0494e+01
2	1.0141e+00
3	3.4897e-01
4	1.1138e-01
5	2.9974e-02
6	5.6653e-03
7	1.1400e-03
8	6.4668e-05

Tabela 2: Sistema de 57 barras

iteração	grad. do Lagrangeano
1	2.1407e+02
2	7.5749e+01
3	1.9875e+01
4	1.3723e+01
5	7.4064e+00
6	4.5129e+00
7	2.4919e+00
8	1.3279e+00
9	5.9365e-01
10	6.9662e-01
11	2.2048e-01
12	3.0670e-02
13	1.6098e-02
14	7.8667e-03
15	2.3935e-03
16	2.6047e-04

Podemos observar que o processo de convergência fica gradativamente mais lento a medida que a dimensão aumenta, com 8 iterações em 30 barras e 16 iterações em 57 barras.

Um fato relevante a ser destacado é que, para as dimensões acima reportadas, o método de Newton não exigiu o procedimento de busca, ou seja,  $\alpha_k = 1$  sempre foi aceito como parâmetro inicial sendo aceito para o teste de descida suficiente. Isto não

ocorre em geral com dimensões maiores, onde o procedimento de busca geralmente é requisitado (veja [1]).

Apesar do passo completo de Newton ser aceito, esperar-se-ia que o processo de convergência apresentasse ordem quadrática [3]. Este fato tem repercussões tanto teóricas quanto práticas. Do lado teórico, este comportamento linear da convergência do método de Newton nos leva a supor que a Hessiana do Lagrangeano é singular na solução [3]. Este fato também nos leva a cogitar como estaria se comportando o ângulo das direções de Newton com o gradiente do Lagrangeano, dado que o avanço de cada iteração é relativamente pequeno.

Do ponto de vista prático, as observações acima induz um estudo da numérico e da estrutura do problema. Na prática, pelo comportamento dos resultados, deve-se verificar um mal condicionamento da matriz Jacobiana, em paralelo à sua singularidade teórica.

O emprego de um método iterativo linear, considerando futuras abordagens ao problema com dimensão maior poderia ser aconselhável. Entretanto, os resultados acima mostram, em princípio, que esta estratégia pode não ser interessante, pois um preconditionador baseado numa fatoração incompleta de uma matriz mal condicionada dificilmente apresentará bons resultados [5]. Uma observação na estrutura da matriz é necessária para afirmar com mais clareza sobre usar ou não os métodos iterativos.

## **Conclusões e Alternativas**

Neste trabalho descrevemos o método de resolução do problema de quadrados mínimos não lineares com restrições. Descrevemos também um importante problema em redes de potência de energia elétrica cuja formulação do modelo conduz ao problema matemático citado. Apresentamos resultados numéricos que comprovam a recuperação das soluções como era o objetivo do modelo. Efetuamos um estudo numérico na expectativa de fornecer uma explicação da taxa de convergência do método de Newton, adequado à mínimos quadrados restritos. Podemos assim detectar que o número de iterações do

método de Newton cresce quando a dimensão aumenta. Isso parece ser indicativo de mal condicionamento dos sistemas. De fato o referido método apresenta convergência apenas linear quando o Hessiano é singular na solução. Este fato tem várias implicações. Uma delas é que mesmo considerando um sistema esparso, não é conveniente usar métodos iterativos lineares para a etapa de linearização. O problema do mal condicionamento tende a se agravar com o aumento da dimensão do problema, tornando a implementação numérica cada vez mais complexa e exigente. Um estudo numérico mais aprofundado é necessário para identificar a origem do mal condicionamento do sistema e assim melhorar o desempenho do método.

Uma alternativa interessante para o problema em questão consiste em trabalhar com a função Lagrangeano Aumentado definida por:

$\mathcal{L}_A : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^{++} \rightarrow \mathbb{R}$  para o problema (11),

$$\mathcal{L}_A(x, \lambda, \rho) = \frac{1}{2} \|F(x)\|^2 + h(x)^T \lambda + \frac{\rho}{2} \|h(x)\|^2. \quad (12)$$

Nesse método o passo principal consiste em resolver uma sequência de subproblemas do tipo: encontre  $x_l$  tal que

$$\|\nabla_x \mathcal{L}_A(x_l, \lambda_l, \rho_l)\| \leq \tau_l. \quad (13)$$

em que  $\lambda_l$  é o multiplicador e  $\tau_l$  é uma tolerância definidos a priori nas iteração externa.

Uma vez fixado  $\lambda_l \in \mathbb{R}^m$  e  $\rho_l \in \mathbb{R}^{++}$ , resolver o problema:

$$\min_{x \in \mathbb{R}^n} \mathcal{L}_A(x, \lambda_l, \rho_l)$$

é equivalente (veja [4]) a

$$\min_{x \in \mathbb{R}^n} \frac{1}{2} \left\| \begin{bmatrix} F(x) \\ \sqrt{\rho_l} h(x) \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\lambda_l}{\sqrt{\rho_l}} \end{bmatrix} \right\|^2, \quad (14)$$

que é um problema de mínimos quadrados não-linear. Uma alternativa interessante usar o mtodo de Levenberg-Marquardt [4]. A grande vantagem desta estratégia é não demandar iformação de segunda ordem como exige o método original. Nessa direção que encaminharemos as futuras investigações.

# Referências Bibliográficas

- [1] Barbosa L. V. [2001]: Análise e Desenvolvimento de Metodologias Corretivas para a Restauração da Solução das Equações da Rede Elétrica. Tese de Doutorado. Dept. Eng. Elétrica, Universidade Federal de Sta. Catarina - UFSC.
- [2] Barbosa L. and Salgado R. [2000]: Corrective Solutions of Steady State Power Systems via Newton Optimization Method. Revista SBA Controle e & automação, Vol. 11, p 182-186
- [3] Dennis, J. E. jr.; Schnabel, R. B. [1996]: Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM Philadelphia.
- [4] J. Fan e Y. Yuan. On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption, *Computing*, **74** (2005), 23-39.
- [5] Golub G. A. and Van Loan C. F. [1996]: Matrix Computations. 3rd. Edition. The John Hopkins University Press Ltda. London.
- [6] Monticelli, Alcir José. [1983]: Fluxo de Carga em Rede de Energia Elétrica. Edgard Blücher, São Paulo.
- [7] Nocedal J. and Wright J. [1999]: Numerical Optimization 1999 Springer Series in Operations Research, Springer Verlag New York Inc.
- [8] Pärt-Enander E. and Sjöberg A. [1999]: The Matlab Handbook 5. Addison Wesley, Harlow UK.



# CONSTRUÇÃO DOS NÚMEROS REAIS ATRAVÉS DOS CORTES DE DEDEKIND

Emerson Lazzarotto, Marcos de Abreu dos Santos.  
Universidade Estadual do Oeste do Paraná – UNIOESTE,  
emerson.lazzarotto@gmail.com; Universidade Estadual do Oeste do Paraná – UNIOESTE,  
abreu\_mat@hotmail.com.

**Resumo:** O progresso da humanidade e da ciência em geral estimulou a extensão dos conjuntos numéricos. Estudamos em matemática que o conjunto dos números racionais ( $\mathcal{Q}$ ) não é completo do ponto de vista algébrico, uma vez que é possível resolver equações do tipo  $ax + b = 0$  sempre que  $a \neq 0$ , porém, equações como  $x^2 - 2 = 0$  não tem solução em  $\mathcal{Q}$ . Pretendemos, nesse trabalho, mostrar como são construídos os conjuntos numéricos a partir do conjunto dos números naturais e quais propriedades algébricas são ganhas com tais construções. Em particular, destacamos a forma de definir um número real através do método de cortes de Dedekind no conjunto dos números racionais.

**Palavras-chave:** Axiomas de Peano, Número Real, Cortes de Dedekind.

## 1. Números Naturais ( $\mathfrak{N}$ )

O referencial teórico usado em todo processo de construção de conjuntos numéricos adotado neste trabalho foi extraído de Lobeiro (2000), incrementado pelas visualizações geométricas destes conjuntos.

Atualmente é difícil encontrar uma situação que não esteja direta ou indiretamente relacionada com números ou a questão da quantificação de uma ou mais variáveis.

Os números naturais surgem da necessidade de contagem e de enumerar objetos. A construção axiomática desse conjunto nos possibilita a construção de todos os demais, uma vez que serve de base para a construção do conjunto dos inteiros ( $Z$ ), que por sua vez servirá de base para a construção do conjunto  $Q$  e, por fim, chegar ao conjunto  $\mathfrak{R}$ .

Toda teoria dos números naturais pode ser deduzida dos três axiomas abaixo, conhecidos como *axiomas de Peano*, atribuído ao italiano Giuseppe Peano.

- I. Existe um conjunto  $\mathfrak{N}$  tal que  $1 \in \mathfrak{N}$
- II. Para todo  $n \in \mathfrak{N}$  existe um único sucessor ( $n'$ ) para  $n$ , para cada  $n \in \mathfrak{N}$ ,  $n' \neq 1$  e se  $m' = n'$  então  $m = n$ ,
- III. Se  $A \subset \mathfrak{N}$ ,  $1 \in A$  e  $n' \in A$  sempre que  $n \in A$  então  $A = \mathfrak{N}$

Os axiomas acima asseguram a existência de um conjunto cujo primeiro elemento é simbolizado pelo número 1, que não é sucessor de nenhum outro natural, que se dois elementos possuem o mesmo sucessor, então esses elementos são iguais e por fim o princípio de indução, que estabelece que qualquer número natural pode ser atingido começando com o 1 e contando os sucessores consecutivos. Os símbolos 2,3,4,... servem, respectivamente, para indicar o sucessor de 1,2,3,...

A partir disso, define-se a adição de naturais da seguinte forma:  $n+1=n'$  e  $m+n'=(m+n)'$  sempre que  $n+m$  está definido. A adição nos naturais assim definida é fechada e goza das propriedades associativa, comutativa e satisfaz a lei do corte, isto é,  $m+n=m+p \Rightarrow n=p$ , propriedades que se verificam, essencialmente, pelo princípio de indução deduzida imediatamente do axioma III. Em seguida, é definida uma ordem no conjunto dos naturais:  $m < n$  se, e somente se, existe  $p \in \mathfrak{N}$  tal que  $m+p=n$ . Disto segue a transitividade:  $m < n, n < p \Rightarrow m < p$ , monotonicidade:  $m < n \Leftrightarrow m+p < n+p$  e a tricotomia:  $m < n$ ,  $m = n$  ou  $n < m$  e o princípio do menor elemento: todo subconjunto não

vazio  $A \subset \mathbb{N}$  possui menor elemento. O Princípio do Menor Elemento é equivalente ao Princípio de Indução (Santos, 2000).

Define-se também a multiplicação de números naturais:  $n \cdot 1 = n$ , para todo  $n \in \mathbb{N}$  e  $n \cdot m' = n \cdot m + n$ , sempre  $n \cdot m$  está definido, tal operação além de fechada nos permite a verificação de todas as propriedades usuais como distributiva, associativa e comutativa.

## 2. Números Inteiros ( $Z$ )

Para construir  $Z$  a partir de  $\mathbb{N}$  acrescentamos os simétricos aos números naturais em relação à adição, essa motivação pode partir da resolução da equação  $m + x = n$  que, em  $\mathbb{N}$ , só tem solução para  $m < n$ . Com a construção de  $Z$  tal equação terá sempre solução, embora a natureza de seus elementos seja distinta da natureza dos números naturais, como veremos a seguir, eles podem ser vistos como generalizações dos naturais.

A relação de equivalência  $R$  em  $\mathbb{N} \times \mathbb{N}$  definida por  $(a,b)R(c,d) \Leftrightarrow a + d = b + c$  induz uma partição de  $\mathbb{N} \times \mathbb{N}$  e seu conjunto quociente  $\mathbb{N} \times \mathbb{N} / R = \{\overline{(a,b)} : (a,b) \in \mathbb{N} \times \mathbb{N}\}$  é definido como sendo o conjunto dos números inteiros, denotado por  $Z$ . Por exemplo, a classe do elemento  $(1,1)$  é  $\overline{(1,1)} = \{(n,n) : n \in \mathbb{N}\}$  e a classe do elemento  $(2,1)$  é  $\overline{(2,1)} = \{(n+1,n) : n \in \mathbb{N}\}$ . Desta forma, um número inteiro é, na verdade, um conjunto.

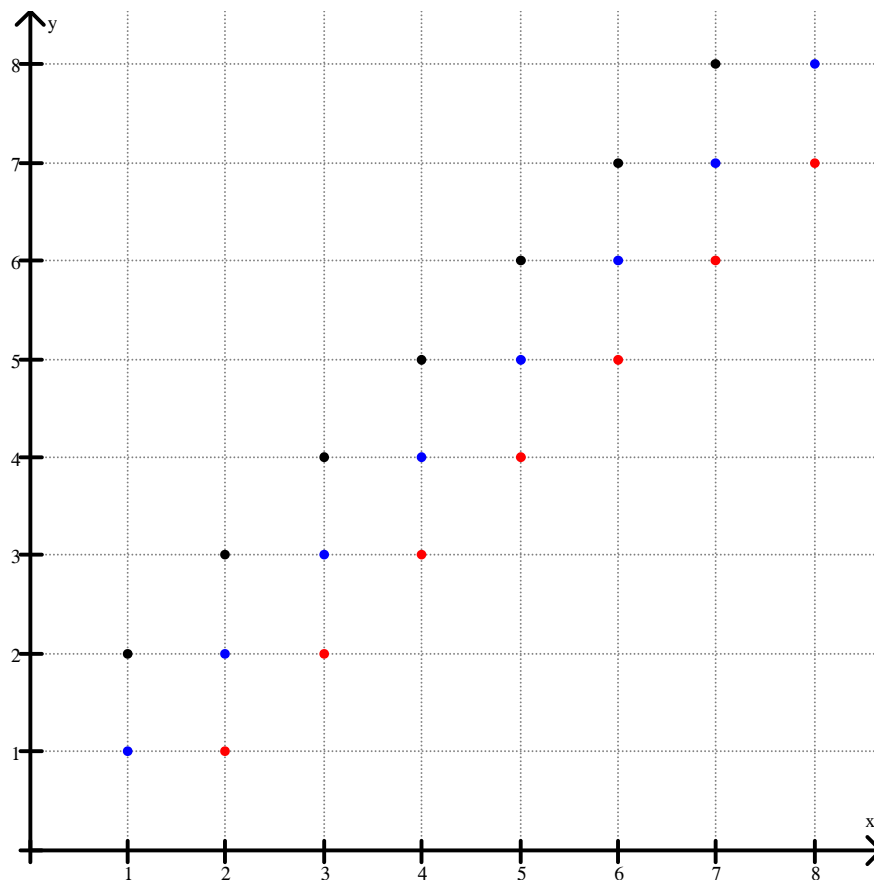


Figura 1: visualizando um número inteiro

A figura 1 ilustra alguns números inteiros. Cada cor representa um conjunto de pontos que estão numa mesma classe da relação de equivalência. Mais adiante, vamos identificar os conjuntos de pontos pretos, azuis e vermelhos, respectivamente, pelos números inteiros  $-1, 0$  e  $1$ . Uma classe da relação de equivalência ou, um número inteiro, será designado daqui por diante simplesmente por  $x = \overline{(a,b)}$  ou  $y = \overline{(c,d)}$ . A adição em  $Z$  é definida por  $x + y = \overline{(a+c, b+d)}$  e a multiplicação em  $Z$  é definida por  $x \cdot y = \overline{(ac+bd, ad+bc)}$ . É importante destacar, neste ponto, que a adição definida é a usual (soma coordenada com coordenada) e a multiplicação não é a usual. Há duas razões para isto: a primeira é que, pela natureza dos elementos do conjunto dos inteiros (são, por si só, conjuntos), tais operações devem ser independentes do elemento que representa a classe e, a segunda, é que as propriedades que se desejam obter não seriam válidas caso fossem

utilizadas outras operações. A adição, acima definida, é associativa, comutativa, possui como elemento neutro a classe  $\{(n,n) : n \in \mathbb{N}\}$  (que será denotada por  $0$ ) e, para cada classe,  $x = \overline{(a,b)}$ , a classe  $\overline{(b,a)}$  que será representada, convenientemente, por  $-x$  (oposto aditivo), é tal que  $x + (-x) = 0$ . Quanto à multiplicação, ela é associativa, comutativa, distributiva em relação à adição. São válidas as tradicionais leis do corte nas duas operações, as regras de sinais e a propriedade que confere estrutura de anel de integridade aos inteiros, isto é,  $x \cdot y = 0 \Rightarrow x = 0$  ou  $y = 0$ . A ordem em  $Z$  é dada por:  $x < y \Leftrightarrow a + d < b + c$ . Mais uma vez, é importante destacar que esta ordem não usual, (a usual seria  $x < y \Leftrightarrow a < c$  ou  $a = c$  e  $b < d$ ), é fundamental para independer do representante tomado na classe. Dentre uma série de propriedades que são obtidas através da ordem definida em  $Z$ , podemos definir um inteiro positivo  $x$  quando  $x > 0$ , denotados por  $Z_+$ , vamos mostrar que  $Z_+ = \{x = \overline{(a,b)} : b < a\}$ . De fato,  $0 = \overline{(n,n)} < \overline{(a,b)} = x \Leftrightarrow n + b < n + a \Leftrightarrow b < a$ . A figura a seguir ilustra alguns números inteiros positivos.

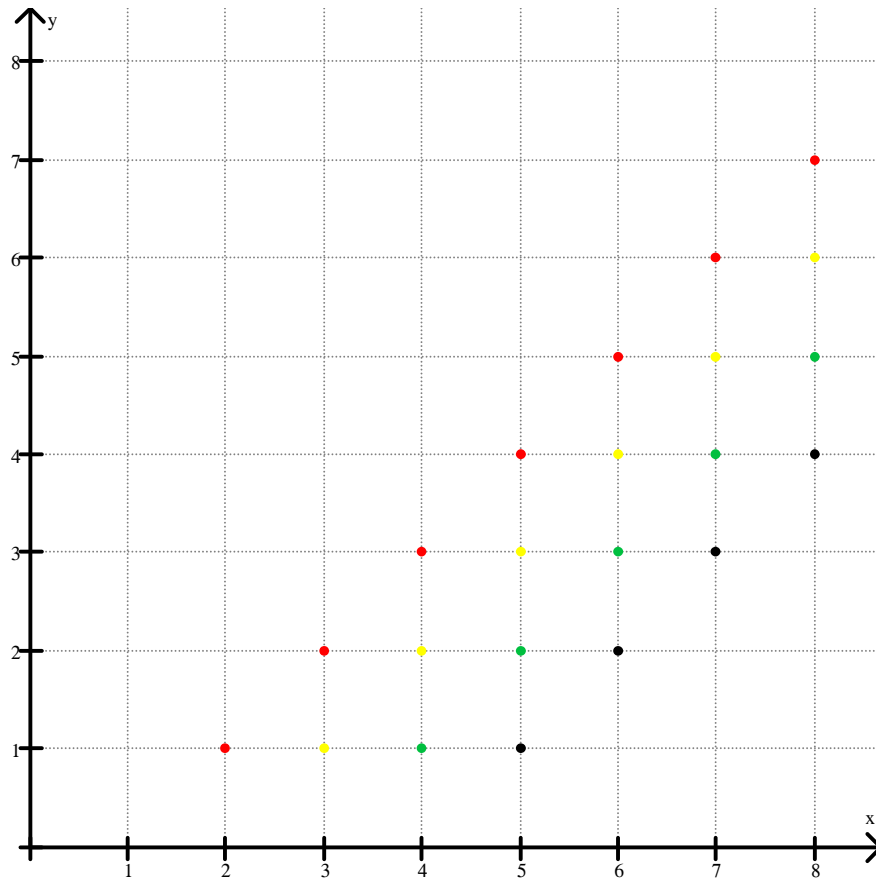


Figura 2: inteiros positivos

Cada conjunto de pontos de uma cor representa um inteiro positivo. Agora, é possível identificar o conjunto dos naturais com os inteiros positivos, considerando a aplicação  $\phi: \mathfrak{N} \rightarrow Z_+$  dada por  $\phi(n) = \overline{(n+1,1)}$ . Primeiramente,  $\phi$  é bijetora, pois:

$$\phi(m) = \phi(n) \Rightarrow \overline{(m+1,1)} = \overline{(n+1,1)} \Rightarrow (m+1,1)R(n+1,1) \Rightarrow m+1+1 = n+1+1 \Rightarrow m = n$$

Dado  $x = \overline{(a,b)} \in Z_+$ ,  $b < a$ , existe  $n \in \mathfrak{N}$  tal que  $b+n = a$ , portanto,  $n+1+b = a+1 \Rightarrow \overline{(n+1,1)} = \overline{(a,b)} = x \Rightarrow \phi(n) = x$ .

A aplicação  $\phi$  também é um isomorfismo:

$$\phi(m+n) = \overline{(m+n+1,1)} = \overline{(m+1+n+1,1+1)} = \overline{(m+1,1)} + \overline{(n+1,1)} = \phi(m) + \phi(n)$$

$$\begin{aligned} \phi(m.n) &= \overline{(m.n+1,1)} = \overline{(m.n+m+n+1+1, m+n+1+1)} = \\ &= \overline{((m+1)(n+1)+1, (m+1)+(n+1))} = \overline{(m+1,1)}. \overline{(n+1,1)} = \phi(m).\phi(n) \end{aligned}$$

O fato de  $\phi$  ser um isomorfismo significa que a estrutura algébrica atribuída à adição e a multiplicação nos inteiros (um conjunto cujos elementos têm natureza distinta dos naturais) é compatível com as respectivas operações nos naturais, em outras palavras, as operações são preservadas quando usamos esta identificação de  $\mathfrak{N}$  com  $Z_+$ . Além disso,  $m < n \Rightarrow \exists p \in \mathfrak{N} : m + p = n \Rightarrow \phi(n) = \phi(m + p) = \overline{(m + p + 1, 1)} > \overline{(m + 1, 1)} = \phi(m)$ , ou seja,  $\phi$  preserva a ordem, isto quer dizer que, além das operações, as respectivas ordens nos conjuntos  $\mathfrak{N}$  e  $Z_+$  também são compatíveis. Representando por  $Z_-$  os inteiros  $\overline{(a, b)}$  tais que  $a < b$  pode-se decompor  $Z = Z_- \cup \{0\} \cup Z_+$  e, denotando por  $-n$  a classe  $\overline{(1, n+1)}$ , dizer que  $Z$  é igual, a menos de um isomorfismo, ao conjunto  $\{\dots, -n, \dots, -1, 0, 1, \dots, n, \dots\}$ .

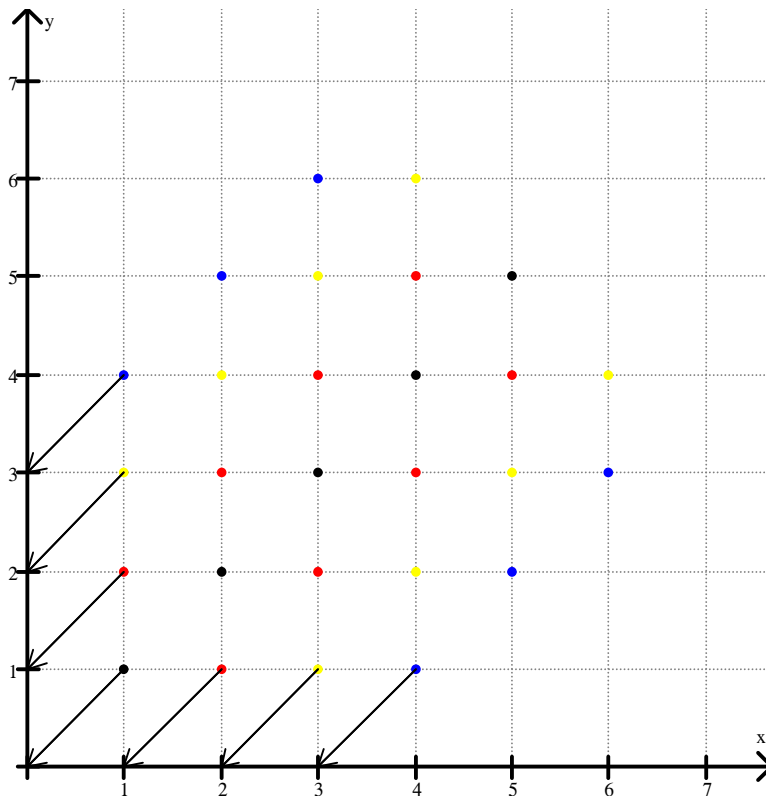


Figura 3: a identificação dos inteiros

A figura 3 mostra como identificamos os números inteiros com o convencional conjunto  $\{\dots, -n, \dots, -1, 0, 1, \dots, n, \dots\}$ . Por exemplo, a classe  $\overline{(1,2)} = \{(1,2), (2,3), (3,4), \dots, (n, n+1), \dots\}$  é identificada com o número  $-1$ , a classe  $\overline{(1,1)} = \{(1,1), (2,2), (3,3), \dots, (n, n), \dots\}$  é identificada com número  $0$ , enquanto que a classe  $\overline{(2,1)} = \{(2,1), (3,2), (4,3), \dots, (n+1, n), \dots\}$  é identificada com o número  $1$ . Uma reflexão de  $45^\circ$  dos eixos coordenados, mostra que a aplicação  $\phi$  consiste de uma identificação natural daquilo que “construímos” como número inteiro com o que manipulamos como número inteiro.

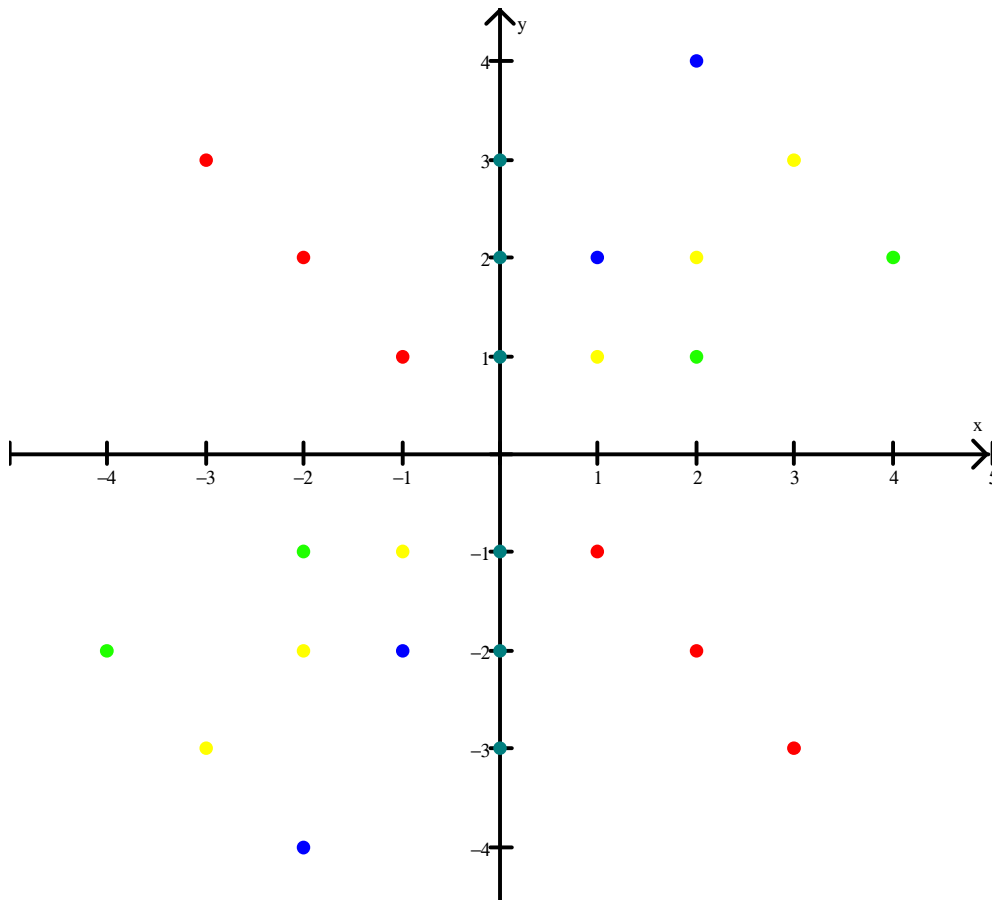
### 3. Os Números Racionais ( $\mathcal{Q}$ )

Mediante a identificação acima, daqui por diante, o conjunto dos inteiros  $Z$  será considerado  $\{\dots, -2, -1, 0, 1, 2, \dots\}$ . Embora não se tenha dado uma apresentação pormenorizada das propriedades da estrutura algébrica do conjunto dos inteiros, é sabido que, uma de suas “falhas” é que nem todo elemento não nulo é inversível para a multiplicação. A construção dos racionais vem preencher exatamente esta lacuna. Dada a relação de equivalência  $R$  em  $Z \times Z^*$  definida por  $(a,b)R(c,d) \Leftrightarrow ad = bc$ , ela induz uma partição de  $Z \times Z^*$  e, seu conjunto quociente,  $Z \times Z^* / R$ , denotado por  $\mathcal{Q}$  é chamado de conjunto dos números racionais. Mais precisamente,  $\mathcal{Q} = \{\overline{(a,b)} : a \in Z, b \in Z^*\}$ . Por exemplo,

$$\begin{aligned}\overline{(0,1)} &= \{(a,b) : (a,b)R(0,1)\} = \{(a,b) : a = 0\} = \{\dots, (0,-2), (0,-1), (0,1), (0,2), \dots\} \\ \overline{(1,1)} &= \{(a,b) : (a,b)R(1,1)\} = \{(a,b) : a = b\} = \{\dots, (-2,-2), (-1,-1), (1,1), (2,2), \dots\} \\ \overline{(-1,1)} &= \{(a,b) : (a,b)R(-1,1)\} = \{(a,b) : b = -a\} = \{\dots, (-2,2), (-1,1), (1,-1), (2,-2), \dots\}\end{aligned}$$

A figura a seguir ilustra alguns números racionais.





**Figura 4: Números racionais**

Cada conjunto de pontos de uma cor representa um número racional; cada número racional é um conjunto de pontos colineares e cada um destes conjuntos é colinear à origem, assim, é possível classificar os números racionais como o conjunto de “retas” (pontos colineares com coordenadas inteiras) que passa pela origem. Esta abordagem geométrica tem a vantagem de permitir afirmar, mesmo que intuitivamente, que os racionais são enumeráveis.

Uma classe da relação de equivalência, ou um número racional, será designado daqui por diante por  $r = \overline{(x, y)}$  e  $s = \overline{(w, z)}$ . Dessa maneira definimos a adição em  $\mathcal{Q}$  por  $r + s = \overline{(x, y)} + \overline{(w, z)} = \overline{(x \cdot z + y \cdot w, y \cdot z)}$  e a multiplicação em  $\mathcal{Q}$  por  $r \cdot s = \overline{(x, y)} \cdot \overline{(w, z)} = \overline{(x \cdot w, y \cdot z)}$ .

A adição é associativa, comutativa, possui como único neutro o elemento  $0 = \{(0, x) : x \in Z^*\}$  e, para cada  $r = \overline{(x, y)} \in Q$  existe uma única  $\overline{(-x, y)} \in Q$  que será denotada por  $-r$  (oposto aditivo) tal que  $r + (-r) = 0$ . A multiplicação é comutativa, associativa, distributiva em relação à adição, possui um único elemento neutro, a saber,  $1 = \{(x, x) : x \in Z^*\}$  e para todo  $r = \overline{(x, y)} \in Q$  existe um único elemento  $r^{-1} = \overline{(y, x)} \in Q$  tal que  $r \cdot r^{-1} = 1$ .

Assim como o conjunto  $Z$  são válidas as tradicionais leis do corte nas duas operações, as regras de sinais e a propriedade que confere estrutura de anel de integridade aos inteiros, isto é,  $r \cdot s = 0 \Rightarrow r = 0$  ou  $s = 0$ .

A ordem em  $Q$  é definida a seguir: seja  $r = \overline{(x, y)} \in Q$ , diremos que  $r$  é maior que 0 e, escrevemos  $r > 0$ , se existir  $(x, y) \in r$  tal que  $x \cdot y > 0$ .

Se considerarmos a aplicação  $\Phi : Z \rightarrow Q$  dada por  $\Phi(x) = \overline{(x, 1)}$  é possível identificar os inteiros como subconjunto dos racionais.

Note que  $\Phi$  é injetora, pois:

$$\Phi(x) = \Phi(y) \Rightarrow \overline{(x, 1)} = \overline{(y, 1)} = (x, 1)R(y, 1) \Rightarrow x \cdot 1 = 1 \cdot y \Rightarrow x = y$$

A aplicação  $\Phi$  é também um homomorfismo:

$$\Phi(x) + \Phi(y) = \overline{(x, 1)} + \overline{(y, 1)} = \overline{(x \cdot 1 + y \cdot 1, 1 \cdot 1)} = \overline{(x + y, 1)} = \Phi(x + y)$$

$$\Phi(x) \cdot \Phi(y) = \overline{(x, 1)} \cdot \overline{(y, 1)} = \overline{(x \cdot y, 1 \cdot 1)} = \overline{(x \cdot y, 1)} = \Phi(x \cdot y)$$

$$x < y \Leftrightarrow x \cdot 1 < y \cdot 1 \Leftrightarrow \overline{(x, 1)} < \overline{(y, 1)} \Leftrightarrow \Phi(x)$$

Desta forma, podemos abusar da notação e escrever  $Z = \{\overline{(x, 1)} \in Q : x \in Z\}$  ainda que  $x$  e  $\overline{(x, 1)}$  sejam objetos distintos.

Para que essa identificação fique mais clara é conveniente usar a seguinte notação:

$$1. \quad r \cdot s^{-1} = \frac{r}{s} = r/s$$

Com isso, se  $r = \overline{(x, y)}$ , teremos  $r = \frac{\overline{(x, 1)}}{\overline{(y, 1)}}$ .

Por outro lado como os racionais  $\overline{(x, 1)}$  e  $\overline{(y, 1)}$  podem ser identificados com os números inteiros  $x$  e  $y$ , respectivamente, conclui-se que

$$2. \quad r = \frac{x}{y} = x/y$$

Ou seja, o número racional  $r$  está sendo escrito como uma fração de dois números inteiros.

#### 4. Números Reais

O conjunto  $Q$  dos números racionais contém algumas imperfeições, como por exemplo, não são suficientes para medir todos os comprimentos em geometria, ou mesmo para que a equação  $x^2 = 2$  tenha solução em  $Q$ . “As teorias não métricas do irracional passaram a exigir um maior rigor. O uso da palavra “quantidade” como uma magnitude geométrica sem referência a número e como um número que mede alguma grandeza era desconcertante, especialmente porque não existia uma base sólida para admitir que as mesmas regras de operação eram aplicáveis” (Cajori, 2007). Isso tinha uma implicação até mesmo sobre a própria definição de limite de uma função.

Surgiu então da necessidade da “criação” de um novo conjunto. O objetivo era completar o conjunto dos números racionais criando um sistema contínuo de números.

“A definição de número irracional (e conseqüentemente de número real) foi alcançada independentemente por Méray, Weierstrass, Dedekind e Cantor” (Cajori, 2007).

Vamos focar nesse trabalho o método de cortes no conjunto dos números racionais. Essa noção de corte foi introduzida por R. Dedekind.

Definimos um corte em  $Q$ , através de um par ordenado  $(A, B)$  de subconjuntos de números racionais. Mais precisamente, um par  $(A, B)$  é um corte no conjunto dos números racionais se as seguintes condições são verificadas:

- 1)  $A \neq \emptyset$  e  $B \neq \emptyset$
- 2)  $A \cup B = Q$
- 3)  $a \in A$  e  $b \in B \Rightarrow a < b$
- 4)  $A$  não tem elemento máximo.

O conjunto  $A$  e seus elementos são chamados de conjunto minorante e elementos minorantes, respectivamente. O conjunto  $B$  e seus elementos são chamados de conjunto majorante e elementos majorantes, respectivamente.

A título de exemplificação seja  $r$  um número racional e considere os conjuntos:

$$A_r = \{x \in \mathcal{Q}; x < r\}$$

$$B_r = \{x \in \mathcal{Q}; x \geq r\}$$

Então, o par  $(A_r, B_r)$  é um corte, que será denotado por  $r^\#$ .

Uma observação importante a se fazer é que um corte  $(A, B)$  no qual o conjunto majorante tem elemento mínimo denomina-se um corte racional. Assim, o corte do exemplo acima define um corte racional. Por outro lado um corte  $(A, B)$  no qual o conjunto majorante não tem elemento mínimo (em  $\mathcal{Q}$ ) denomina-se um corte irracional, como por exemplo, sejam:

$$A = \{x \in \mathcal{Q}; x^2 < 2 \text{ ou } x \leq 0\}$$

$$B = \{x \in \mathcal{Q}; x^2 > 2 \text{ e } x > 0\}$$

O par  $(A, B)$  é um corte, contudo, o conjunto  $B$  não tem elemento mínimo (em  $\mathcal{Q}$ ).

É possível agora definir o conceito de número real. O conjunto  $\mathfrak{R}$  formado por todos os cortes, racionais ou irracionais, é chamado de conjunto dos números reais e seus elementos de números reais.

Denotaremos um corte  $(A, B)$  ou um número real  $(A, B)$  por letras gregas minúsculas, ou seja,  $\alpha = (A, B)$ ,  $\beta = (C, D)$  números reais. Diremos que  $\alpha > \beta \Leftrightarrow \{\exists x \in \mathcal{Q}; x \in A \cap D\}$ , por outro lado diremos que  $\alpha < \beta \Leftrightarrow \{\exists x \in \mathcal{Q}; x \in B \cap C\}$ .

Disso seguem algumas propriedades facilmente verificadas como, por exemplo, a Lei da Tricotomia. Com base na Lei da Tricotomia temos que, se  $\alpha$  e  $\beta$  são números reais diremos que  $\alpha$  é maior ou igual a  $\beta$  e escrevemos  $\alpha \geq \beta$ , se  $\alpha > \beta$  ou  $\alpha = \beta$ , por outro lado diremos que  $\alpha$  é menor ou igual a  $\beta$  e escrevemos  $\alpha \leq \beta$  se  $\alpha < \beta$  ou  $\alpha = \beta$ . É importante salientar que  $\mathfrak{R}$  com a definição acima de menor ou igual é totalmente ordenado.

Por abuso de notação diremos que um número  $\alpha$  é positivo se  $\alpha > 0^\#$  e negativo se  $\alpha < 0^\#$ . O elemento  $0^\#$  é o corte racional exemplificado anteriormente, com  $r = 0$ .

A soma de dois conjuntos racionais  $A$  e  $B$  é dada por:  $A + B = \{a + b : a \in A, b \in B\}$ . Dados dois números reais  $\alpha = (A, B)$  e  $\beta = (C, D)$  o par

$\gamma = (A + C, (A + C)^c)$  é um corte e, portanto, um número real, que é chamado de soma dos números reais  $\alpha$  e  $\beta$  e denotado por  $\gamma = \alpha + \beta$ .

A adição em  $\mathfrak{R}$ , assim definida, é comutativa, associativa, possui um único elemento  $0^\#$  tal que para todo  $\alpha \in \mathfrak{R}$  tem-se  $\alpha + 0^\# = \alpha$  e existe um único elemento  $\beta$  denotado por  $-\alpha$  tal que  $\alpha + (-\alpha) = 0^\#$ .

De modo análogo, se  $\alpha = (A, A^c)$  e  $\beta = (B, B^c)$  são números reais positivos, então o par  $\gamma = ((A_+, B_+) \cup Q_-, ((A_+, B_+) \cup Q_-)^c)$  é um corte e define o número real  $\gamma$  chamado de produto de  $\alpha$  por  $\beta$ . O produto de dois números reais quaisquer, que leva em conta o conceito de módulo, sofre pequenas alterações e satisfaz as regras de sinais:  $(-\alpha)\beta = \alpha(-\beta) = -(\alpha\beta)$  e  $(-\alpha)(-\beta) = \alpha\beta$ , além da comutatividade, associatividade, distributividade em relação à adição e as leis de integridade e compatibilidade.

Por fim, forneceremos as idéias da identificação dos racionais como subconjunto dos reais. Primeiramente, definimos  $\mathfrak{R}_Q = \{r^* : r \in Q\}$  onde  $r^* = (A_r, B_r)$  é um corte racional (número real racional). Desta forma,  $\mathfrak{R}_Q$  é o conjunto dos números reais racionais. Podem ser verificadas (Lobeiro, 2008) as propriedades:  $r^* > (<)0^\# \Leftrightarrow r > (<)0$ ,  $-(r^*) = (-r)^*$ ,  $r^* + (.)s^* = (r + (.)s)^*$ , ou seja: a ordem, a adição e a multiplicação definida em  $\mathfrak{R}_Q$  são compatíveis com as respectivas ordem, adição e multiplicação definidas em  $Q$ . Portanto, mediante a aplicação  $\phi: Q \rightarrow \mathfrak{R}_Q$  dada por  $\phi(r) = r^*$ , podemos identificar o racional  $r$  com o real  $r^*$  e ver uma cópia de  $Q$  em  $\mathfrak{R}$ . Com isso, mostramos a construção de cada um dos conjuntos numéricos: inteiros, racionais e reais a partir dos números naturais e que, embora a natureza de cada um deles seja distinta, podemos dizer que  $\mathfrak{N} \subset Z \subset Q \subset \mathfrak{R}$ . É importante observar também que normalmente ao se iniciar um curso de análise, como em (Lima, 2007) acaba-se admitindo como axioma para a construção de  $\mathfrak{R}$  o teorema do supremo, que nada mais é que a idéia de corte introduzida neste trabalho. A contrapartida geométrica de um número real, em comparação com a figura apresentada de número racional (figura 4), seria de uma “reta” que passa pela origem.

**Referências Bibliográficas:**

LIMA, E. L. **Análise real. Funções de uma variável.** Rio de Janeiro: IMPA, vol. 1, 9ª edição, 2007.

LOBEIRO, A. M. **Construção dos reais: um enfoque usando cortes de Dedekind.** Universidade Estadual de Maringá, [http://www.dma.uem.br/kit\\_antigo/](http://www.dma.uem.br/kit_antigo/), acessado em 10/07/2008.

SANTOS, JOSÉ P. O. **Introdução à Teoria dos Números.** Rio de Janeiro: IMPA, 2000.

CAJORI, FLORIAN. **Uma história da Matemática.** Rio de Janeiro: LTC, 2007.

PINEDO, CHRISTIAN J. Q. **Fundamentos de Matemática II.** Pato Branco: UTFPR, 2002.

# REPETÊNCIA E EVASÃO NO CÁLCULO: PROPONDO ALTERNATIVAS

Profa. Ms. Renata Camacho Bezerra, Marli Schmitt.

Universidade Estadual do Oeste do Paraná - UNIOESTE,

renatacamachobezerra@gmail.com

Universidade Estadual do Oeste do Paraná - UNIOESTE, marlischmitt@hotmail.com;

**Palavras-chave:** Cálculo, Matemática, Ensino/Aprendizagem.

Este trabalho iniciou-se diante de nossas inquietações e indagações a respeito do processo de ensino e aprendizagem da disciplina de Cálculo Diferencial e Integral I, que há muito tem sido palco de diversas pesquisas nas Instituições de Ensino Superior, tanto em âmbito nacional quanto internacional, segundo Nasser (2007), e isso vêm de encontro com as percepções e as diversas opiniões acerca da disciplina, acompanhado de altos índices de reprovação, em diferentes cursos da área de Exatas.

A disciplina de Cálculo Diferencial e Integral geralmente é vista por alunos de diferentes cursos como uma das disciplinas da Matemática mais difícil, é a mais temida, e os números mostram que é também a que mais reprova. É senso entre os alunos que passar em Cálculo é para poucos, compreender Cálculo é para ainda menos. De acordo com Franchi (1995) *“para eles esta disciplina é de difícil compreensão e responsável por altos índices de reprovação”*.

Escuta-se com frequência diversas opiniões e desabafos de alunos que acabaram de ingressar na universidade, em especial nos cursos de Ciências Exatas, que não imaginavam que encontrariam uma disciplina tão difícil na grade curricular, como é o Cálculo Diferencial e Integral.

Brolezzi (2007, p. 21) faz referência a esta concepção acrescentando que:

A Matemática apresentada na Escola Básica, frequentemente como um conjunto de regras e fórmulas, processos mecânicos de resolução de determinados tipos de problemas, questões fechadas, com pouquíssima, às vezes nenhuma investigação, acarreta uma postura passiva por parte dos estudantes. Na Universidade, porém, a Matemática adquire um caráter distinto. É cobrada dos alunos uma experiência anterior que eles em geral não têm. Os professores

chegam a conclusão que aquilo que os alunos sabem de pouco vale para o aprendizado da Matemática em nível superior.

Conforme Barufi (1999), *“o Cálculo é uma ferramenta poderosa e importantíssima para as mais diversas áreas do conhecimento”*, que pode proporcionar ao aluno, além de ferramentas e metodologias para a resolução de problemas, tais como explicar o formato de latas ou a localização e a formação de um arco-íris, desenvolve o raciocínio, a crítica, a habilidade e a agilidade para enfrentar os desafios da atual e crescente globalização mundial. Esta mesma ferramenta que se faz indispensável para a criação, inovação e o desenvolvimento tecnológico, tem-se mostrado frequentemente como promotora de frustrações e de alarmantes índices de retenção de alunos nas universidades.

Nos cursos do Centro de Engenharias e Ciências Exatas – CECE da Universidade Estadual do Oeste do Paraná - UNIOESTE Campus de Foz do Iguaçu, atualmente localizado no Parque Tecnológico ITAIPU – PTI, situado na Usina Binacional de ITAIPU, no qual encontram-se os cursos de Ciência da Computação, Engenharia Elétrica, Engenharia Mecânica e Matemática, percebe-se um problema em relação à repetência, os quais já apontam por si só a necessidade de se buscar e traçar estratégias metodológicas para suprir tais problemas, que não diz respeito apenas à comunidade acadêmica, mas aos professores e a comunidade em geral, uma vez que afastam grande parte desta geração que conquistou o ingresso na universidade e que se vê a mercê de um fracasso profissional, ao invés de desempenhar sua função de desenvolvimento do conhecimento científico e tecnológico para a formação de uma sociedade mais instruída, crítica e consciente, responsável de seus atos e de sua participação para a evolução do meio em que se está inserido.

Isto tem se mostrado em âmbito nacional e internacional, que há tempo vem promovendo questionamentos por parte dos pesquisadores, em geral da área de Educação Matemática. Desta maneira, professores e pesquisadores desta linha de pesquisa têm procurado conhecer quais as causas para tamanha insatisfação e insucesso destes alunos, no que tange a disciplina de Cálculo Diferencial e Integral a fim de tentar auxiliá-los a reverter este quadro.

A partir de uma abordagem de pesquisa qualitativa pretende-se investigar o processo de ensino e aprendizagem de Cálculo Diferencial e Integral I nos cursos do CECE / UNIOESTE – FOZ.

Para D’Ambrósio (2004, p. 21 Apud Borba):



A pesquisa qualitativa é o caminho para escapar da mesmice. Lida e dá atenção às pessoas e às suas idéias, procura fazer sentido de discurso e narrativas que estariam silenciosas. E a análise dos resultados permitirá propor os próximos passos.

A pesquisa aqui relatada iniciou-se em 2006, com o objetivo principal de tentar identificar e compreender os principais motivos pelos quais um grande número de alunos não consegue aprender Cálculo, visto os altos índices de reprovações, bem como as principais falhas no processo de ensino e aprendizagem.

Abordou-se à priori as ementas destes cursos, elencando as semelhanças e diferenças existentes além do perfil do profissional que cada curso pretende formar.

Em seguida realizou-se um levantamento dos índices de aprovações e reprovações nos quatro cursos do centro para a disciplina de Cálculo I, no período de 2001 a 2005.

Ainda neste mesmo ano realizaram-se questionários com professores e alunos, sendo que os professores entrevistados foram os que lecionaram Cálculo I no período de 2003 a 2006 e alunos que cursaram por mais de uma vez esta disciplina neste mesmo período, sendo escolhidos aleatoriamente cinco (5) alunos de cada série de todos os cursos, totalizando oitenta (80) alunos.

Após, em 2007 fez-se uma entrevista onde se contou com o apoio e a participação da acadêmica Cátia Piano que está desenvolvendo o projeto de iniciação científica intitulado “O Cálculo no Curso de Licenciatura em Matemática”. Com a entrevista, recolheram-se dados com os professores que lecionaram a disciplina a fim de detectar quais os possíveis obstáculos que prejudicam o desempenho dos alunos no Ensino Médio, onde muitas vezes é apontado como fator predominante na baixa aprendizagem de seus alunos em Cálculo I, segundo Brolezzi (2007) e Nasser (2007), bem como identificar e sugerir estratégias voltadas a reverter esta problemática de forma a tentar contribuir com o processo de ensino e aprendizagem de Cálculo Diferencial e Integral I.

Durante as entrevistas, os professores mostraram interesse na possibilidade da realização de um curso Pré-Cálculo e demonstraram grande apoio à realização do mesmo.

De posse desses dados partiu-se para a elaboração de estratégias que pudessem contribuir para a aprendizagem dos alunos, a partir de um curso Pré-Cálculo.

O público alvo do curso foi os acadêmicos matriculados na disciplina de Cálculo Diferencial e Integral I dos quatro cursos do CECE, podendo ser alunos com histórico de repetência na disciplina ou não.

Foi confeccionada uma apostila, com conteúdos sugeridos pelos professores entrevistados. Os exercícios presentes na apostila foram extraídos de livros da Coleção de Matemática Elementar e do próprio livro de Cálculo I, utilizado pelos professores e presente nas ementas desta disciplina para os respectivos cursos.

Após a confecção da apostila deu-se início a organização do curso. A idéia era realizar este curso com apenas uma turma de mais ou menos 20 alunos. Porém teve-se mais de 120 inscritos, e devido a esta grande procura, fez-se necessário atender 4 turmas com capacidade de 30 alunos cada. Acredita-se que este número de turmas também poderá contribuir para o resultado final do projeto, pois permitirá uma análise comparativa entre as turmas e o desempenho dos alunos.

Objetivou-se com este curso Pré-Cálculo auxiliar a integração dos alunos de diferentes cursos, a formação de grupos de estudo, permitindo assim a troca de experiências e mesmo de ajuda com o conteúdo de Cálculo, bem como colaborar para a aprendizagem de conteúdos matemáticos, apontados pelos professores como pré-requisitos para um bom desempenho em Cálculo I.

Pretende-se com esta pesquisa também, investigar quais as possíveis contribuições do Pré-Cálculo para estes alunos. Para tal far-se-á um acompanhamento dos alunos que participaram do curso e que obtiveram frequência superior a cinquenta por cento (50%) das aulas ministradas, de forma a abranger o maior número possível de participantes.

Este acompanhamento dar-se-á através de duas entrevistas, uma ao término do curso e outra no mês de setembro e ainda será utilizado como meio de acompanhá-los o parâmetro nota, podendo ser estas parciais e semestrais, uma vez que este processo avaliativo poderá não ter sido concluído em prazo adequado a construção desta monografia.

Por se tratar de uma pesquisa com abordagem qualitativa e devido a multiplicidade de procedimentos adotados no decorrer de seu desenvolvimento, pretende-se ter uma compreensão mais abrangente dos resultados da pesquisa.

Bicudo (2004, p. 104) afirma que:

O qualitativo engloba a idéia do subjetivo, passível de expor sensações e opiniões. O significado atribuído a essa concepção de pesquisa também engloba noções a respeito de percepções de diferenças e semelhanças de aspectos comparáveis de experiências.

Assim, espera-se que esta pesquisa em andamento possa contribuir com o processo de ensino-aprendizagem do Cálculo Diferencial e Integral I nos cursos do CECE e que a mesma possa fornecer subsídios para investigações futuras.

### **Referências Bibliográficas:**

BARUFI, M. C. B. A Construção/Negociação de Significados no Curso Universitário Inicial de Cálculo Diferencial e Integral. **Tese de Doutorado**, Universidade de São Paulo, 1999.

BICUDO, M. A. V. Pesquisa Qualitativa e Pesquisa Quantitativa Segundo A Abordagem Fenomenológica. In: BORBA, M. C. **Pesquisa Qualitativa em Educação Matemática**. Belo Horizonte: Autêntica, 2004.

BORBA, M. C. **Pesquisa Qualitativa em Educação Matemática**. Belo Horizonte: Autêntica, 2004.

BROLEZZI, A. C.; BARUFI, M. C. B. **História da Matemática e Ensino de Cálculo: Reflexões Sobre o Pensamento Reverso**. Guarapuava: SBHMat, 2007.

FRANCHI, R. H. O. L. Cursos de Cálculo: Uma Proposta Alternativa. **Temas e Debates: O Ensino de Cálculo**. Sociedade Brasileira de Educação Matemática (SBEM) Ano VIII, n. 6, (1995), 39-43.

NASSER, L. Ajudando a Superar Obstáculos na Aprendizagem de Cálculo. In: Encontro Nacional de Educação Matemática, 9., 2007, Belo Horizonte. **Anais IX ENEM**, Belo Horizonte: UniBH, 2007. CD-ROM.

# RESOLUÇÃO DO PROBLEMA DE STEINER

Guilherme Baggio Ferla, Carlos Alexandre Ribeiro Martins

Universidade Tecnológica Federal do Paraná - UTFPR, guilhermebaggioferla@hotmail.com;

Universidade Tecnológica Federal do Paraná - UTFPR, carlos@utfpr.edu.br

**Palavras-chave:** Geometria Euclidiana, Ponto de Steiner, Arco-Capaz.

## Resumo:

O problema, inicialmente proposto por Fermat, era encontrar um ponto no plano cuja soma das distâncias a três pontos distintos e não-colineares fosse a mínima possível. Constata-se que este problema é relevante para minimizar custos, como na instalação de redes de telefonia, de água, de esgoto e rodovias. A solução encontrada para o caso de 3 pontos parte de alguns conhecimentos de Geometria Euclidiana Plana.

## 1 Introdução.

A primeira vez que o nome Problema de Steiner foi utilizado foi em 1941, no livro *What is mathematics?* de R. Courant e H. Robbins. Entretanto, muito antes de Steiner (1796-1863), outros matemáticos como Fermat (1601-1665), Torricelli (1608-1647), Cavalieri (1598-1647) e Simpson (1710-1761) já haviam trabalhado neste problema.

Torricelli encontrou o ponto-solução para este problema nos casos em que os ângulos internos do triângulo formado pelos três pontos distintos e não-colineares são todos menores ou iguais a  $120^\circ$ . Cavalieri percebeu uma propriedade importante do ponto de Torricelli: quando o ponto está dentro do triângulo, os ângulos entre os segmentos que ligam o ponto de Torricelli aos pontos do triângulo são todos iguais a  $120^\circ$ .

Simpson, em seu livro Fluxions, propôs o problema de Fermat em sua forma geral: encontrar um ponto no plano cuja soma das distâncias a  $n$  pontos dados  $A_1, \dots, A_n$  fosse mínima.

Neste presente trabalho o problema de Steiner se restringe ao caso de três pontos, onde nenhum ângulo interno excede  $120^\circ$ . São apresentadas, de forma resumida, algumas questões levantadas ao longo do seu desenvolvimento. Resultando, por fim, no encontro de soluções matemáticas para o problema com base em postulados colocados por matemáticos ao longo da história.

## 2 Materiais e Métodos.

O estudo da Geometria Euclidiana Plana, feito de forma axiomático, possibilitou a criação de meios que tornaram possível estudar o Problema de Steiner.

O caso  $n=3$  foi inicialmente dividido em dois teoremas: dado três pontos não colineares e distintos, o ponto-solução está no interior do triângulo e não no seu exterior; e no caso de três pontos colineares o ponto-solução é o ponto que se encontra entre os outros dois. Todo o desenvolvimento do trabalho foi verificado experimentalmente no software livre de Geometria Dinâmica GeoGebra versão 3.0 ([www.geogebra.at](http://www.geogebra.at)).

## 3 Resultados e Discussões.

### 3.1 Teoremas I e II.

O Teorema I considera três pontos distintos e não colineares. Serve para mostrar que dado um ponto qualquer exterior ao triângulo formado pelos pontos é sempre possível encontrar um ponto no triângulo cuja soma das distâncias deste ponto aos vértices do triângulo é menor do que a soma das distâncias do ponto exterior a estes vértices.

**Teorema 1** *Dado um ponto  $Q$  exterior ao triângulo, conforme a Figura 01, sempre é*

possível encontrar um ponto  $P$  no triângulo  $ABC$  tal que:  $\overline{QA} + \overline{QB} + \overline{QC} > \overline{PA} + \overline{PB} + \overline{PC}$ .

**Prova.** Dado que  $P \neq Q$  e  $Q$  exterior ao triângulo  $ABC$ ,  $QA + QC > AC$ . Como  $AC = PA + PC$ , logo  $QA + QC > PA + PC$ . Tendo  $QB > PB$ , então de  $QA + QC > PA + PC$  e  $QB > PB$  segue  $QA + QB + QC > PA + PB + PC$ , como se queria provar.

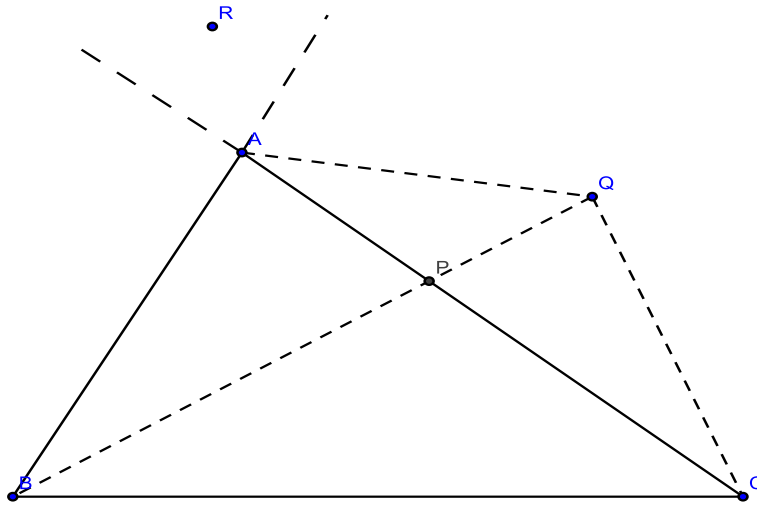


Figura 01

**Observação:** No teorema acima, existe o caso em que o ponto fora do triângulo é o ponto  $R$ , como Figura 01, tal demonstração é próxima à anterior.

**Teorema 2** Dados três pontos colineares,  $A$ ,  $B$  e  $C$  nesta ordem, e  $Q$  fora da reta que contém os três pontos. Então  $QA + QB + QC > BA + BB + BC$ .

**Prova.**  $QA + QC > AC$ , ou seja,  $QA + QC > BA + BC$ . Sendo  $Q \neq B$  e  $BB = 0$ , então  $QB > BB$ , donde  $QA + QC + QB > BA + BC + BB$ , como se queria provar. Logo, o ponto que minimiza a soma das distâncias é o ponto  $B$ .

### 3.2 Teorema III

Tomando como base o Teorema I, partimos para o caso em que se consideram apenas triângulos com ângulos internos menores que  $120^\circ$ . A resolução do Teorema III mostra

que o ponto-solução é um ponto no interior do triângulo que "olha" para os três lados do triângulo sob o mesmo ângulo, como Cavalieri já havia mostrado.

**Teorema 3** *Dado um triângulo  $ABC$  conforme Figura 02, onde nenhum dos ângulos internos excede  $120^\circ$ , e um ponto  $P$  interior ao triângulo, quer encontrar-se o valor mínimo  $d$  para a soma  $AP + BP + CP$ .*

**Prova.** *Dado o triângulo  $ABP$ , considerando o centro em  $B$ , gira-se o triângulo em  $60^\circ$  no sentido anti-horário, formando o triângulo  $C'P'B$ . Por congruência de triângulos temos que  $AP + BP + CP = C'P' + P'P + PC$ , sendo assim o comprimento  $d$  é a igual ao comprimento da poligonal  $C'P'PC$ . Conclui-se que  $d$  será mínimo quanto menor for o comprimento da poligonal. Este valor será encontrado quando os pontos  $C', P', P$  e  $C$  estiverem alinhados, atingindo o menor comprimento possível. Neste caso, prova-se que  $\hat{B}PC = 180^\circ - \hat{P}'PB = 120^\circ$  e  $\hat{A}PB = \hat{C}'P'B = 180^\circ - \hat{P}'P'B = 120^\circ$ . Concluí-se, portanto que o ponto-solução é um ponto que "olha" para os três lados com o ângulo de  $120^\circ$ .*

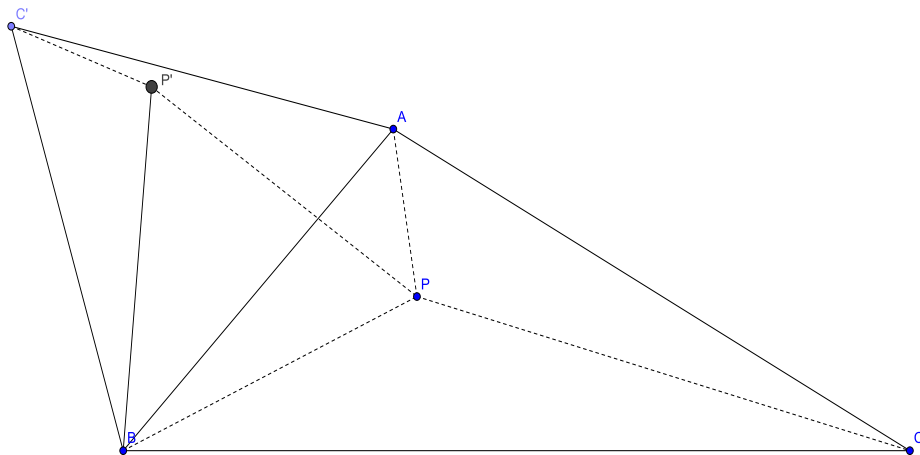


Figura 02

### 3.3 Materialização do Ponto.

O meio utilizado para materializar o ponto - solução para o Teorema III foi construindo três arcos capazes. Todo ângulo inscrito num arco-capaz  $AB$  qualquer, cujos lados

passam por  $A$  e  $B$ , são congruentes, o que significa que o segmento  $AB$  é sempre visto sob um mesmo ângulo. Dado um triângulo  $ABC$ , construindo três arcos-capazes que olham para os segmentos  $AB$ ,  $AC$  e  $BC$  sob um ângulo de  $120^\circ$ , o ponto - solução será dado pelo ponto na intersecção destes três arcos-capazes.

## 4 Conclusão.

O desenvolvimento da solução do Problema de Steiner nos remete a duas situações: I) três pontos distintos e colineares; e II) três pontos distintos e não colineares onde nenhum ângulo interno do triângulo excede  $120^\circ$ . Para a primeira situação o ponto - solução é aquele que se encontra entre os outros dois e para a segunda situação o ponto - solução é aquele que "olha" para os três lados do triângulo sob o ângulo de  $120^\circ$ , podendo ser materializado pela construção de três arcos-capazes.

### Referências Bibliográficas:

- BALDIN, Y.Y., VILLAGRA, G.A.L.: **Atividades com Cabri-Géomètre II**. Editora da Universidade Federal de São Carlos: São Carlos, 2002.
- BARBOSA, J.L.M.: **Geometria Euclidiana Plana**. 6. ed. Sociedade Brasileira de Matemática: Rio de Janeiro, 2004.
- GERÔNIMO, J.R, FRANCO, V.S.: **Geometria Plana e Espacial: um estudo axiomático**. 1. Ed. Editora Massoni: Maringá, 2005.
- REZENDE, E.Q.F., QUEIROZ, M.L.B.: **Geometria Euclidiana Plana e construções geométricas**. Editora da Unicamp: Campinas, 2000.
- SOCIEDADE BRASILEIRA DE MATEMÁTICA.: **Revista do Professor de Matemática**. Nº 16, 1º semestre de 1990.



# **MATEMÁTICA: RESULTADO DE UMA CONSTRUÇÃO SOCIAL**

Profª. Ms. Renata Camacho Bezerra, Danieli Aquino Somensari

Universidade Estadual do Oeste do Paraná – UNIOESTE, renatacamachobezerra@gmail.com;

Universidade Estadual do Oeste do Paraná – UNIOESTE, danielisomensari@hotmail.com.

**Palavras-chaves:** Matemática, Construção Social e Aprendizagem.

## **Resumo:**

Ao longo de seu desenvolvimento percebemos que a Matemática é resultado de uma construção social, que nasceu da necessidade humana de se localizar no tempo e no espaço, através da contribuição dos homens. No entanto muitas pessoas hoje, não sabem como relaciona-la com o seu dia-a-dia.

Através de sua história pudemos ver que várias concepções foram criadas e (re) criadas a seu respeito, o dicionário a define como estudo de padrões de quantidade, estrutura, mudanças e espaço. As pessoas num geral também possuem uma definição a seu respeito e hoje a mais popular delas é que a Matemática é uma disciplina difícil e em alguns casos a grande responsável pela evasão e repetência escolar.

Esta visão já faz parte da mentalidade de alunos das séries iniciais do Ensino Fundamental e em alguns casos antes mesmo de entrarem no ambiente escolar, concepção esta transmitida pelos pais e pela sociedade de uma forma geral.

A aprendizagem da Matemática na sala de aula na maioria das vezes ocorre de forma mecânica e tradicional e isso pode contribuir com esta aversão, no entanto não queremos dizer que o quadro e o giz não sejam importantes, mas é necessário diferenciar o método de ensino para que desperte nos alunos o interesse em aprender.

Percebemos também que a Matemática é considerada por muitos como uma ciência perfeita, incapaz de errar, isso é transmitido pela mídia e também pela escola, porém sabemos que isto não é verdade e como já citado anteriormente ela faz parte de uma construção social, e muito ainda se pode criar, e até mesmo recriar.

Foi pensando em tudo isto que criamos o projeto denominado como Estação Ciência – Módulo de Matemática que esta sendo desenvolvido desde o ano de 2006, nas dependências do Parque Tecnológico Itaipu – PTI, com o objetivo de desmistificar e

socializar a Matemática e este trabalho é desenvolvido com crianças do 1º e 2º ciclo do Ensino Fundamental através de atividades lúdicas, ou seja, atividades relacionadas com jogos e divertimento para facilitar o ensino e a aprendizagem da Matemática.

Segundo Chateu (1908) o jogo para a criança é em primeiro lugar brincadeira, e também uma atividade séria em que o faz-de-conta e a imaginação possuem uma grande importância, ou seja, o lúdico, o divertimento faz parte de seu estágio de vida e seria essa uma alternativa de facilitar a aprendizagem.

Algumas alternativas, como diferentes metodologias de ensino e materiais didáticos que facilitem o processo de ensino e aprendizagem dos alunos tem sido o interesse dos professores, percebe - se que estes procuram participar de palestras e cursos que apresentem maneiras diferentes de se ensinar Matemática.

Há muitas crianças que não gostam de Matemática porque o método tradicional de ensino faz memorizar coisas que elas não entendem ou pelas quais não se interessam, como as etapas dos algoritmos a da tabuada. além do mais, as crianças têm medo de Matemática porque a professora categoricamente diz: “isto está certo” ou “isto esta errado.” (Kamii – Revista Aprendizagem, 2008)

Mas será que o trabalho desenvolvido neste espaço esta gerando uma mudança no conceito que os alunos têm do que é Matemática? Este é o objetivo da pesquisa, ou seja, pretendemos investigar qual a visão que os alunos possuem da Matemática ensinada na sala de aula, além de propor alterações e sugestões para o módulo de Matemática.

No momento esta sendo aplicado um questionário aos alunos que vão até a Estação Ciência – Módulo de Matemática para saber qual é sua concepção sobre a Matemática antes de ser desenvolvida a atividade e depois que participam da atividade.

Antes de participarem das atividades os alunos respondem algumas questões rápidas sendo elas:

- 1) Você gosta de Matemática? ( )sim ( )não
- 2) Você acha a Matemática difícil? ( )sim ( )não

Após participarem das atividades desenvolvidas no Módulo os alunos respondem algumas questões a respeito da atividade:

- 1) O que você achou da atividade desenvolvida no Módulo de Matemática? ( ) ruim  
( ) bom ( ) ótima
- 2) Gostaria que a professora realizasse alguma atividade do Módulo de Matemática na escola? ( ) sim ( ) não
- 3) Gostaria de voltar a Estação Ciência – Módulo de Matemática e participar de alguma atividade diferente da que está participando hoje? ( ) sim ( ) não

Através de uma conversa informal que temos com os alunos antes de cada atividade parece-nos que existe uma forma de rejeição por parte destes em relação à disciplina, alguns até dizem que não é possível aprender Matemática sem que possua números e faça continhas, outro caso citado é que é impossível aprender Matemática brincando.

Com o desenvolver do trabalho esta visão que as crianças possuíam da Matemática vem se alterando e elas percebem que é possível aprender Matemática de forma diferente do tradicional, ou seja, é possível aprender alguns conceitos matemáticos de forma e dinâmica, lúdica e através do divertimento, mas para que tenhamos dados concretos, faz-se necessária a aplicação de um questionário aos alunos que visitam o módulo de Matemática.

Além da investigação com os alunos, pretendemos também fazer uma investigação com os professores do Ensino Fundamental das séries iniciais com o objetivo de conhecer qual a necessidade que eles possuem para que se possa facilitar o processo de ensino e aprendizagem das crianças, e no ponto de vista deles existe um receio por partes dos alunos pela disciplina e uma dificuldade de relacionarem a Matemática aprendida na sala de aula com situações do seu dia-a-dia.

Estamos desta forma promovendo a indissociabilidade entre o ensino, a pesquisa e a extensão. Neste caso, na seguinte ordem, extensão (projeto Estação Ciência), ensino (as atividades desenvolvidas no módulo de Matemática) e a pesquisa (através da investigação da concepção de alunos e professores antes e após as atividades).

O brincar e se divertir não representa somente a infância, desenvolver a coordenação motora da criança e seu convívio social muito, além disso, pode-se despertar nela o interesse em aprender algo. “Para “educar” as crianças, nada melhor do que as brincadeiras, pois a ação lúdica é um princípio próprio da atividade infantil e

indica o modo como a criança pensa, ou seja, a capacidade de reflexão se desenvolve pelo brincar.”(MELO – Revista Aprendizagem, 2008)..

Sabe – se que a Matemática esta presente em diversas situações do nosso dia-a-dia, por isso a grande preocupação e importância para que as pessoas num todo saibam como lidar com ela quando necessário, e isso deve ser trabalhado desde a infância. Mostrando que todos podem gostar e aprender, quebrar o paradigma de que só os considerados “inteligentes” aprendem. Este ensino deve iniciar desde as séries iniciais do Ensino Fundamental, pois é nessa fase que se constrói a base para um conhecimento maior, não somente da Matemática, mas de todas as disciplinas.

Um dos nossos objetivos enquanto professor e futuros professores de Matemática é socializar a mesma e tornar o conhecimento científico acessível a todos, através de atividades interativas e experimentos lúdicos de forma a despertar a curiosidade e o interesse pela disciplina, por isso a importância do trabalho ser avaliado, de forma que possamos analisar se estamos alcançando nossos objetivos, ou seja, se as concepções dos alunos estão mudando em relação a esta disciplina que vem sendo discutida até o momento e se o processo de ensino e aprendizagem tem sido facilitado.

Através desta pesquisa realizada a respeito dessa maneira diferente de ensinar Matemática, com atividades lúdicas e relacionadas com o dia-a-dia poderemos confirmar o que acreditamos ser uma forma de ensinar e transmitir o conhecimento de maneira que facilite o processo de ensino e aprendizagem da Matemática, tornando-a mais humana, livre dos paradigmas de ciência fria e infalível.

### **Referências Bibliográficas**

CHATEAU, J. **O jogo e Criança** (tradução: Guido de Almeida) – São Paulo: Summus, 1987.

BARALDI, I. M.; **Matemática na Escola: Que Ciência é esta?** Bauru, SP: EDUSC, 1999.

BORBA, M. C.; SKOVSMOSE, O.; **A Ideologia da Certeza em Educação Matemática. In. Educação Matemática Crítica: a questão da democracia.** Campinas Papyrus, 2001 pp.127-148. (Coleção Perspectivas em Educação Matemática).

KAMII, C.; Jogos Autonomia e Matemática. Aprendizagem: a revista da prática pedagógica 2008, nº4, 30.

MELLO, P.; Educação infantil valoriza o lúdico, reconhecendo sua importância para o desenvolvimento de aspectos cognitivos, sociais, emocionais, afetivos e psicomotores das crianças. *Aprendizagem: a revista da prática pedagógica* 2008, nº4, 14.

Artigo: **Jogar e brincar** – Maria da Conceição aparecida Leira Ramos e Prof<sup>a</sup>. Orien. Ms. Íris Weiduschat, disponível no site <http://www.icpg.com.br/artigos/rev01-07.pdf> acesso em 07/10/2007.

[http://www.educacional.com.br/articulas/outrosEducacao\\_artigo.asp?artigo=artigo0071](http://www.educacional.com.br/articulas/outrosEducacao_artigo.asp?artigo=artigo0071) acesso em 30/07/07.

## **(RE) CONSTRUINDO O TANGRAM**

Prof<sup>a</sup>. Ms. Renata Camacho Bezerra, Danieli Aquino Somensari, Luciana Misturini,  
Marizete da Silva, Marli Schmitt.

Universidade Estadual do Oeste do Paraná – UNIOESTE,  
renatacamachobezerra@gmail.com; UNIOESTE/Foz do Iguaçu,  
danielisomensari@hotmail.com; UNIOESTE/Foz do Iguaçu,  
luciana\_sti@hotmail.com; UNIOESTE/Foz do Iguaçu,  
mari\_zt\_sti@hotmail.com; UNIOESTE/Foz do Iguaçu,  
marlischmitt@hotmail.com; UNIOESTE/Foz do Iguaçu

**Palavras-chaves:** Matemática, Ensino e Tangram

### **Resumo:**

Desde o ano de 2006 vem sendo realizado o trabalho de extensão intitulado Estação Ciência, que tem como objetivo levar o conhecimento científico a toda à comunidade. O foco principal do projeto são as crianças do 1<sup>a</sup> e 2<sup>a</sup> ciclos do ensino fundamental, porém esta aberta a todos os interessados.

A Estação Ciência está localizada na Fundação Parque Tecnológico de Itaipu – FPTI, e é um projeto desenvolvido em parceria entre o Programa de Desenvolvimento Tecnológico Avançado - PDTA e a Universidade Estadual do Oeste do Paraná – UNIOESTE.

Os acadêmicos da UNIOESTE sob orientação dos professores coordenadores do projeto são responsáveis em elaborar atividades para serem desenvolvidas nos módulos de Biologia, Física, Informática e Matemática, sendo os cursos de Ciência da Computação, Enfermagem, Matemática e Engenharia Mecânica que contribuem para o desenvolvimento deste trabalho.

No Módulo de Matemática temos como objetivo principal tornar o conhecimento matemático acessível a todos, levando a matemática a crianças das séries iniciais do Ensino Fundamental, de forma lúdica e relaciona-la com o dia-a-dia.

“Atualmente, pelo menos no campo da matemática, se considera que para a matemática deixar de ser vista como algo somente para uns poucos gênios, torna-se necessário concebê-la como um processo de

construção ligado – tanto em elaboração históricas quanto no desenvolvimento das idéias matemática nas pessoas – à resolução de problemas concretos, muitos deles gerados em outros campos do conhecimento ou na atividade humana.” (PAVANELLO 2004).

Várias atividades foram criadas e desenvolvidas neste período no Módulo de Matemática, entre elas: A História dos Números, Jogos Boole, Volume, Simetria, Cidade Geométrica e Tangram, todas através de materiais concretos facilitando assim o processo de ensino e aprendizagem, sem que o aluno precise imaginar uma situação problema para desenvolver sua atividade, mas ele possui em mãos todo o material necessário.

No presente artigo vamos relatar a atividade do (Re) Construindo o Tangram desenvolvido no Módulo de Matemática.

Esta atividade tem como objetivo levar o conhecimento do Tangram aos alunos da educação básica. É um conceito presente nos livros didáticos, mas na maioria das vezes não chega ao conhecimentos dos alunos.

Por esse fato é que desenvolvemos esta atividade com o Tangram no Módulo de Matemática para explorar alguns conceitos matemáticos presente nele e em alguns casos apresentar a alunos que não conheciam ou até mesmo nunca ouvido falar.

O Tangram é um conceito ensinado na disciplina de matemática, porém somente conhecido por muitos como um jogo de distração para desenvolver o raciocínio lógico.

Este é um jogo muito antigo que segundo algumas lendas é de origem oriental. Composto por sete peças todas num mesmo padrão para que possam se encaixar, das quais são:

- Cinco triângulos de tamanhos diferentes;
- Um quadrado;
- Um paralelogramo,

Neste jogo primeiramente exploramos o conteúdo geométrico presente nas formas que a compõe, depois usamos a imaginação para criar outras formas utilizando todas as peças.

Não se sabe bem ao certo qual a real origem do Tangram, seu criador, sua data e local, porém alguns dizem ser de origem chinesa e seu nome significa “tábua das sete sabedorias” ou “tábua das sete sutilezas”.

De acordo com algumas das lendas sua origem se deu no momento em que um jovem chinês, ao se despedir de seu mestre para iniciar uma grande viagem recebeu dele um espelho de forma quadrada para o seguinte dizer:

- Com este espelho você irá registrar tudo o que ver para me mostrar quando retornar.

O discípulo curioso e sem entender perguntou:

- Mestre como poderei registrar tudo o que ver em um espelho?

No exato momento que perguntava o espelho escapando de sua mão caiu-se no chão e se quebrou em sete pedaços, assim o mestre respondeu:

- Agora com estas sete peças você poderá construir figuras para representar tudo o que ver em sua viagem.

Existem diversas versões sobre a lenda que conta sua origem, esta contada acima é uma das mais conhecidas.

O tangram proporciona ao professor diversas maneiras de trabalhar geometria, ângulo, perímetro, reta, segmento, área, explorar as formas planas, características que as compõem através da construção realizada pelos próprios alunos assim como apresentar aos seus alunos sólidos geométricos.

Uma maneira de se construir o tangram de forma criativa que fará com que os alunos aprendam e dificilmente se esqueçam, é através de dobraduras a forma com que realizamos na Estação Ciência – Módulo de Matemática.

Iniciávamos a atividade perguntando aos alunos se conheciam o tangram, alguns diziam que sim e outros se perguntavam o que poderia ser isso.

Apresentamos assim o quebra-cabeça e então contávamos à lenda que o originou citada acima. Muito conhecido como um jogo que possui como um de seus objetivos o desenvolvimento do raciocínio lógico na construção de figuras assim como o ensino da matemática encontrada nele.

Assim os alunos descobriam que o tangram no início de seu surgimento era apenas um espelho de forma quadrada que acabou caindo no chão e se quebrando em sete peças.

Após isso construíamos junto aos alunos um tangram através de dobraduras, cada aluno recebia um pedaço de papel cartolina de forma quadrada e com as orientações recebidas pelo monitor do módulo ele poderia construir as peças que o compõem através de dobras, por fim deveriam recorta-las e com estas representar diversas figuras.



Desta maneira exploramos o conteúdo geométrico e as características principais que cada peça compõe, entre elas o nome específico de cada figura.

Os alunos construíram cada um o seu tangram através de dobradura e usando a imaginação para representarem outros objetos.

Segundo (FIORENTINI e MIORIM, 2001) “a matemática deve ter uma forma de ensinar mais ligada a realidade, que não obrigue os alunos a apenas decorar ou aplicar regras.”, ou seja, os alunos devem se interessar em aprender matemática só dessa forma eles realmente irão aprender.

Assim procuramos contribuir com o processo de ensino e aprendizagem da matemática, de forma diferente da sala de aula, do tradicional, fazendo com que os alunos participem mais ativamente das atividades de forma que cada um possa construir o seu experimento.

### **Referências Bibliográficas:**

FIORENTINE, D. MIORIM, M.A. **Por trás da porta, o que acontece?**, Campinas, SP: Editora Graf. FE/UNICAMP – CEPEN, 2001.

PAVANELLO, R. M. **Matemática nas Séries Iniciais do Ensino Fundamental: A pesquisa em sala de aula**. São Paulo: Coleção SBEM, volume 2, 2004.

DREWS, S. B. T. KINALSKI, N. M. Artigo: **Tangram e a geometria no ensino fundamental e no ensino médio**. Disponível no site

<http://ccet.uces.br/eventos/outros/egem/minicursos/mc62.pdf> acesso no dia 05/11/2007

<http://paje.fe.usp.br/~labmat/edm321/1999/geometr/tangram.html>, acesso em 05/11/2007

<http://pedagogia.brasilecola.com/trabalho-docente/a-configuracao-geometrica-tangram.htm>, acesso em 05/11/2007

<http://pt.wikipedia.org/wiki/Tangram>, acesso no dia 20/06/2008

**PESQUISA E EXTENSÃO:**  
**CONSTRUINDO ALTERNATIVAS METODOLÓGICAS PARA**  
**AUXILIAR O PROCESSO DE ENSINO E APRENDIZAGEM DO**  
**CÁLCULO DIFERENCIAL E INTEGRAL NOS CURSOS DO CECE**

Cátia Piano, Renata Camacho Bezerra.

Universidade Estadual do Oeste do Paraná - UNIOESTE, catiapiano@gmail.com; Universidade  
Estadual do Oeste do Paraná - UNIOESTE, renatacamachobezerra@gmail.com

**Palavras-chave:** Cálculo, Matemática, Ensino/Aprendizagem.

**Resumo:**

Considerada uma das vilãs dos cursos superiores nas áreas de exatas, a disciplina de Cálculo Diferencial e Integral I, também conhecida apenas como Cálculo I, sempre foi a grande responsável por altos índices de reprovação acadêmica e por grande parte do abandono de cursos superiores, não sendo diferente nos cursos do Centro de Engenharias e Ciências Exatas – CECE, da Universidade Estadual do Oeste do Paraná – UNIOESTE, Campus de Foz do Iguaçu.

Apesar da “má fama”, é inegável a importância do Cálculo I para o desenvolvimento do raciocínio do aluno e, além disso, para Catapani (2001), a disciplina de Cálculo Diferencial e Integral não só serve de base para as diversas carreiras da área de exatas, bem como é uma ferramenta importante para a resolução de problemas.

Procurando conhecer os fatores que levam aos alarmantes índices de reprovação, realizou-se durante o ano de 2006 e início de 2007<sup>1</sup>, uma pesquisa com os professores e alunos da disciplina dos quatro cursos do CECE. Esta pesquisa deu-se através da análise das ementas de Cálculo I de Ciência da Computação, Engenharia Elétrica, Engenharia Mecânica e Matemática. Também foram analisados dados sobre as aprovações e reprovações na disciplina no período de 2001 a 2005 dos quatro cursos fornecidos pela Secretaria Acadêmica da UNIOESTE. E ainda, realizaram-se duas entrevistas, uma com os professores que ministraram a disciplina durante o ano de 2006 e outra com os

---

<sup>1</sup> Pesquisa realizada pela acadêmica Marli Schmitt sob orientação da Profa. Ms. Renata Camacho Bezerra.

acadêmicos que haviam cursado a disciplina. Nestas entrevistas foi apontado, entre outros fatores, a falta de maturidade dos acadêmicos em Cálculo I e a deficiência de conteúdo matemático básico.

Para esclarecer o que os professores queriam dizer com “falta de maturidade” e qual seria realmente este conteúdo matemático que os alunos apresentam deficiência, decidiu-se realizar, ao final do ano de 2007, uma nova entrevista somente com os professores que lecionaram a disciplina no decorrer do período letivo de 2007.

Os professores atribuem a falta de maturidade dos alunos a alguns fatores, entres eles, a baixa faixa etária dos mesmos ao ingressarem na universidade (cerca de 18 anos), a diferença entre as metodologias de ensino adotadas no período escolar e na universidade, o despreparo dos alunos frente ao conteúdo, e principalmente, a sensação equivocada que os acadêmicos tem de que a faculdade é uma continuação da escola (grande parte dos alunos chega a universidade com a idéia que para se formar basta “tirar” nota e passar em todas as disciplinas, o fato de ser a preparação para a profissão escolhida é muitas vezes desconsiderado).

Quanto ao conteúdo matemático, foram apontados aqueles conteúdos de matemática básica, ou seja, conteúdos que são trabalhados no decorrer da vida escolar, tais como Equações, Expressões Algébricas, Funções (domínio, contradomínio, imagem, interpretação geométrica, interpretação analítica, funções de 1º e 2º grau, exponenciais, logarítmicas, ...), Simplificações (de frações, de expressões, ...), Trigonometria, Geometria, Gráficos de funções básicas, e interpretação de situações problema. Perguntamos ainda aos professores o que é necessário para aprender Cálculo, como resposta, obtivemos três requisitos sendo eles: uma base matemática sólida (conteúdo), raciocínio lógico-matemático e disposição para o estudo.

Como alternativa para auxiliar no processo de ensino-aprendizagem do Cálculo Diferencial e Integral foi desenvolvido um projeto de extensão<sup>2</sup>, na forma de um curso Pré-Cálculo, que buscava sanar as deficiências de conteúdo básico e também integrar os alunos dos cursos de forma a possibilitar a formação de grupos de estudo. Para a realização do curso foi desenvolvido um material com teoria e exercícios englobando os conteúdos apontados pelos professores. Para a confecção deste material, foram utilizados como referência livros de matemática de ensino médio, livros de fundamentos da matemática e o próprio livro de Cálculo Diferencial e Integral.

---

<sup>2</sup> Projeto de Extensão realizado pelas acadêmicas Cátia Piano e Marli Schmitt sob coordenação da Profa. Ms. Renata Camacho Bezerra.

O público alvo do curso foi os acadêmicos matriculados na disciplina de Cálculo I dos quatro cursos. Devido ao grande número de inscrições, foram montadas 4 turmas com cerca de 30 alunos em cada turma. As aulas foram ministradas nas instalações da UNIOESTE, no Parque Tecnológico Itaipu – PTI. Durante a realização do curso os alunos mostraram-se bastante interessados, questionando e buscando esclarecer dúvidas referentes ao conteúdo, fora do horário de aula, sempre que podiam ou que necessitavam os alunos nos procuravam para esclarecer um ou outro ponto que não havia ficado claro, e também buscavam auxílio com os exercícios, não somente aqueles propostos na apostila do Pré-Cálculo, mas também aqueles relativos à própria disciplina de Cálculo I. O índice de frequência dos acadêmicos às aulas do Pré-Cálculo ficou em torno de 60 %.

A avaliação dos alunos participantes deu-se através de questionários que testavam não só o conteúdo matemático, mas também o raciocínio e a interpretação das questões. Para acompanhar o rendimento destes acadêmicos que participaram do curso Pré-Cálculo e verificar se o curso realmente surtiu algum efeito, estarão sendo realizadas entrevistas com alunos que participaram e também estará sendo verificado o aproveitamento dos mesmos na disciplina, este aproveitamento estará sendo avaliado de forma quantitativa e também qualitativa. Buscaremos ainda criar novas propostas de auxílio do processo de ensino e aprendizagem do Cálculo, não somente no âmbito da sala de aula, bem como fora dela.

A realização da pesquisa reflete a preocupação e o incômodo que sentimos diante da situação, e esperamos estar contribuindo para uma melhora no processo de ensino e aprendizagem do Cálculo Diferencial e Integral como um todo.

#### **Referências Bibliográficas:**

CATAPANI, E. C. **Cálculo em Serviço: Um Estudo Exploratório**. In BORBA, M. C; BICUDO, M. A. V. (org) Bolema, Ano 14, nº 16. UNESP, Rio Claro, 2001.

# PARAMETRIZAÇÃO DE PARÁBOLAS EM COORDENADAS POLARES

Manuel J. M. Negrón, Ruben E. P. Pazos.

Universidade de Santa Cruz do Sul - UNISC, [pepe@unisc.br](mailto:pepe@unisc.br);

Universidade de Santa Cruz do Sul - UNISC, [rpazos@unisc.br](mailto:rpazos@unisc.br)

**Palavras-chave:** Parametrização, Parábolas, Coordenadas polares.

## Resumo:

Neste trabalho se estuda a representação da parábola em coordenadas polares. Para isso se analisam apenas as propriedades geométricas da parábola, procurando a forma mais genérica. A prova é feita em várias etapas, considerando primeiro que a parábola é côncava para cima, com vértice nos quadrantes 1 ou 2, e que tem valor  $p > 0$ . Depois se considera que a parábola é côncava para cima, com vértice nos quadrantes 3 ou 4, e que tem valor  $p > 0$ .

## 1. Introdução

O interesse pelo cálculo de áreas de regiões fechadas no plano surge em muitas áreas, e cresce dia a dia pelos problemas que aparecem especialmente no processamento de imagens. A aproximação mais elementar é construir uma região poligonal que aproxime a região dada, ou ainda aplicar o Teorema de Green no plano para fronteiras de fácil manipulação por integrais de linha. O emprego de fórmulas que envolvam coordenadas polares não é muito extenso. Além disso, a importância de realizar regressões de tipo quadrático em coordenadas polares revela sua utilidade em alguns problemas no processamento de imagens, especialmente na área da biofísica. Neste trabalho se aplicam fórmulas que envolvem expressões em coordenadas polares  $r, \theta$ . Para isso, se emprega uma expressão geral da parábola nestas coordenadas, cuja derivação é o tópico da próxima seção.

## 2. Derivação da fórmula geral da parábola em coordenadas polares

É fato conhecido (MURDOCH) que a representação polar da parábola, cuja equação em coordenadas cartesianas é:

$$y + p = \frac{1}{4p} x^2$$

onde *Latus Rectum* =  $4p$  e  $p > 0$

é dada pela expressão 
$$r(\theta) = \frac{2 \cdot p}{1 - \text{sen}(\theta)}$$

isto é, se conhece um resultado particular (para quando o foco coincide com o **polo**).

Agora vamos deduzir a forma polar de uma parábola qualquer

$$y - y_0 = \frac{1}{4p} (x - x_0)^2$$

Por enquanto, a única suposição que fazemos é que o vértice  $V = (x_0, y_0)$  da parábola esteja no primeiro ou no segundo quadrante e que  $p > 0$  (Figura 1).

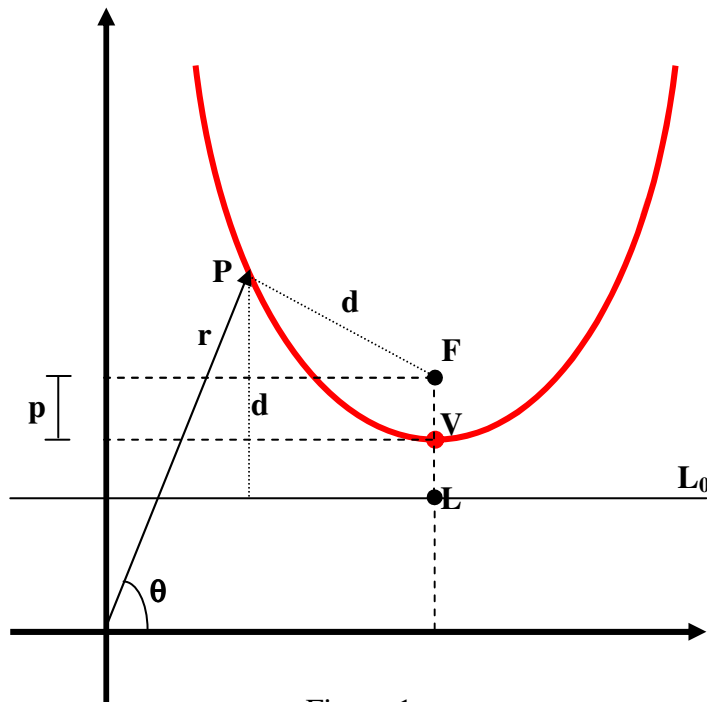


Figura 1

Se representarmos os pontos  $F$ ,  $P$ ,  $V$  e  $L$  de maneira complexa:

$$F = r_F \exp(\theta_F i) = r_F \cos(\theta_F) + r_F \text{sen}(\theta_F) i$$

$$P = r \exp(\theta i) = r \cos(\theta) + r \text{sen}(\theta) i$$

$$V = F - p i = r_F \cos(\theta_F) + (r_F \text{sen}(\theta_F) - p) i$$

$$L = V - p i = r_F \cos(\theta_F) + (r_F \text{sen}(\theta_F) - 2p) i$$

onde  $r_F = |F|$ ,  $r = |P|$ ,  $\theta_F = \arg(F)$  e  $\theta = \arg(P)$ , temos que

$$\begin{aligned} d &= |P - F| = |r \cos(\theta) + r \operatorname{sen}(\theta) i - r_F \cos(\theta_F) - r_F \operatorname{sen}(\theta_F) i| \\ &= |r \cos(\theta) - r_F \cos(\theta_F) + (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F)) i| \\ &= \sqrt{(r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F))^2} \end{aligned} \quad (1)$$

e

$$\begin{aligned} d &= \operatorname{dist}(P, \text{reta } L_0) = r \operatorname{sen}(\theta) - \operatorname{Im}(L) \\ &= r \operatorname{sen}(\theta) - (r_F \operatorname{sen}(\theta_F) - 2p) = r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F) + 2p \end{aligned} \quad (2)$$

Segue de (1) e (2) que

$$\sqrt{(r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F))^2} = r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F) + 2p \quad (3)$$

Elevando ao quadrado ambos membros de (3) e simplificando adequadamente obtemos a seqüência de resultados:

$$\begin{aligned} (r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F))^2 &= (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F) + 2p)^2 \\ (r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F))^2 &= ((r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F)) + 2p)^2 \\ (r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F))^2 &= \\ &= (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F))^2 + 4p(r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F)) + 4p^2 \\ (r \cos(\theta) - r_F \cos(\theta_F))^2 &= 4p(r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F)) + 4p^2 \\ r^2 \cos^2(\theta) - 2r r_F \cos(\theta) \cos(\theta_F) + r_F^2 \cos^2(\theta_F) &= 4p r \operatorname{sen}(\theta) - 4p r_F \operatorname{sen}(\theta_F) + 4p^2 \\ r^2 \cos^2(\theta) - 2r r_F \cos(\theta) \cos(\theta_F) + r_F^2 \cos^2(\theta_F) - 4p r \operatorname{sen}(\theta) + 4p r_F \operatorname{sen}(\theta_F) - 4p^2 &= 0 \\ \cos^2(\theta) r^2 - (2r r_F \cos(\theta) \cos(\theta_F) + 4p \operatorname{sen}(\theta)) r + r_F^2 \cos^2(\theta_F) + 4p r_F \operatorname{sen}(\theta_F) - 4p^2 &= 0 \end{aligned} \quad (4)$$

Resolvendo a equação (4) para  $r$  tem-se:

$$\begin{aligned} r &= \frac{2r_F \cos(\theta) \cos(\theta_F) + 4p \operatorname{sen}(\theta)}{2 \cos^2(\theta)} \pm \\ &\quad \frac{\sqrt{(2r_F \cos(\theta) \cos(\theta_F) + 4p \operatorname{sen}(\theta))^2 - 4 \cos^2(\theta) (r_F^2 \cos^2(\theta_F) + 4p r_F \operatorname{sen}(\theta_F) - 4p^2)}}{2 \cos^2(\theta)} \end{aligned}$$

ou simplesmente

$$\begin{aligned}
r &= \frac{2r_F \cos(\theta) \cos(\theta_F) + 4p \sin(\theta) \pm \sqrt{16pr_F \cos(\theta) \cos(\theta_F) \sin(\theta) + 16p^2 - 16pr_F \cos^2(\theta) \sin(\theta_F)}}{2\cos^2(\theta)} \\
&= \frac{2r_F \cos(\theta) \cos(\theta_F) + 4p \sin(\theta) \pm \sqrt{16pr_F \cos(\theta) [\cos(\theta_F) \sin(\theta) - \cos(\theta) \sin(\theta_F)] + 16p^2}}{2\cos^2(\theta)} \\
&= \frac{2r_F \cos(\theta) \cos(\theta_F) + 4p \sin(\theta) \pm \sqrt{16pr_F \cos(\theta) \sin(\theta - \theta_F) + 16p^2}}{2\cos^2(\theta)} \quad (5)
\end{aligned}$$

Assim, temos que

$$r(\theta) = \frac{2r_F \cos(\theta) \cos(\theta_F) + 4p \sin(\theta) + \sqrt{16pr_F \cos(\theta) \sin(\theta - \theta_F) + 16p^2}}{2\cos^2(\theta)} \quad (6)$$

ou

$$r(\theta) = \frac{2r_F \cos(\theta) \cos(\theta_F) + 4p \sin(\theta) - \sqrt{16pr_F \cos(\theta) \sin(\theta - \theta_F) + 16p^2}}{2\cos^2(\theta)} \quad (7)$$

Se considerarmos a equação geral de uma parábola:

$$y = a x^2 + b x + c \quad (8)$$

então temos que

$$p = \frac{1}{4a}, \quad V = -\frac{b}{2a} + (c - \frac{b^2}{4a}) i \quad \text{e} \quad F = V + p i = -\frac{b}{2a} + (p + c - \frac{b^2}{4a}) i$$

**Exemplo 1.** Consideremos a parábola

$$y = \frac{1}{8} x^2 - \frac{3}{4} x + \frac{41}{8} = \frac{1}{4(2)} (x - 3)^2 + 4$$

Assim, neste caso,  $p = 2$ ,  $V = 3 + 4 i$  e  $F = 3 + 6 i$  e temos

$$RF = r_F = |F| = 3\sqrt{5} \quad \text{e} \quad AF = \theta_F = \arg(F) = \arctan(2)$$

Por isso, neste caso, a parametrização da parábola (8) vem dada por:

$$\begin{aligned}
r(\theta) &= \frac{2r_F \cos(\theta) \cos(\theta_F) + 4p \sin(\theta) \pm \sqrt{16pr_F \cos(\theta) \sin(\theta - \theta_F) + 16p^2}}{2\cos^2(\theta)} \\
&= \frac{6\cos(\theta) + 8\sin(\theta) \pm 4\sqrt{6\sqrt{5}\cos(\theta) \sin(\theta - \arctan(2)) + 4}}{2\cos^2(\theta)}
\end{aligned}$$

de onde



$$r(\theta) = \frac{6\cos(\theta) + 8\sin(\theta) + 4\sqrt{6\sqrt{5}\cos(\theta)\sin(\theta - \arctan(2)) + 4}}{2\cos^2(\theta)}$$

ou

$$r(\theta) = \frac{6\cos(\theta) + 8\sin(\theta) - 4\sqrt{6\sqrt{5}\cos(\theta)\sin(\theta - \arctan(2)) + 4}}{2\cos^2(\theta)}$$

Agora, se supomos que o vértice  $V = (x_0, y_0)$  esteja no terceiro ou no quarto quadrante e que  $p > 0$  (Figura 2).

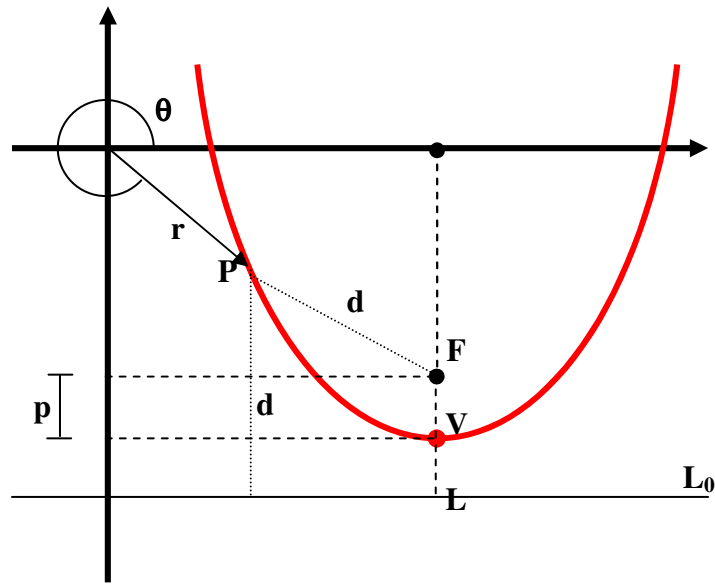


Figura 2

Neste caso, porque

$$\begin{aligned} d &= |P - F| = |r \cos(\theta) + r \sin(\theta) i - r_F \cos(\theta_F) - r_F \sin(\theta_F) i| \\ &= \sqrt{(r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \sin(\theta) - r_F \sin(\theta_F))^2} \end{aligned} \quad (9)$$

e

$$\begin{aligned} d &= \text{dist}(P, \text{reta } L_0) = \text{Im}(L) + r \sin(2\pi - \theta) = \text{Im}(L) - r \sin(\theta) \\ &= (r_F \sin(\theta_F) - 2p) - r \sin(\theta) = r_F \sin(\theta_F) - 2p - r \sin(\theta) \end{aligned} \quad (10)$$

então segue de [8] e [9] que

$$\sqrt{(r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \sin(\theta) - r_F \sin(\theta_F))^2} = r_F \sin(\theta_F) - 2p - r \sin(\theta)$$

e como na prova anterior segue, outra vez, que

$$r(\theta) = \frac{2r_F \cos(\theta) \cos(\theta_F) + 4p \sin(\theta) \pm \sqrt{16pr_F \cos(\theta) \sin(\theta - \theta_F) + 16p^2}}{2\cos^2(\theta)} \quad (11)$$

Agora vamos deduzir a forma polar de uma parábola qualquer

$$x - x_0 = \frac{1}{4p}(y - y_0)^2$$

Por enquanto, a única suposição que fazemos é que o vértice  $V = (x_0, y_0)$  esteja no primeiro ou no quarto quadrante e que  $p > 0$  (Figura 3).

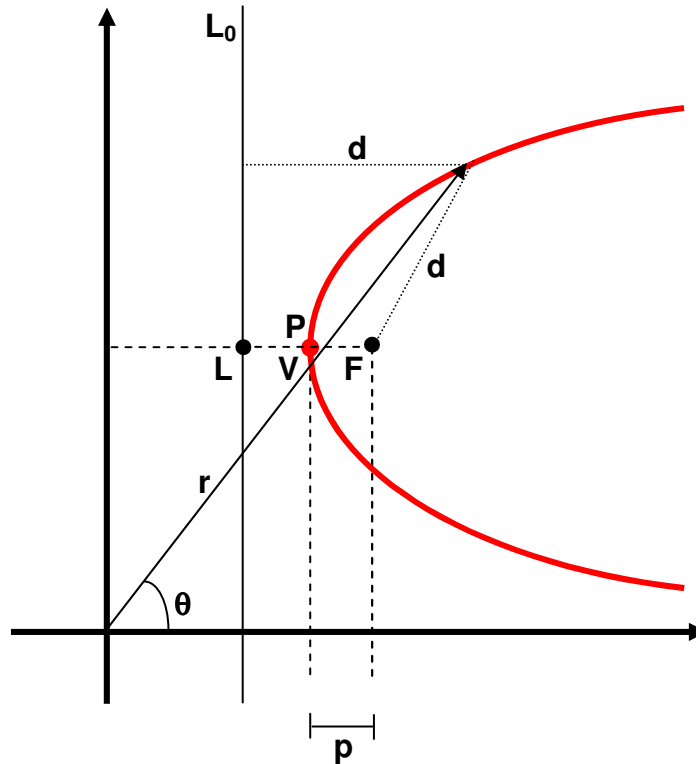


Figura 3

Porque, neste caso,

$$\begin{aligned} d &= |P - F| = |r \cos(\theta) + r \operatorname{sen}(\theta) i - r_F \cos(\theta_F) - r_F \operatorname{sen}(\theta_F) i| \\ &= |r \cos(\theta) - r_F \cos(\theta_F) + (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F)) i| \\ &= \sqrt{(r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \operatorname{sen}(\theta) - r_F \operatorname{sen}(\theta_F))^2} \end{aligned} \quad (12)$$

e

$$\begin{aligned} d &= \operatorname{dist}(P, \text{reta } L_0) = r \cos(\theta) - \operatorname{Re}(L) \\ &= r \cos(\theta) - (r_F \cos(\theta_F) - 2p) = r \cos(\theta) - r_F \cos(\theta_F) + 2p \end{aligned} \quad (13)$$

então segue de (12) e (13) que

$$\sqrt{(r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \sin(\theta) - r_F \sin(\theta_F))^2} = r \cos(\theta) - r_F \cos(\theta_F) + 2p \quad (14)$$

Elevando ao quadrado ambos membros de (14) e simplificando adequadamente obtemos a seqüência de resultados:

$$\begin{aligned} (r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \sin(\theta) - r_F \sin(\theta_F))^2 &= (r \cos(\theta) - r_F \cos(\theta_F) + 2p)^2 \\ (r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \sin(\theta) - r_F \sin(\theta_F))^2 &= (r \cos(\theta) - r_F \cos(\theta_F) + 2p)^2 \\ (r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \sin(\theta) - r_F \sin(\theta_F))^2 &= ((r \cos(\theta) - r_F \cos(\theta_F)) + 2p)^2 \\ (r \cos(\theta) - r_F \cos(\theta_F))^2 + (r \sin(\theta) - r_F \sin(\theta_F))^2 & \\ &= (r \cos(\theta) - r_F \cos(\theta_F))^2 + 4p(r \cos(\theta) - r_F \cos(\theta_F)) + 4p^2 \\ (r \sin(\theta) - r_F \sin(\theta_F))^2 &= 4p(r \cos(\theta) - r_F \cos(\theta_F)) + 4p^2 \\ r^2 \sin^2(\theta) - 2r r_F \sin(\theta) \sin(\theta_F) + r_F^2 \sin^2(\theta_F) &= 4p r \cos(\theta) - 4p r_F \cos(\theta_F) + 4p^2 \\ r^2 \sin^2(\theta) - 2r r_F \sin(\theta) \sin(\theta_F) + r_F^2 \sin^2(\theta_F) - 4p r \cos(\theta) + 4p r_F \cos(\theta_F) - 4p^2 &= 0 \\ \sin^2(\theta) r^2 - (2r r_F \sin(\theta) \sin(\theta_F) + 4p \cos(\theta)) r + r_F^2 \sin^2(\theta_F) + 4p r_F \cos(\theta_F) - 4p^2 &= 0 \end{aligned} \quad (15)$$

Resolvendo a equação (15) para  $r$  tem-se:

$$r(\theta) = \frac{2r_F \sin(\theta) \sin(\theta_F) + 4p \cos(\theta) \pm \sqrt{16p r_F \sin(\theta) \sin(\theta_F - \theta) + 16p^2}}{2 \sin^2(\theta)} \quad (16)$$

Agora, se supomos que o vértice  $V = (x_0, y_0)$  esteja no segundo ou terceiro quadrante e que  $p > 0$  (Figura 2).

Se considerarmos a equação geral de uma parábola:

$$x = a y^2 + b y + c \quad (17)$$

então temos que

$$p = \frac{1}{4a}, \quad V = \left(c - \frac{b^2}{4a}\right) - \frac{b}{2a} i \quad \text{e} \quad F = V + p = \left(p + c - \frac{b^2}{4a}\right) - \frac{b}{2a} i$$

### **3. Conclusão e extensões**

O caso quando a diretriz é horizontal ou vertical foi analisado, ficando apenas o caso quando a diretriz não é paralela aos eixos cartesianos, mas isso se resolve mediante a rotação e translação, o qual é uma tarefa mais de cálculo.

O uso desta forma de parametrização permite resolver problemas da computação gráfica, quando se empregam regressões de tipo quadrático entre o raio vetor e o ângulo  $\theta$ .

#### **Referências Bibliográficas:**

MURDOCH, DAVID C.: **Geometria Analítica**, Livros Técnicos e Científicos, Rio de Janeiro, (1977).

STRANG, Gilbert, **Calculus**, Wellesley-Cambridge Press, (1991).

# CÓDIGOS PERFEITOS EM ESPAÇOS DE HAMMING

Jerry Anderson Pinheiro<sup>1</sup>, Luciano Panek

Universidade Estadual do Oeste do Paraná - UNIOESTE, Campus Foz do Iguaçu

jerrysuper@gmail.com, lucpanek@gmail.com

**Palavras-chave:** Códigos corretores de erros, códigos de Hamming, códigos de Golay.

## Resumo:

Seja  $\mathbb{F}_q$  um corpo finito e  $\mathbb{F}_q^n$  o espaço vetorial  $n$ -dimensional das  $n$ -uplas sobre  $\mathbb{F}_q$ . Diremos que um subconjunto  $C \subseteq \mathbb{F}_q^n$  é um código linear se este for um subespaço vetorial de  $\mathbb{F}_q^n$ . O número de coordenadas distintas entre  $x, y \in \mathbb{F}_q^n$  é a chamada distância de Hamming entre  $x$  e  $y$ , e denotada por  $d(x, y)$ . A distância mínima de um código  $C$  é o mínimo das distância de Hamming entre vetores distintos de  $C$ . Se a dimensão de  $C$  é igual a  $k$  e  $d$  é a distância mínima de  $C$ , diremos que  $C$  é um  $[n, k, d]$  código linear. Um código  $C$  é dito um código perfeito se cada vetor de  $\mathbb{F}_q^n$  estiver contido em  $B(c; t)$  para algum  $c \in C$  com  $t := \lfloor (d-1)/2 \rfloor$ . O inteiro  $t$  é máximo com a propriedade de que as bolas de raio  $t$  centradas nos vetores de  $C$  são disjuntas. O número  $t$  é a chamada capacidade de correção de erros do código  $C$ .

Sabendo que cada bola contém  $\sum_{i=0}^t \binom{n}{i} (q-1)^i$  vetores, e supondo que  $C$  possua  $M$  palavras-código, temos que

$$M \leq \frac{q^n}{\sum_{i=0}^t \binom{n}{i} (q-1)^i}.$$

Essa desigualdade é conhecida como limitante de Hamming. Quando o limitante é atingido, temos que  $C$  é um código perfeito (nesta condição as bolas são disjuntas e cobrem todo o espaço).

Os códigos de Hamming foram apresentados independentemente por Marcel J. E. Golay em 1949 e Richard Hamming em 1950. Essa classe de códigos lineares possuem

---

<sup>1</sup>Bolsista de Iniciação Científica PTI-PDTA.

parâmetros  $[(q^r - 1)/(q - 1), (q^r - 1)/(q - 1) - r, 3]$  com  $r \geq 2$ . Por meio de equivalência monomial não é difícil mostrar que todos os códigos sobre  $\mathbb{F}_q$  com parâmetros iguais aos códigos de Hamming são equivalentes aos códigos de Hamming. Esses códigos possuem a capacidade de correção de um erro e admitem um mecanismo simples de decodificação, mais ainda, os códigos de Hamming são perfeitos.

Os códigos de Golay foram desenvolvidos por Marcel J. E. Golay em 1949. Essa classe de códigos lineares possuem duas ramificações, a dos códigos binários com parâmetro  $[23, 12, 7]$  e capacidade de correção de no máximo três erros, e a ramificação dos códigos de Golay ternário que possuem parâmetros  $[11, 6, 5]$  e podem corrigir até dois erros. Utilizando o limitante de Hamming não é difícil ver que esses códigos são perfeitos.

Outros códigos perfeitos sobre espaços de Hamming conhecidos são os códigos triviais, e são eles: os subespaços triviais de  $\mathbb{F}_q^n$ , os códigos binários de repetição de comprimento par e os códigos binários de comprimento par consistindo de um vetor  $c$  e seu complementar  $\bar{c}$ .

A classificação dos códigos lineares sobre espaços de Hamming está completa. Em 1973 Tietäväinen deu a solução para esse problema. Porém a classificação dos códigos não lineares com mesmos parâmetros de um código de Hamming é um problema ainda em aberto na Teoria dos Códigos.

Em nosso trabalho apresentamos a classificação dos códigos perfeitos lineares sobre espaços de Hamming, mostrando assim que os únicos códigos não triviais perfeitos sobre esse espaço são os códigos de Hamming e de Golay.

**Teorema 1** (*Tietäväinen e Van Lint*) *Um código perfeito não trivial sobre  $\mathbb{F}_q$  deve ter os mesmos parâmetros  $[n, k, d]$  dos códigos de Hamming ou Golay.*

### **Referências Bibliográficas:**

HEFEZ, A.; VILLELA, Maria Lúcia T.: **Códigos Corretores de Erros**, IMPA, Rio de Janeiro, (2002).

MACWILLIAMS, Florence J.; SLOANE, Neil J. A.: **The Theory of Error-Correcting**

**Codes**, North Holland, New York, (1977).

HUFFMAN W. C.; PLESS V.: **Fundamental of Error-Correcting Codes**, Cambridge University Press. (2003).

# CONJUNTOS PARCIALMENTE ORDENADOS, LEMA DE ZORN E AXIOMA DA ESCOLHA

Luciano Panek, Diego Diéferson Apolinário.

Universidade Estadual do Oeste do Paraná - UNIOESTE, lupanek@gmail.com; Universidade  
Estadual do Oeste do Paraná - UNIOESTE, didiapolinario@bol.com.br

**Palavras-chave:** Conjuntos Ordenados, Axioma da Escolha, Lema de Zorn.

## **Resumo:**

Neste trabalho apresentamos uma introdução à teoria dos conjuntos parcialmente ordenado dando enfoque aos clássicos lema de Zorn, axioma da escolha e lema de Zermelo.

Iniciamos o trabalho com uma introdução aos conceitos básicos de conjuntos ordenados, definindo o que é um conjunto parcialmente ordenado e utilizando-se desta definição trabalhamos com vários exemplos. Em muitas situações a realização geométrica de uma ordem parcial mostra-se eficiente. Neste sentido apresentamos uma introdução aos diagramas de Hasse, que são por excelência tipos especiais de grafos. Na seqüência definimos o conceito de fecho transitivo e funções que preservam ordem.

Prosseguindo no trabalho definimos o que é uma cadeia, conjunto indutivamente ordenado, limite superior, elemento maximal. Em seguida enunciamos o lema de Zorn que diz o seguinte: Seja  $P$  um conjunto indutivamente ordenado. Então  $P$  contém um elemento máximo  $M$ .

O lema de Zorn uma ferramenta padrão para matemáticos, sendo usado por exemplo na álgebra para estabelecer a existência de ideais máximos e na análise funcional. Neste trabalho nós assumimos o lema de Zorn como axioma e obtemos como consequência o axioma da escolha e o lema de Zermelo.

Após termos enunciado o lema de Zorn obtemos o axioma da escolha como descrito acima, definimos os conjuntos bem ordenados e demonstramos várias proposições sobre conjuntos bem ordenados. Na seqüência obtemos, também a partir do lema de Zorn, o teorema da boa ordenação.

Posteriormente definimos anti cadeias, filtro e por fim concluímos o trabalho descrevendo o teorema de decomposição de cadeias de Dilworth.



**Referências Bibliográficas:**

BERND S.W. SCHRÖDER .: **Ordered Sets – An Introduction**, Birkhäuser 2003.

GERÔNIMO, JOÃO ROBERTO.: **Fundamentos de Matemática: uma introdução à lógica matemática, teoria dos conjuntos, relações e funções**, Gerônimo, Valdeni Soliani Franco. – Maringá, PR: Eduem, 2006.

DOMINGUES, HYGINO H.; IEZZI, GELSON.: **Álgebra Moderna**, 3ª Edição, Atual Editora, 1995.

# MODELAGEM PROBABILÍSTICA NO DIMENSIONAMENTO DA PROGRAMAÇÃO DE UM SEMÁFORO

Luz Delicia Castillo Villalobos, Onerio Cambuzzi Filho.

Universidade Tecnológica Federal do Paraná - UTFPR, luz\_delicia@yahoo.com.br;

Universidade Tecnológica Federal do Paraná - UTFPR, onerio@yahoo.com.br

**Palavras-chave:** probabilidades, dimensionamento, semáforo.

## **Resumo:**

O objetivo do presente trabalho de iniciação científica é desenvolver uma metodologia para o estudo e à análise dos fluxos veiculares que chegam a um cruzamento cujo tráfego é crítico e precisa ser controlado por semáforo. Uma das ferramentas para o desenvolvimento desta metodologia é a modelagem probabilística, que permite identificar a distribuição dos fluxos veiculares no cruzamento observado. Como consequência, desta análise poderão determinar-se os fluxos por unidade de tempo mais freqüente que servirão de base para o dimensionamento da programação do semáforo a ser instalado. Os fluxos veiculares que serão utilizados na análise serão observados em horário de pico (RIBEIRO, 1991).

Quando se instalam semáforos, para colocá-los em operação, a programação do tempo de ciclo e suas repartições de verde, vermelho e amarelo precisam de alguns critérios em relação ao fluxo de veículos que passam pela interseção controlada, como por exemplo, em alguns casos é usado o fluxo máximo em outros casos é usado o fluxo médio por unidade de tempo e assim por diante.

Na presente pesquisa, o critério está relacionado com a distribuição de probabilidade da variável discreta “fluxo de veículos”. Informações sobre distribuição de probabilidades podem ser encontradas em Feller (1976), Larson (1982) e Velis (1983). A distribuição será dividida em quartis. Logo será selecionado o terceiro quartil onde será identificado o fluxo com maior probabilidade de ocorrência por unidade de tempo. Este critério foi utilizado devido à característica do fenômeno, já que os fluxos mais altos estão no terceiro quartil, tais que o fluxo veicular com maior probabilidade de ocorrência por unidade de tempo no terceiro quartil justificaria o cálculo do tempo de ciclo do semáforo. Isto é, numa amostra ordenada de tamanho  $n$  de fluxos de veículos

por unidade de tempo pertencente ao terceiro quartil, observados em uma aproximação, serão maiores dos  $3n/4$  dos fluxos na distribuição de probabilidades. Destes fluxos, será selecionado o fluxo com maior probabilidade de ocorrência. Isto é, se  $P(x_i = A)$  é a máxima probabilidade da distribuição no terceiro quartil, então  $A$  é o volume de veículos da aproximação observada que será selecionado para o cálculo da programação do semáforo. O número de fluxos/unidade de tempo, selecionados dependerá do número de aproximações do cruzamento que está sendo estudado.

No cálculo do dimensionamento da programação do semáforo será utilizado o método de Webster e Cobbe (1966). Eles desenvolveram uma expressão para o tempo de ciclo em relação ao atraso total do cruzamento. Este tempo é chamado de ciclo ótimo, ou seja, aquele que causa o mínimo atraso total no cruzamento. Esta expressão é:

$$C_0 = \frac{1,5Tp + 5}{1 - Y}, \text{ onde: } C_0 \text{ é tempo de ciclo ótimo (s); } Tp \text{ é o tempo total perdido}$$

no cruzamento durante um ciclo (s);  $Y$  é o somatório da taxa de ocupação crítica de cada fase do cruzamento. Calculado o tempo de ciclo ótimo, torna-se necessário determinar os tempos de verde efetivo de cada fase. Webster concluiu que o atraso causado ao tráfego que utiliza o cruzamento seria mínimo se o tempo total de verde efetivo do ciclo ótimo fosse distribuído proporcionalmente às taxas de ocupação crítica. Maiores informações sobre programação de semáforos podem ser encontrados em Contran-Denatran (1979).

A metodologia desenvolvida foi aplicada na cidade de Pato Branco, PR., no cruzamento entre a Rua Itacolomi com a Rua Paraná. O cruzamento selecionado é composto por duas vias de mão dupla com quatro aproximações e doze movimentos, controlados por semáforo operando a tempo fixo modo isolado. Os fluxos veiculares foram observados em horários de pico (16:30 a 17:30) em intervalos de 10 minutos durante quatro dias da semana de terça a sexta-feira. Nestes dados, para sua melhor manipulação, cada intervalo observado de 10 minutos foi convertido a veículos por minuto. O gráfico da distribuição dos dados nos sugere que a tendência da variável

aleatória “fluxo de veículos” segue uma distribuição de Poisson:  $P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!}$ ,

onde  $k$  são as diferentes categorias medidas em veículos/minuto.

Para verificar nossa hipótese serão realizados dois testes: o teste de aderência de chi-quadrado e o de Kolmogorov-Smirnov, que podem ser consultados em Barbeta, Menezes e Bornia (2004). O teste de chi-quadrado se inicia com o cálculo do estimador

$\lambda$  do parâmetro da distribuição de Poisson, dada a seguir:  $\lambda = \frac{\sum_{i=0}^n x_i * f_i}{\sum_{i=0} f_i} = 2,875$ , onde:

$f_i$  é a frequência absoluta da  $i$ -ésima observação;  $x_i$ , é a  $i$ -ésima observação da variável aleatória  $X$  (número de veículos/minuto) e  $n$  é o numero de observações. Em seguida serão calculadas as probabilidades para cada valor de  $X$ , possíveis frequências teóricas.

$$p_1 = P(X = 0) = e^{-(2,875)} = 0,05642; \quad p_2 = P(X = 1) = \frac{e^{-(2,875)} 2,875}{1!} = 0,1622$$

$$p_3 = P(X = 2) = \frac{e^{-(2,875)} (2,875)^2}{2!} = 0,233; \quad p_4 = P(X = 3) = \frac{e^{-(2,875)} (2,875)^3}{3!} = 0,223$$

$$p_5 = P(X \geq 4) = 1 - P(X < 4) = 1 - 0,66 = 0,3247$$

Os resultados das frequências observadas e teóricas da variável aleatória são apresentados no quadro 1 e o gráfico de suas distribuições na figura 1.

Em seguida calcula-se a estatística  $D^2$ , que mede a aderência à distribuição especificada:  $D^2 = \sum_{i=1}^5 \frac{(f_i - np_i)^2}{np_i} = 2,336$ , Onde  $n$  é o número de observações da amostra;  $p_i$  é a probabilidade da  $i$ -ésima observação.

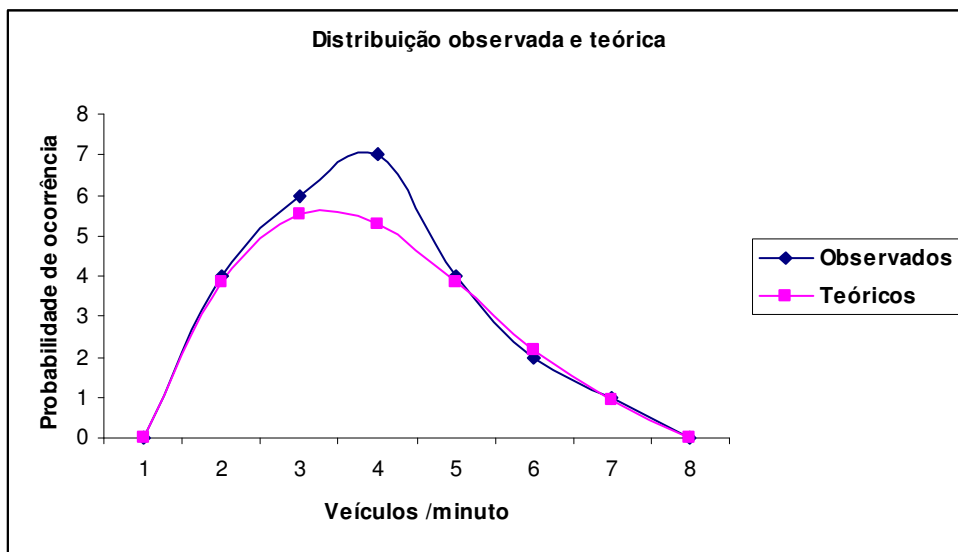
Para construir a regra de decisão, obtém-se o valor crítico  $X_c^2$  da distribuição chi-quadrado, com  $gl=6$  graus de liberdade e com um nível de significância do 5%, o qual leva à  $P(D^2 > X_c^2) = 0,05$  de rejeitar a hipótese de que a variável  $X$  segue uma distribuição de Poisson. Segundo a tabela de chi-quadrado, obtém-se  $X_c^2 = 12,59$  e  $D^2 = 2,336$ . Como  $D^2$  é menor que  $X_c^2$ , então cai na região de aceitação da distribuição chi-quadrado. Portanto, aceitamos a hipótese de que a distribuição dos fluxos veiculares segue uma distribuição de Poisson. O teste Kolmogorov-Smirnov também confirma a hipótese, portanto se tem:

$$P(X = k) = \frac{e^{-2,875} 2,875^k}{k!}, k = 0, 1, 2, \dots, 7$$

Estes resultados correspondem à aproximação 1 do cruzamento. O mesmo procedimento foi realizado nas três outras aproximações onde também se mostrou que seguem uma distribuição de Poisson.

Nº Veículos por minuto	Frequência Absoluta ( $f_i$ )	Frequência Acumulada ( $F_i$ )	Frequência relativa observada		Frequência Relativa Teórica ( $p_i$ )	Frequência Teórica
			Simple (r)	Acumulada ( $R_i$ )		
0	0	0	0,00000	0,0000	0,05640	1,4
1	4	4	0,16667	0,1667	0,16220	3,9
2	6	10	0,25000	0,4167	0,23320	5,6
3	7	17	0,29167	0,7083	0,22350	5,4
4	4	21	0,16667	0,8750	0,16100	3,9
5	2	23	0,08333	0,9583	0,09240	2,2
6	1	24	0,04167	1,0000	0,04430	1,1
Toral	24		1,00000		0,97300	23,4

**Quadro 1-** Frequências observada e teórica dos fluxos veiculares.



**Figura 1 –** Distribuição observada e teórica

Tendo certeza das condições anteriores, procede dimensionar o plano semaforico de modo que o tempo de ciclo seja ótimo, cujos resultados são dados no quadro 2, onde também se apresenta a comparação com a programação em operação.

	Programação proposta	Programação em operação
Tempo de Ciclo (s)	44	42
Tempo de verde real do estagio 1 (s)	18	14
Tempo de verde real do estagio 2 (s)	20	24
Tempo de amarelo/ fase (s)	3	2

**Quadro 2 –** Comparação das programações do semáforo da interseção selecionada

Concluindo, a metodologia desenvolvida permitiu identificar que a distribuição da variável aleatória fluxo veicular nas quatro aproximações segue uma distribuição de Poisson com parâmetros  $\lambda_1 = 2,875$ ;  $\lambda_2 = 3,7$ ;  $\lambda_3 = 2,8$ ;  $\lambda_4 = 4,8$  veic./min., respectivamente. Tendo identificado a distribuição dos fluxos que chegam na interseção, poderá determinar-se a probabilidade de ocorrência de certo volume de veículos que

poderiam chegar ao cruzamento por unidade de tempo. Por exemplo, observando o quadro 1, a probabilidade de chegar 4 veic./min. é de 0,161.

No trabalho desenvolvido, os fluxos selecionados para o cálculo do tempo de ciclo segundo o critério estabelecido pela metodologia proposta com maior probabilidade no terceiro quartil em cada aproximação da interseção, são:  $q_1 = 342$ ;  $q_2 = 288$ ;  $q_3 = 276$ ;  $q_4 = 384$  veic./h respectivamente.

Comparando o tempo de ciclo calculado segundo o critério da presente metodologia que é de 44 seg, com o tempo de ciclo do semáforo em operação que é de 42 seg, não existe uma diferença significativa. A diferença de dois segundos se deve que na metodologia proposta está considerando 3 segundos de luz amarela por fase cujos tempos foram dados para uma melhor segurança de motoristas e pedestres. Estes resultados implicam que o tempo de ciclo do semáforo em operação não precisa ser atualizado, mas se nota uma diferença significativa nas repartições de verde real.

Como se pôde observar, a metodologia desenvolvida permitiu definir como se pode prever a quantidade de veículos que passam pelo cruzamento durante um período de tempo determinado. Também através da presente metodologia, pode-se verificar se a programação do semáforo em operação precisa ser atualizado, possibilitando tomar decisões sobre seu funcionamento. Portanto, obteve-se um método tecnicamente definido e estruturado, capaz de resolver adequadamente os objetivos propostos.

### **Referencias bibliográfica**

BARBETA, P. A., MENEZES, M. R., BORNIA, A. C.: **Estatística para Cursos de Engenharia e Informática**, Ed. Atlas, São Paulo, 2004.

CONTRAN-DENATRAN: **Manual de Semáforos**. 1979.

FELLER, W.: **Introdução à Teoria das Probabilidades e suas Aplicações: Espaços Amostrais Discretos**, Ed. Edgard Blücher LTDA. 1976.

LARSON, H. J.: **Introduction to Probability Theory and Statistical Inference**, 1982

RIBEIRO, P.C.M.: **Influence of Flow Variation on Fixed-Time Signal Control**, Tese de Doutorado. University College London, 1991.

VELIS C. C.: **Estatística Aplicaciones**, 2<sup>o</sup> Ed. Lima- Peru, 1993.

WEBSTER, F. V.; COBBE, B. M.: **Traffic signals. Road Research Technical**, Paper, n. 56. 1966.

# **SIMULAÇÃO DA AERAÇÃO DE GRÃOS ARMAZENADOS VARIANDO AS CONDIÇÕES DO AR AMBIENTE**

Luiz Carlos Radtke, Oleg Khatchatourian, Círio Thomas.

Universidade Regional do Noroeste do Estado do Rio Grande do Sul - UNIJUÍ,  
lcradtke@yahoo.com.br; Universidade Regional do Noroeste do Estado do Rio Grande do Sul -  
UNIJUÍ, olegkha@unijui.edu.br; Universidade do Oeste de Santa Catarina - UNOESC,  
ciriothomas@yahoo.com.br

## **Resumo:**

Neste trabalho será implementado um modelo matemático para simulação da aeração de grãos variando as condições do ar ambiente, com o objetivo de prever o comportamento dos sistemas de aeração para diferentes regiões. O modelo adotado é um modelo unidimensional formulado por Thorpe (2001) algumas alterações foram feitas a fim de tornar o modelo menos computador – intensivo, para validação será comparado os dados obtidos através do modelo com dados encontrados na bibliografia.

**Palavras-chaves:** Simulação, Aeração, Grão.

## **Introdução:**

O Brasil ocupa a terceira posição na produção mundial de grãos, isto em números significa 113,6 milhões de toneladas produzidas na safra 2004/2005, e tem como principais produtores os estados do Mato Grosso, Paraná, Rio Grande do Sul, Goiás e Minas Gerais.

Mas a produção de grãos não consiste somente no plantio e na colheita, há outro elemento importante que faz parte desse processo, o qual se traduz no armazenamento desta produção.

Tendo em vista o armazenamento destes grãos, há a preocupação de manter a qualidade dos mesmos, visando a não ocorrência da deterioração dos grãos, a qual pode ser causada por diversos fatores, como ransificação, ataque de fungos e insetos, entre outros.

Diante desta situação, buscando primordialmente a manutenção da qualidade dos grãos armazenados, é utilizada a técnica de aeração, que consiste num método de arejamento periódico dos grãos.

### Modelo Matemático:

O modelo matemático usado neste trabalho é baseado no modelo formulado por Thorpe (2001). Este modelo relaciona relações psicométricas do ar com balanços de massa e de energia. Algumas mudanças foram feitas nas equações originais para simplificar o modelo. As equações diferenciais que descrevem o calor e transferência de massa são expressas como:

$$\begin{aligned} \frac{\partial \theta}{\partial t} \left\{ \rho_b [c_g + c_w^U] + \varepsilon \rho_a \left[ c_a + R \left( c_w + \frac{\partial h_v}{\partial T_a} \right) \right] \right\} = \\ = \rho_b h_s \frac{\partial U}{\partial t} - u_a \rho_a \left[ c_a + R \left( c_w + \frac{\partial h_v}{\partial T_a} \right) \right] \frac{\partial \theta}{\partial y} + \rho_b \frac{dm_s}{dt} (Q_r - 0.6h_v) \end{aligned} \quad (1)$$

$$\frac{\partial U}{\partial t} \rho_b = -\rho_a u_a \frac{\partial R}{\partial y} + \frac{dm_s}{dt} (0.6 + U) \quad (2)$$

onde:  $c_a$  é o calor específico do ar ( $\text{J kg}^{-1} \text{ } ^\circ\text{C}^{-1}$ ),  $c_w$  é o calor específico da água ( $\text{J kg}^{-1} \text{ } ^\circ\text{C}^{-1}$ ),  $h_v$  é o calor latente de vaporização da água ( $\text{J kg}^{-1}$ ),  $h_s$  é o calor diferencial de absorção ( $\text{J kg}^{-1}$ ),  $U$  é o conteúdo de umidade do grão (%) em uma base seca,  $m_s$  é a perda de matéria seca do grão (%),  $Q_r$  é o calor de oxidação do grão ( $\text{J s}^{-1} \text{ m}^{-3}$ ),  $T_a$  é a temperatura do ar em equilíbrio com o grão ( $^\circ\text{C}$ ),  $t$  é o tempo (s),  $u_a$  é a velocidade de aeração ( $\text{m s}^{-1}$ ),  $R$  é a razão de umidade do ar ( $\text{kg kg}^{-1}$ ),  $y$  é a coordenada vertical (m),  $\varepsilon$  é o coeficiente de porosidade do grão (decimal),  $\rho_a$  é a densidade do ar intergranular ( $\text{kg m}^{-3}$ ),  $\rho_b$  é a massa específica dos grãos ( $\text{kg m}^{-3}$ ) e  $\theta$  é a temperatura do grão ( $^\circ\text{C}$ ).

O modelo acima descrito foi resolvido através do método de diferenças finitas, considerou-se que a massa de grãos foi dividida em várias camadas finas na direção do fluxo do ar (direção vertical). Os limites entre as camadas foram chamados de nó conforme figura 1, durante o processo de simulação o conteúdo de umidade e a temperatura do grão são calculados a cada intervalo de tempo para cada seção de um modo repetitivo.



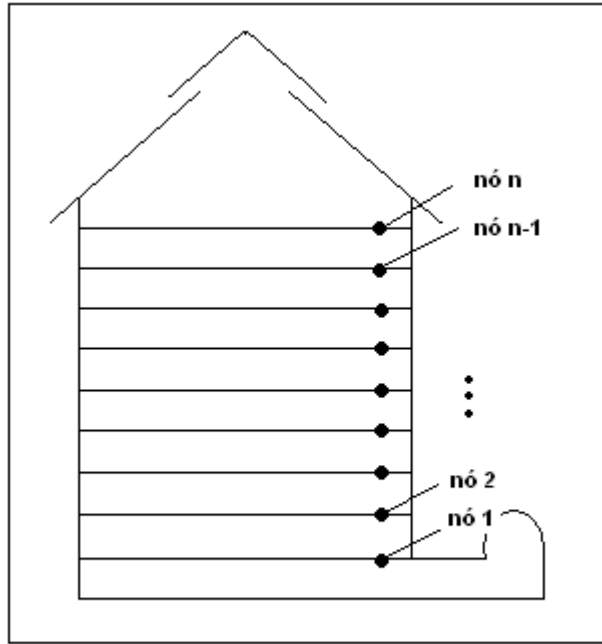


Figura 1: Mostra como foi dividido os nós no interior do silo

Neste trabalho o calor específico da água e o calor específico do ar são considerados constantes ( $4186$  e  $1000 \text{ J } ^\circ\text{C}^{-1}$ , respectivamente), e o calor específico dos grãos é calculado em cada intervalo de tempo para cada camada de grão usando a Eq.(3).

$$c_g = 1000(K + QU_p) \quad (3)$$

onde: K e Q são constantes que dependem do produto armazenado e  $U_p$  é o conteúdo de umidade em base molhada (%).

A entalpia diferencial de sorção é calculada em cada iteração aplicando a Eq.(4):

$$h_s = h_v \left( 1 + \frac{A \exp(-BU)}{(\theta + C)^2} (\theta + 273.15) \left( -5 + \frac{6800}{\theta + 273.15} \right) \right) \quad (4)$$

onde: A, B e C são valores constantes que dependem do produto armazenado.

A entalpia de vaporização ou calor latente de vaporização é obtida a cada iteração aplicando a Eq.(5).

$$h_v = 2501330 - 2363T_a \quad (5)$$

O calor de oxidação do grão é considerado uma constante seu valor é igual a  $15,778 \text{ kJ s}^{-1}\text{m}^{-3}$ , a porosidade da massa de grãos também é considerado uma constante igual a (40%) baseado em Brooker et al. (1992).

A massa específica do ar intergranular é calculada usando a Eq.(6), apresentada por Alé (2001).

$$\rho_a = \frac{352.8P_{atm}}{101.325(T_{aera} + 273.15)} \quad (6)$$

onde:  $T_{aera}$  é a temperatura de aeração e  $P_{atm}$  é a pressão barométrica.

Para calcular a taxa de perda de matéria seca é utilizado um método desenvolvido por Thompson (1972) (Eq.(7)).

$$\frac{dm_s}{dt} = \frac{14.72 \times 10^{-10} \{ \exp(1.667 \times 10^{-6} t / M_M M_T) \} + 2.833 \times 10^{-9}}{M_M M_T} \quad (7)$$

onde:  $M_M$  e  $M_T$  são parâmetros utilizados para ajustar o tempo de aeração com o teor de água e o teor de temperatura dos grão. Assim quando a temperatura dos grãos for menor ou igual a  $15,5 \text{ }^\circ\text{C}$  ou o teor de menor ou igual a 19% b.u., a Eq.(8) deve ser utilizada para calcular  $M_T$ .

$$M_T = 32,2 \exp(-0.1044\theta - 1.856) \quad (8)$$

No caso da temperatura dos grãos ser maior que  $15,5 \text{ }^\circ\text{C}$  e o teor de água estar entre 19% b.u. e 28% b.u., o valor de  $M_T$  deverá ser calculado utilizando a Eq.(9).

$$M_T = 32,2 \exp(-0.1044\theta - 1.856) + \frac{U_p - 19}{100} \exp(0.0183\theta - 0.2847) \quad (9)$$

Se a temperatura dos grãos for maior que  $15,5 \text{ }^\circ\text{C}$  e o teor de água for maior que 28% b.u.,  $M_T$  deverá ser calculado utilizando a Eq.(10).

$$M_T = 32,2 \exp(-0.1044\theta - 1.856) + 0.09 \exp(0.0183\theta - 0.2847) \quad (10)$$

O valor de  $M_M$  é obtido aplicando-se a Eq.(11).

$$M_M = 0.103 \left( \exp \left( \frac{455}{(100U)^{1.53}} \right) - 0.845U + 1.558 \right) \quad (11)$$

O modelo será implementado e seus dados serão comparados com dados encontrados na bibliografia, esperasse que o modelo possa predizer o comportamento do sistema de aeração para diferentes regiões.

### **Referências Bibliográficas:**

ALAGUSUNDARAM, K., JAYAS, D.S., WHITE, N.D.G. and MUIR, W.E.: **Finite difference model of three-dimensional heat transfer in grain bins**. Can. Agric. Eng. 32, (1990) 315-321.

ALÉ, J.F.: **Sistemas de ventilação industrial-ventiladores**. Departamento de Engenharia Mecânica e Mecatrônica, LSFM, PUCRS, (2001).

ANDRADE, E.T.: **Simulação da variação de temperatura em milho armazenado em silo metálico**. Viçosa, Minas Gerais: Imprensa Universitária, Universidade Federal de Viçosa, Tese de doutorado em Engenharia Agrícola 174pp. (2001).

BROOKER, D.B., BAKKER-ARKEMA, F.W. and HALL, C.W.: **Drying and store of grains and oilseeds**. AVI BooK, USA, 450pp. (1992).

CANCHUN, J., DA-WEN, S.E. and CHONGWEN, C.: **Computer simulation of temperature changes in a wheat storage bin**. J. Store Prod. Res. 37, (2001), 165-167.

DAVIDSON, M.R.: **Natural convection of gas/vapor mixtures in porous medium**. Int. J. Heat Mass Transfer 29, (1986), 1371-1381.

DEVILLA, I.A.: **Simulação de deterioração e de simulação de temperatura e umidade em uma massa de grãos armazenados em silos com aeração**. Viçosa Minas Gerais, Imprensa Universitária. Universidade Federal de Viçosa, Tese de Doutorado em Engenharia Agrícola 84pp, (2002).

LOPES, D.C., MARTINS, J.H., MELO, E.C. and MONTEIRO, P.M.B.: **Aeration simulation of stored grain variable air ambient conditions**. P. Biology and Technology 42, (2006), 115-120.

MUIR, W.E., FRASER, B.M. and SINHA, R.N.: 1980. **Simulation model of two-dimensional heat transfer in controlled atmosphere grain bins**. Can. Agric. Eng. 12, (1980), 21-24.

NAVARRO, S. and NOYES, R.T.: **The mechanics and physics of modern grain aeration management**. CRC Press, Boca Raton, 647 p. (2001).

SINÍCIO, R., MUIR, W.E. and JAYAS, D.S.: **Sensitivity analysis of a mathematical model to simulate aeration of wheat stored in Brazil.** *Postharvest Biol. Technol.* 11, (1997), 107-122.

# ESTIMATIVA DE ERROS DE ITERAÇÃO EM MÉTODOS DE RELAXAÇÃO

Márcio André Martins, Maria Regina Carvalho Macieira Lopes.

Universidade Estadual do Centro Oeste do Paraná - UNICENTRO, mandre@unicentro.br;

Universidade Estadual do Centro Oeste do Paraná - UNICENTRO; mrlopes@unicentro.br.

**Palavras-chave:** Erros de Iteração, Métodos de Relaxação .

## Resumo:

Ao se aplicar um método numérico iterativo na resolução de um sistema de equações, o número de iterações é um fator que deve ser considerado. O erro numérico pode ser propagado com um número muito elevado de iterações. Tratando-se de resolução de sistemas lineares de grande porte é relevante buscar alternativas para diminuir o número de repetições na execução de um processo iterativo. Ou seja, faz-se necessário o emprego de técnicas que propiciem a aceleração da taxa de convergência. Métodos que alteram a taxa de convergência iterativa são tratados na literatura por **Métodos de Relaxação** (BURDEN e FAIRE, 1997) . Considerando-se o sistema linear de equações  $AX = B \Leftrightarrow (D - L - U)X = B$ , onde  $A$  é uma matriz  $N \times N$ ,  $X, B \in \mathbb{R}^N$ ,  $D$  diagonal,  $L$  triangular inferior e  $U$  triangular superior, a forma iterativa matricial dos **Métodos de Relaxação** é dada por:  $X^{(k)} = T_\omega X^{(k-1)} + C_\omega$  onde o superescrito de  $X$  é a representação dos níveis iterativos,  $T_\omega = (D - \omega L)^{-1}[(1 - \omega)D + \omega U]$  e  $C_\omega = \omega(D - \omega L)^{-1}$ . Kahn (1958) demonstrou que **Métodos de Relaxação** não convergem se  $\omega$  estiver fora do intervalo  $(0, 2)$ . Para escolhas de  $\omega$ , com  $\omega > 1$ , os procedimentos são denominados **Métodos de Sobre-relaxação** e são empregados com o objetivo de acelerar a convergência de sistemas que têm a convergência garantida, ou seja, satisfazem os critérios de convergência do método iterativo a ser adotado. Para escolhas de  $0 < \omega < 1$ , os procedimentos são conhecidos por **Métodos de Sub-relaxação** e podem servir para

obter a convergência de alguns sistemas que não convergem ao se utilizar um determinado método iterativo. No emprego destes métodos deve-se ter a garantia de que os resultados obtidos são confiáveis. A confiabilidade de um resultado está relacionada à magnitude do erro numérico. O processo que quantifica o **erro numérico** é denominado verificação (OBERKAMPF e BLOTTNER, 1998) ou validação (MALISKA, 1995). A verificação é obtida através do emprego de **estimadores de erro**. O resultado fornecido por um estimador de erro é denominado incerteza numérica. Neste trabalho, analisa-se a eficiência de um estimador de erro para processos iterativos que sofrem a influência de Métodos de Relaxação. Este estimador calcula a incerteza numérica com base na taxa de convergência da variável de interesse. O parâmetro considerado para avaliar a qualidade das estimativas obtidas foi a **efetividade**  $\theta$ , definida pela razão entre incerteza e erro (ZHU e ZIENKIEWICZ, 1990). Se  $\theta > 1$ , a estimativa do erro é **confiável** (MARCHI, 2001). Se  $\theta \approx 1$ , a estimativa do erro é **acurada** (CHAPRA e CANALE, 1994). Inicialmente foram realizados alguns testes com a aplicação de Métodos de Relaxação em sistemas lineares de ordem 10, 20, e 30. Nestes casos, o estimador mostrou-se **acurado e confiável**. Em seguida, foram realizados testes com sistemas lineares gerados a partir da aplicação do método de diferenças finitas em um problema difusivo unidimensional com absorção de calor, o que resulta em uma Equação de Poisson. O modelo matemático deste teste é dado por:  $\frac{d^2\Lambda}{dx^2} - \frac{q}{k} = 0$  onde  $q$  representa a absorção de calor,  $k$  a condutividade térmica do material e  $\Lambda$  a variável dependente do problema, que é um escalar difundido. Ou ainda,  $\frac{d^2\Lambda}{dx^2} + f(x) = 0$ , sendo  $x$  a variável independente e  $f(x)$  o termo fonte. Considerou-se  $f(x) = -12x^2$  (INCROPERA, 1996) com as seguintes condições de contorno de Dirichlet:  $\Lambda(0) = 0$  e  $\Lambda(L) = 1$  onde  $L$  é o comprimento do domínio de cálculo ( $L = 1$ ). O resultado numérico, erro de iteração e incerteza numérica foram analisados em três nós (ou pontos) O resultado numérico, erro de iteração e incerteza numérica foram analisados em três nós (ou pontos) específicos da malha. Considerou-se condições iniciais nula e unitária e a variação do número de nós da malha (11, 101 e

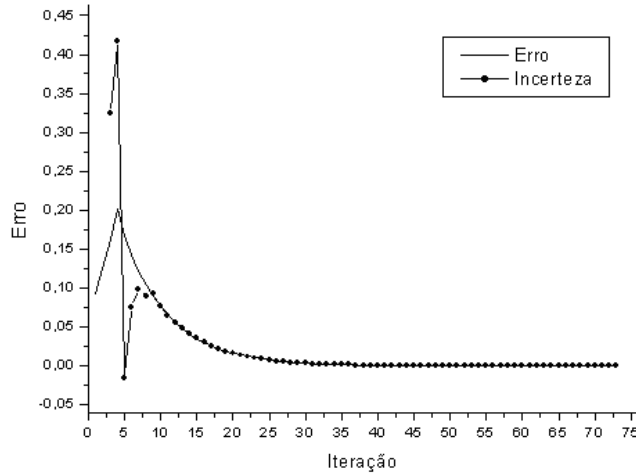


Figura 1: Estimativa do Erro Numérico,  $0 < \omega < 1, 2$

1001). A Figura 1 apresenta o comportamento da incerteza numérica para  $0 < \omega < 1, 2$ . Observa-se dois tipos de comportamento: nas iterações iniciais, as incerteza numérica é inacurada e sem confiabilidade. Após estas iterações, aproximadamente 10% do número total de iterações, as estimativas foram confiáveis e acuradas.

Para  $\omega > 1, 2$ , casos em que a aceleração de convergência é mais significativa, o erro de iteração e a incerteza não apresentaram comportamento exponencial (Figura 2).

Nestes casos, o estimador mostrou-se eficiente na maioria das iterações, ou seja, acurado e confiável. O valor médio identificado para efetividade do estimador foi de aproximadamente 1,08. Quantitativamente uma estimativa de erro é considerada acurada de acordo com o quão próximo da unidade está sua efetividade.

### Referências Bibliográficas:

BLOTTNER, F. G.: **Accurate Navier-Stokes results for the hypersonic flow over a spherical nosetip**, Journal of Spacecraft and Rockets, vol. 27, 2, (1990), 113-

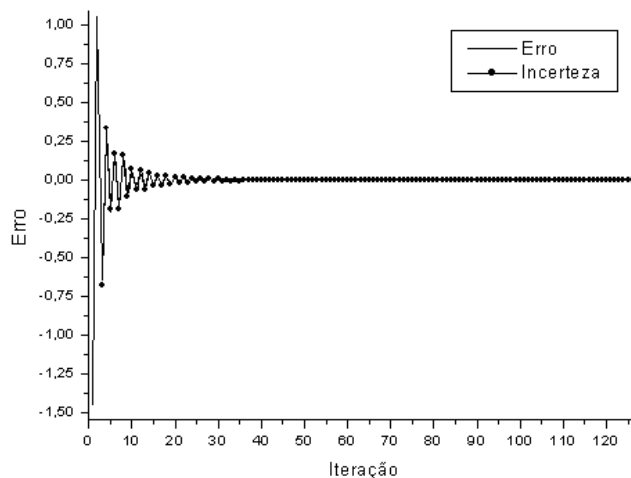


Figura 2: Estimativa do Erro Numérico,  $\omega > 1, 2$

122.

BURDEN, R. L. and FAIRES, J. D.: **Numerical Analysis**, Brooks/Cole Publishing Company, 2nd ed, (1997).

CHAPRA, S. C. and CANALE, R. P.: **Introduction to computing for Engineers**, McGraw-Hill, New York, 2nd ed, (1994).

INCOPRERA, F. P. and WITT, D. P.: **Fundamentals of Heat and Mass Transfer**, 4th ed. New York, Wiley, 4nd ed, (1996).

KAHAN, W.: **Gauss-Seidel Methods of Solving Large Systems of Linear Equations**, University of Toronto, Toronto, (1958).

MALISKA, C. R.: **Transferência de Calor e Mecânica dos Fluidos Computacional**, LTC, Rio de Janeiro, (1995).

MARCHI, C. H.: **Verificação de soluções numéricas unidimensionais em dinâmica dos fluidos**, Universidade Federal de Santa Catarina, Florianópolis, (2001).

OBERKAMPF, W. L. and BLOTTNER, F. G.: **Issues in computational fluid dynamics code verification and validation**, AIAA Journal, vol. 36, 5, (1998), 687-695.



ZHU, J. Z. and ZIENKIEWICZ, O. C.: **Superconvergence recovery technique and a posteriori error estimates**, International Journal for Numerical Methods in Engineering, vol. 30, (1990), 1321-1339.

# Condições de Otimalidade e de Qualificação para Problemas de Programação Não Linear

André Luiz Barbosa Arantes dos Santos.

Universidade Federal do Paraná - UFPR, druida.andre@gmail.com

**Palavras-chave:** programação não linear, condições de otimalidade, condições de qualificação.

**Resumo:** Estudamos problemas de programação não linear que consistem em minimizar uma função  $f$  com restrições de igualdade  $h(x) = 0$  e de desigualdade  $g(x) \leq 0$ , onde  $f : \mathbb{R}^n \rightarrow \mathbb{R}$ ,  $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$  e  $g : \mathbb{R}^n \rightarrow \mathbb{R}^p$  são funções continuamente diferenciáveis. As condições de Karush-Kuhn-Tucker (KKT) afirmam que se  $x^*$  é solução local do problema e uma condição de qualificação é satisfeita, então existem multiplicadores de Lagrange,  $\lambda \in \mathbb{R}^m$  e  $\mu \in \mathbb{R}^p$ , tais que

$$-\nabla f(x^*) = \sum_{i=1}^m \lambda_i \nabla h_i(x^*) + \sum_{i=1}^p \mu_i \nabla g_i(x^*)$$

$$\mu_i \geq 0, i = 1, \dots, p,$$

$$\mu_i g_i(x^*) = 0, i = 1, \dots, p,$$

$$h_i(x^*) = 0, i = 1, \dots, m,$$

$$g_i(x^*) \leq 0, i = 1, \dots, p.$$

A figura abaixo ilustra a situação em que há somente restrições de desigualdade. Neste caso, as condições de KKT significam que o oposto do gradiente da função objetivo pertence ao cone gerado pelos gradientes das restrições ativas. Estudar as condições de KKT através de uma abordagem cônica facilita o entendimento e a visualização geométrica. Provamos, então, as condições de KKT por meio de uma abordagem cônica supondo a condição de qualificação mais fraca possível, ou seja, a igualdade do polar dos cones

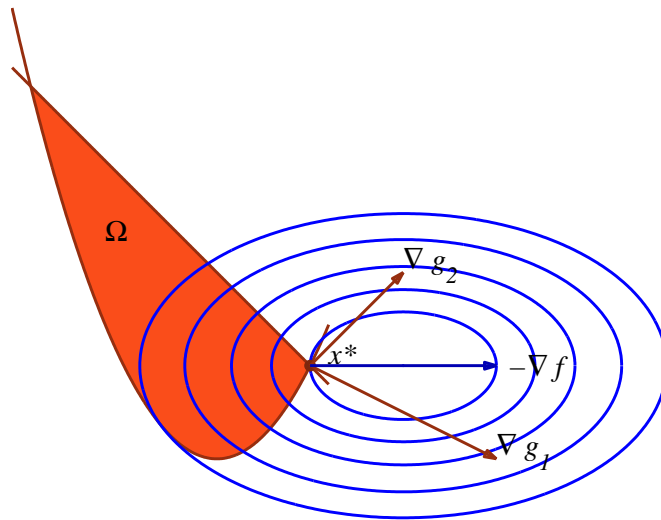


Figura 1: Condições de Karush-Kuhn-Tucker

tangente e viável linearizado. Neste contexto, o Lema de Farkas é uma ferramenta fundamental. Estudamos várias outras condições de qualificação mais fortes, porém mais fáceis de serem verificadas tais como: a condição de Slater, independência linear dos gradientes das restrições ativas (LICQ), Mangasarian-Fromovitz, e quase-regularidade. Discutimos as implicações entre elas e exibimos contra-exemplos para as recíprocas. Durante a realização do trabalho procuramos demonstrar os resultados formalmente, mas sempre enfatizando o aspecto geométrico da teoria.

### Referências Bibliográficas:

Rodrigo Garcia Eustáquio: **Condições de otimalidade e de qualificação para problemas de programação não linear**, Dissertação de mestrado, (2007).

Izmailov, A. and Solodov, M.: **Otimização - Condições de otimalidade, elementos de análise convexa e de dualidade**, IMPA, Volume 1, (2005).

Friedlander, A.: **Elementos de programação não linear**, Editora da Unicamp, Campinas, (1994).

# TÉCNICAS DE ESTATÍSTICA MULTIVARIADA APLICADAS NA CLASSIFICAÇÃO DOS CANDIDATOS AO VESTIBULAR DA FECILCAM

Tatiane Cazarin da Silva<sup>1</sup>, Gislaine Aparecida Pericaro<sup>2</sup>

**Palavras-chave:** Padrões. Estatística Multivariada. Variáveis.

O presente trabalho integra um projeto de Iniciação Científica vinculado ao Núcleo de Pesquisas Multidisciplinar da Faculdade Estadual de Ciências e Letras de Campo Mourão - FECILCAM, tendo como objetivo principal abordar técnicas estatísticas multivariadas utilizadas no processo de reconhecimento de padrões, por meio da análise do desempenho de 1157 candidatos ao Vestibular de Verão 2007 da instituição.

A educação é uma questão que preocupa a esfera social, pois atua na formação e qualificação pessoal, exigindo que todos os setores da educação trabalhem num objetivo comum: a qualidade. De acordo com Zanella [2], considerando a população e a realidade a que se destina, é possível traçar metas e projetos que melhorem o processo pelo qual se dá o ensino. E como a educação é marcada por diversas variáveis, numéricas ou latentes, relacionadas a diversos elementos, a estatística atua como ferramenta base nessa análise.

A estatística surgiu como uma ciência que buscava explicar e validar diversos fenômenos marcados numericamente, pelo levantamento de hipóteses e observações reais que conduziam a formulações matemáticas. Atualmente, é utilizada nas mais diversas áreas do conhecimento, garantindo não só a validação dos resultados, mas também a simplificação, a predição e o direcionamento do processo que descreve.

Nesse sentido, as instituições de ensino superior fornecem aos seus alunos o preenchimento de um questionário sócio-educacional, possibilitando estabelecer um perfil de seus candidatos. Nessa análise, cada elemento é marcado por características, ou variáveis, distintas que quando verificadas num conjunto, podem explicitar

---

<sup>1</sup> Email: [taticazarin@yahoo.com.br](mailto:taticazarin@yahoo.com.br), Faculdade Estadual de Ciências e Letras de Campo Mourão, Graduação em Matemática (PIC).

<sup>2</sup> Email: [gpericaro@gmail.com](mailto:gpericaro@gmail.com), Mestre, Faculdade Estadual de Ciências e Letras de Campo Mourão, Departamento de Matemática.

características diferentes do que a observada numa análise isolada, seja pela estrutura ou dependência entre as variáveis consideradas. Assim, surge a necessidade de utilizar métodos estatísticos que garantam um conhecimento geral da estrutura numérica das variáveis, por meio de uma análise multivariada que contemple os resultados de forma satisfatória. A esse processo dá-se o nome de Mineração de Dados, sendo a estatística multivariada uma das etapas finais desse. A estatística multivariada é definida por Mingoti [1] como um conjunto de métodos estatísticos utilizados em situações nas quais várias variáveis são medidas simultaneamente, sendo alistadas a um mesmo elemento amostral. Dessa forma, a estatística multivariada tem como característica fundamental a análise multivariada de informações, garantindo a estruturação das mesmas. Com isso, é possível estabelecer padrões e relações na análise proposta, possibilitando a simplificação e generalização dos resultados, ou seja, tem-se o reconhecimento de padrões, a partir da análise exploratória e a aplicação de técnicas multivariadas, como a Análise Fatorial, Análise Discriminante e Análise de Componentes Principais.

A fim de aplicar os métodos citados, e ao mesmo tempo garantir confiabilidade na manipulação dos dados, foi ponderada apenas a análise das variáveis dos candidatos que realizaram as provas, ou seja, dos 1186 inscritos, 1157 foram considerados, de acordo com as 19 variáveis que compunham o questionário. Visando sumarizar informações obtidas e evidenciar conceitos implícitos, foi utilizado como banco de dados os resultados brutos obtidos no questionário dos candidatos ao vestibular. Tais dados, organizados em uma matriz, serviram como base na aplicação da análise multivariada. Com o auxílio do *software Statistica*, foi aplicada a Análise Fatorial nas 19 variáveis verificadas. Essa técnica, que consiste na redução da dimensionalidade dos dados, garante a mesma estrutura numérica das observações iniciais, por meio da identificação dos autovalores, autovetores e os pesos fatoriais referentes, que denotam a mesma variabilidade e explicação das variáveis. Considerou-se os fatores com valores numéricos superiores a 0,85, a fim de garantir uma boa representatividade na variância explicada pelas variáveis normalizadas. Com o auxílio da técnica de Componentes Principais, que ordena as variáveis de acordo com a representatividade geral na amostra, obteve-se 12 fatores, no qual os escores fatoriais respectivos foram utilizados como banco de dados para a análise Discriminante.

Segundo Mingoti [1] a análise discriminante permite a formulação de algoritmos de classificação a novos elementos em populações conhecidas *a priori*. Dessa forma,

utilizou-se essa técnica com o intuito de classificar um candidato como aprovado ou reprovado baseado em suas características sócio-educacionais. Para avaliar a eficiência do modelo, utilizou-se duas técnicas. Na primeira definiu-se um grupo de treinamento composto por 70% dos dados, empregados na estimação dos parâmetros e o restante foi utilizado para validação do modelo. A segunda técnica, conhecida como procedimento de Lachenbruch, considera todas as observações na estimação dos parâmetros. Com essa técnica obteve-se resultados melhores, sendo de 34,04% a probabilidade de classificação errada de um indivíduo como aprovado, sendo na verdade reprovado e de 42,45%, a probabilidade de classificar um candidato aprovado como reprovado. Sendo assim, os resultados mostraram-se significativamente melhores na classificação de reprovados. Com a aplicação da análise discriminante, obteve-se resultados que se mostraram razoáveis na classificação dos candidatos, certamente porque o desempenho deste não se apresenta totalmente vinculado e explicado pelas variáveis consideradas.

Assim, a estatística multivariada atua como uma área de grande importância, seja pelo desenvolvimento dos métodos e softwares computacionais, seja pelo seu amplo meio de aplicação em diversas áreas do conhecimento, por atuar na descoberta de conhecimento em banco de dados. Quando aplicada, permite a estruturação numérica das variáveis envolvidas no fenômeno, possibilitando o levantamento e a predição de informações ou padrões referentes a ele. Quando relacionada à educação permite um estudo detalhado, explicitando as variáveis que melhor explicam o resultado final, por meio dos recursos quantitativos.

Com a utilização dessa ferramenta foi possível analisar a relação entre o desempenho e as variáveis sócio-educacionais dos candidatos ao vestibular da instituição. Mesmo sabendo que tais variáveis não tenham caráter informativo, ou definem o resultado de um candidato, foi possível estruturar o comportamento das mesmas, tornando visível o processo de análise e padronização quando se tem um banco de dados multivariados. Com o auxílio das técnicas de Análise Fatorial, Análise de Componentes Principais e Análise Discriminante foi realizada a redução do banco de dados, e a conseqüente formulação do algoritmo de discriminação. Os resultados obtidos foram suficientes para o desenvolvimento e formulação dos métodos, que visaram a padronização das variáveis associadas ao desempenho dos candidatos e a validação da aplicação do modelo multivariado.

## **Referências**

- [1] S. A. Mingoti, “Análise de Dados através de métodos de estatística multivariada/uma abordagem aplicada”, UFMG, Belo Horizonte, 2005.
- [2] A. Zanella, “Identificação de fatores que influenciam na qualidade do ensino de matemática, através da análise multivariada”, Dissertação de Mestrado, UFSM, 2006.

# **SLOTTING DE ARMAZÉNS UTILIZANDO ALGORITMO GENÉTICO**

Maickel Hubner

Universidade Federal do Paraná - UFPR, maickel.hubner@gmail.com

## **Resumo:**

Este trabalho procura desenvolver uma técnica matemática com o intuito de determinar a melhor localização de produtos na área de *picking* de um armazém, baseado em dados de sua previsão de vendas para um período vindouro e da formulação estatística dos pedidos em períodos passados. Como melhor localização entende-se aquela que, ao final da coleta de um número de pedidos, cada um contendo um conjunto variável de produtos, demandou a menor distância média por produto (ou menor tempo médio), percorrida pelos operadores.

Para atingir este objetivo é utilizado um algoritmo genético padrão, além de alguns operadores adicionais desenvolvidos especificamente para o problema, utilizando uma representação baseada em ordem. Como resultado é apresentado um comparativo da resolução de um problema fictício empregando técnicas habituais e o método proposto utilizando o programa específico construído para a resolução do problema.

**Palavras-chave:** algoritmo genético, *picking*, armazenagem

## **Apresentação:**

Com o aumento da competitividade ocasionado pelo atual movimento de globalização das economias, e com a conseqüente abertura dos mercados em nível internacional, as empresas mais do que nunca estão utilizando as mais variadas técnicas para tentar reduzir os seus custos.

Estudo recente do Banco Mundial estimou os custos logísticos brasileiros em 20% do PIB, em comparação com percentuais na faixa de 10 a 12% em países da OECD e os EUA. Observa-se que, no caso brasileiro, esses custos são excessivos tanto no transporte, como nos serviços de movimentação de carga em terminais,



armazenagem e serviços auxiliares, respondendo por um terço dos custos operacionais das empresas (GUASCH, 2000).

Existem três formas de reduzir os custos de armazenagem, sem alteração das capacidades produtivas das empresas, são elas:

- 1) Agilizar a entrada de produtos num armazém.
- 2) Agilizar a saída de produtos num armazém.
- 3) Organizar o armazém de tal forma que tanto entrada quanto saída sejam favorecidas.

Uma pergunta aparentemente simples e inofensiva tem atormentado por muito tempo os responsáveis pela armazenagem em todo o mundo. “Onde colocar cada produto no armazém”? Qualquer tentativa de resposta acaba recaindo em mais perguntas. Essa otimização é comercialmente conhecida por *slotting* de armazéns e é o tema principal desse trabalho.

Isso ocorre principalmente porque essa etapa do processo de armazenagem se encontra encurralada entre duas outras etapas cruciais, o recebimento e a expedição. Para a etapa do recebimento as perguntas referem-se principalmente a previsões. “Quantos caminhões irão chegar”? “Quanto espaço devo reservar para determinado produto”? A maioria dessas perguntas pode ser respondida a contento através de métodos estatísticos e possuem diversas soluções propostas.

Já na etapa da expedição as perguntas referem-se principalmente a seqüenciamento e organização. “Como coletar determinada lista de pedidos mais rápido”? “Como embarcar os pedidos no menor tempo”? Para essas perguntas também existem diversas soluções utilizando métodos de Pesquisa Operacional muito difundidos.

O problema da pergunta “Onde colocar cada produto no armazém?” é que ela depende das respostas dadas às perguntas do recebimento e da sua resposta dependem as respostas às perguntas da expedição.

Matematicamente, o problema do *slotting* de armazéns é da classe dos NP-Completo, ou seja, problemas cuja solução pode ser testada em um tempo hábil, mas a busca da melhor solução, analisando todas as soluções possíveis, é praticamente impossível. O número de soluções existentes é representado por  $n!$  onde  $n$  é o número de produtos a serem armazenados pela empresa.

Devido à natureza do problema, não é possível achar uma solução exata. Consegue-se apenas encontrar uma solução estatística, pois a solução exata só será encontrada após o conhecimento do conjunto de pedidos em que será aplicada.

Para a formulação do problema, usaremos as convenções a seguir. A letra  $n$  representará sempre o número de produtos diferentes que farão parte do problema. Esses produtos deverão ser distribuídos em  $p$  prateleiras. Os produtos deverão ocupar um número fixo de endereços na prateleira em que estiverem alocados (o vetor  $E$  guardará a quantidade de endereços  $e_j$  ocupados pelo produto  $j$ ) e cada prateleira tem um número fixo de endereços onde podem ser distribuídos produtos (o vetor  $T$  armazenará a quantidade total de endereços  $t_i$  disponíveis na prateleira  $i$ ). Já o vetor  $W$  armazena a distância  $w_i$  da prateleira  $i$  até a doca. Os dados dispostos até aqui são intrínsecos à distribuição física do problema (não mudam em diversas resoluções do problema) e portanto estão disponíveis antes de iniciarmos qualquer resolução.

Também possuímos duas estatísticas que podem ser retiradas através dos dados históricos do problema. Chamaremos de  $P_j$  a probabilidade de encontrarmos o produto  $j$  no conjunto de pedidos do problema (que será armazenada no vetor  $P$ ), e  $MP$  será a matriz de probabilidades, cujo elemento  $mp_{aj}$  mostra, caso o produto  $a$  seja encontrado em um pedido, qual a probabilidade condicionada de também encontrarmos o produto  $j$  no mesmo pedido. Essas duas estatísticas, apesar de mudarem a cada resolução do problema para períodos diferentes (dependem dos dados históricos), também podem ser coletadas antes de determinada resolução.

Precisaremos ainda de alguns dados que dependem diretamente do posicionamento dos produtos nas prateleiras. A matriz  $MD$  guardará a distância média  $md_{aj}$  entre os produtos  $a$  e  $j$ . O vetor  $D$  armazenará a distância média  $d_j$  do produto  $j$  à doca. Esses dados não estarão disponíveis ao iniciar a resolução e dependerão diretamente do posicionamento dos produtos nas prateleiras proposto pela solução.

A função que retorna a distância média percorrida por um operador ao coletar os pedidos de determinado período, por produto seria como a seguinte:

$$Z = \frac{\left\{ \sum_{i=1}^n e_i \cdot p_i \cdot \left( \frac{\left( \sum_{j=1}^n m p_{ij} \cdot m d_{ij} \right) + d_j}{\left( \sum_{j=1}^n m p_{ij} \right) + 1} \right) + \sum_{i=1}^n d_i \cdot e_i \cdot p_i \right\}}{\sum_{i=1}^n e_i}$$

O mínimo dessa função indicará a distribuição ótima de produtos no armazém proposto, respeitando as seguintes restrições:

- 1) *Restrição de alocação* - um produto só pode ser colocado em uma única prateleira

$$\sum_{i=1}^p x_{ij} = 1$$

$$(j = 1, 2, \dots, n)$$

- 2) *Restrição de espaço* - os endereços requeridos por todos os produtos alocados na prateleira não podem ser maior que o total de endereços disponíveis na prateleira

$$\sum_{j=1}^n x_{ij} \cdot e_j = t_i$$

$$(i = 1, 2, \dots, p)$$

Apesar de termos formulado todo o problema, sua função objetivo e suas restrições, observe-se que este não é um problema de otimização binária. Isso ocorre principalmente por causa da matriz  $MD$  e dos demais cálculos de distância, que em sua formulação dependem da solução proposta para o problema. Nesse caso optou-se pela utilização de um algoritmo genético para sua solução.

O algoritmo genético teria as seguintes características:

- Representação cromossomial de ordem;
- Método da roleta para seleção;
- Operadores de *crossover* e mutação baseados em ordem (percentual de mutação de 10%);
- Modulo de população elitista (percentual de sobrevivência de 40% de uma geração para a próxima), com tamanho de população fixo;
- Eliminação preferencial de indivíduos com genótipo repetido da população;

- Indivíduos inválidos, se existirem, são penalizados pela função de avaliação;
- Critérios de parada utilizados: número máximo de gerações, convergência genética e invariabilidade de fitness do melhor indivíduo da população;

Também foi desenvolvido um operador adicional, denominado Mutaç o Dirigida. Esse operador tem por objetivo evitar a exist ncia de indiv duos inv lidos, transformando-os em indiv duos v lidos atrav s de pequenas muta  es em seu c digo gen tico.

O m todo foi ent o aplicado em um problema fict cio, baseado em dados hist ricos de uma empresa real, para que sua efici ncia fosse comprovada. Com as t cnicas emp ricas que s o utilizadas pela empresa, cada operador percorria uma dist ncia m dia de 9,282 metros por produto a ser coletado. A empresa coleta em m dia cinco mil produtos diariamente, o que representa uma dist ncia total percorrida de 46,41 Km.

Utilizando o algoritmo gen tico proposto, foi encontrada uma nova distribui o de produtos. A mesma coleta real hist rica foi simulada contra essa distribui o e chegou a uma dist ncia m dia percorrida de 7,385 metros por produto coletado. Isso representa uma dist ncia total percorrida no dia de 36,925 Km, ou uma redu o de 9,485 Km percorridos (algo em torno de 20%).

Assim comprovou-se que a utiliza o de algoritmos gen ticos para defini o da melhor distribui o de produtos em um armaz m (*slotting*)   uma t cnica vi vel, al m de representar uma redu o de grande escala na dist ncia percorrida pelos operadores para realiza o do *picking* de produtos (independente do m todo de separa o utilizado), o que, conseq entemente, traduz-se na redu o do tempo necess rio para a separa o dos pedidos, aumentando com isso a competitividade da empresa perante o mercado e a sua lucratividade nos processos log sticos.

#### **Refer ncias Bibliogr ficas:**

GOLDBARG, Marco C. & LUNA, H. P. L.: **Otimiza o combinatorial e Programaa o linear**, Elsevier, Rio de Janeiro, Brasil, (2005).

GUASCH, L.: **Diagnostico de la logistica en America Latina: impacto en la Competividad y una reforma por hacer**, Presentation, The World Bank, Wash. D.C., (2000).

HOLLAND, John H.: **A Ordem Oculta – Como Adaptação gera a Complexidade**, Gradiva, Lisboa, Portugal, (1995).

LINDEN, Ricardo: **Algoritmos Genéticos – Uma importante ferramenta da Inteligência Computacional**, Brasport, Rio de Janeiro, (2006).

MULCAHY, David E.: **Warehouse Distribution & Operations Handbook**, McGraw-Hill INC, EUA, (1994).

TOMPKINS, James A. & SIMTH, Jerry D.: **The Warehouse Management Handbook**, second edition, (1998).