

Universidade Federal de Santa Catarina  
Programa de Pós-Graduação em Matemática

CÁLCULO NUMÉRICO DA EXPONENCIAL DE UMA MATRIZ

JOÃO ARTUR DE SOUZA

Florianópolis

1993

Universidade Federal de Santa Catarina  
Programa de Pós-Graduação em Matemática

CÁLCULO NUMÉRICO DA EXPONENCIAL DE UMA MATRIZ

JOÃO ARTUR DE SOUZA

Dissertação submetida a Universidade Federal de Santa  
Catarina para a obtenção do Grau de Mestre em Ciências.

Florianópolis

1993

# CÁLCULO NUMÉRICO DA EXPONENCIAL DE UMA MATRIZ

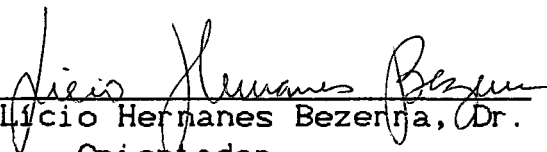
JOÃO ARTUR DE SOUZA

Esta dissertação foi julgada adequada para obtenção do  
Título de Mestre em Ciências.



Prof. Gerardo A. L. Luna, Dr.  
Coordenador do curso

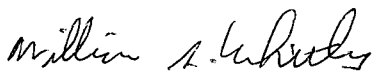
Banca examinadora



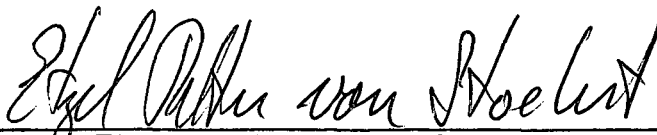
Prof. Lício Hernanes Bezerra, Dr.  
Orientador



Prof. Mário Cesar Zambalde, Dr.



Prof. William Glenn Whitley, Dr.



Prof. Etzel Ritter Von Stockert, Dr.

A Gertrudes

## AGRADECIMENTOS

Quero expressar aqui minha gratidão a todas as pessoas que direta ou indiretamente me ajudaram na elaboração deste trabalho, assim, como na minha formação acadêmica. Agradeço em especial aos professores:

Genaldo Leite Nunes, não só pelas discussões matemáticas, como também pela assessoria técnico computacional,

Nilo Kuelkamp, pelo incentivo constante e valiosa participação na minha formação matemática,

Rogério de Aguiar, pelas discussões matemáticas, e amizade constante,

William Glenn Whitney, pela assessoria matemática, tanto em discussões gerais como na revisão de algumas demonstrações.

O agradecimento maior reservo ao Professor Licio Hernanes Bezerra, por ter me indicado a problemática que deu origem ao presente trabalho e por tê-lo acompanhado com valiosas sugestões.

E finalmente agradeço à Gertrudes A. Dandolini pelo incentivo, companheirismo e pelo auxílio na digitação deste trabalho.

## SUMÁRIO

Resumo	
Abstract	
1 Introdução.....	1
2 Função de Matriz.....	4
2.1 Introdução.....	4
2.1 Definições.....	6
2.2.1 Sylvester e Buchheim.....	6
2.2.2 Weyr.....	10
2.2.3 Cartan.....	13
2.2.4 Fantappiè.....	14
2.2.5 Giorgi.....	15
2.2.6 Schwerdtfeger.....	17
2.2.7 Cipolla.....	18
2.2.8 Richter.....	19
2.3 A equivalência das definições.....	20
3 Função Exponencial de uma Matriz.....	30
3.1 Introdução.....	30
3.2 A Sensibilidade do Problema.....	30
3.3 Limitações de $e^{At}$ .....	32
3.4 A condição da exponencial de uma matriz.....	44
4 Métodos de Aproximação.....	52
4.1 Aproximações pela série de Taylor.....	52
4.2 Aproximações de Padé.....	56
4.3 Aproximações de Padé para a Função Exponencial.....	64
5 Testes Numéricos.....	75
5.1 Introdução.....	75
5.2 Aproximações por Taylor.....	75
5.3 Taylor com balanceamento.....	79
5.4 Padé.....	83
5.5 Padé com balanceamento.....	84

5.6 Runge Kutta.....	87
5.7 Comentários.....	88
6 Conclusão.....	91
7 Referências Bibliográficas.....	92
Apêndice.....	94

## RESUMO

A importância em resolver a Equação Diferencial

$$\begin{cases} x'(t) = Ax(t) \\ x(0) = x_0 \end{cases},$$

onde  $A$  é uma matriz quadrada, levou-nos ao estudo da exponencial de uma matriz  $e^{At}$ . Nosso objetivo é o cálculo numérico de  $e^A$ . Para tanto estudamos como se define função de uma matriz associada com uma função  $f(z)$ , com  $z \in \mathbb{C}$ . Assim, primeiro exploramos teoricamente os problemas em calcular a exponencial de uma matriz e em seguida utilizamos as aproximações de Taylor e Padé para fazer o cálculo numérico de  $e^A$ .

Concluimos, que a magnitude da norma de uma matriz  $A$  influencia nas aproximações do cálculo de  $e^A$ , fato que se tenta contornar utilizando-se a igualdade  $e^A = (e^{A/m})^m$ .



## ABSTRACT

The necessity to resolve the differential equation

$$\begin{cases} x'(t) = Ax(t) \\ x(0) = x_0 \end{cases} ,$$

leads us to study the exponential  $e^{At}$  of a matrix. Our objective is to calculate  $e^A$  numerically. For this end, we study how one defines the matrix function associated with a function  $f(z)$ ,  $z \in \mathbb{C}$ . We begin studying the theoretic aspects necessary to calculate the exponential of a matrix and continue our study using Taylor and Padé approximations to calculate  $e^A$  numerically.

We conclude that the magnitude of the norm of a matrix  $A$  influences the approximations of  $e^A$  and we try to avoid this problem using the equation  $e^A = (e^{A/m})^m$ .

## 1 - INTRODUÇÃO

Um dos problemas importantes em Equações Diferenciais é a resolução do problema de valor inicial

$$\begin{cases} x'(t) = Ax(t) \\ x(0) = x_0 \end{cases}, \quad (1.1)$$

onde  $A$  é uma matriz, cuja solução é dada em termos da função de matriz  $e^{At}$ . Assim torna-se necessário definir função de uma matriz  $f(A)$ , associada a uma função  $f(z)$ , onde  $z \in \mathbb{C}$  (capítulo 2).

Considere, por exemplo, um tubo longo contendo uma determinada substância em difusão no meio principal que preenche o tubo. Este processo se verifica em uma região delimitada entre as secções  $S_0$  e  $S_n$ , figura 1. Se as secções  $S_k$  estão

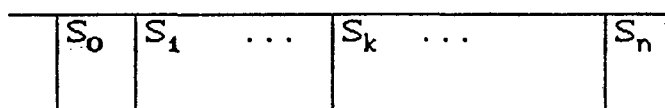


Figura 1- Tubo cheio de liquido

pouco separadas, considere  $\mu_k$  a concentração em média da substância entre  $S_{k-1}$  e  $S_k$ . Segundo a Lei de Fick, a taxa de difusão através de uma destas secções  $S_k$  é proporcional à diferença das concentrações médias destes segmentos adjacentes. O processo é considerado contínuo no tempo e discreto no espaço, o que se costuma chamar de semidiscretização do problema. Desta forma, temos:

$$\frac{d\mu_k}{dt} = \alpha(\mu_{k+1} - \mu_k) + \alpha(\mu_{k-1} - \mu_k) \quad (1.2)$$

$$\frac{1}{\alpha} \frac{d\mu_k}{dt} = \mu_{k+1} - 2\mu_k + \mu_{k-1} = \Delta^2 \mu_{k-1} .$$

Na forma vetorial, temos:

$$\frac{d\mu}{dt} = A\mu , \quad (1.3)$$

onde

$$A = \alpha \begin{pmatrix} -2 & 1 & \dots & 0 \\ 1 & \dots & \dots & \vdots \\ \vdots & \dots & \dots & 1 \\ 0 & \dots & 1 & -2 \end{pmatrix} . \quad (1.4)$$

A solução do sistema é dada por  $U(t) = e^{At}$ . Para  $\alpha = 1$  e um sistema de ordem 3, obtemos:

$$e^{At} = P e^{Dt} P^{-1}, \quad (1.5)$$

onde

$$P = P^{-1} = \frac{1}{2} \begin{pmatrix} 1 & -\sqrt{2} & 1 \\ -\sqrt{2} & 0 & \sqrt{2} \\ 1 & \sqrt{2} & 1 \end{pmatrix}, \quad (1.6)$$

e

$$e^{Dt} = \begin{pmatrix} e^{(-2-\sqrt{2})t} & 0 & 0 \\ 0 & e^{-2t} & 0 \\ 0 & 0 & e^{(-2+\sqrt{2})t} \end{pmatrix} \quad (1.7)$$

Desde 1883 surgiram várias definições de função de uma matriz (capítulo 2). Rinehart [6] mostra que, a menos de generalidades, as definições de função de uma matriz são equivalentes e preservam algumas propriedades da função escalar à qual a função de matriz está associada.

Van Loan [5], Moler e Van Loan [1] realizam uma análise da sensibilidade do problema (capítulo 3) de como definir a função exponencial de uma matriz, estimando algumas limitações da perturbação relativa,  $\varphi(t)$

$$\varphi(t) = \frac{\| e^{(A+E)t} - e^{At} \|}{\| e^{At} \|}, \text{ onde } t \geq 0. \quad (1.8)$$

Os autores concluem que o problema da exponencial de uma matriz é bem condicionado quando a matriz é normal.

Neste trabalho, fizemos testes numéricos (capítulo 5) para comparar as aproximações por Taylor, Runge Kutta de quarta ordem e Padé [1] para a exponencial de matriz.

## 2- FUNÇÃO DE MATRIZ

## 2.1- Introdução

Nosso objetivo é a definição de  $f(A)$ , onde  $A$  é uma matriz em  $\mathbb{C}^{n \times n}$  e  $f$  uma função, ou seja, se tivermos uma função  $f(z)$ ,  $z \in \mathbb{C}$ , queremos definir  $f(A)$ ,  $A \in \mathbb{C}^{n \times n}$ .

Formalmente, poderíamos definir  $f(A)$ , simplesmente substituindo  $z$  por  $A$  em  $f(z)$ . Por exemplo:

$$1) \text{ Se } f(z) = \frac{1 + z/2}{1 - z/2} \quad \text{para } z \neq 2, \text{ então}$$

$$f(A) = (I - A/2)^{-1}(I + A/2) \quad \text{se } 2 \notin \lambda(A)^*$$

$$2) \text{ Se } f(z) = e^z = \sum_{k=0}^{\infty} z^k/k!, \text{ então}$$

$$f(A) = e^A = \sum_{k=0}^{\infty} A^k/k!.$$

$$3) \text{ Polinômios. Se } p(z) = \sum_{i=0}^n c_i z^i, \text{ então}$$

$$p(A) = \sum_{i=0}^n c_i A^i.$$

Observemos que no segundo exemplo temos uma série, o que já envolve convergência.

Uma questão interessante é saber se as propriedades da função escalar são preservadas.

L. Fantappiè foi o primeiro escritor a estabelecer

(\*)  $\lambda(A)$  é o espectro de  $A$ .

explicitamente um conjunto de exigências combinadas para serem cumpridas por uma definição de função de matriz. São as seguintes:

- (i) se  $f(z) = k$ , então  $f(A) = k I$  ;
- (ii) se  $f(z) = z$ , então  $f(A) = A$ ;
- (iii) se  $f(z) = g(z) + h(z)$ , então  $f(A) = g(A) + h(A)$ ;
- (iv) se  $f(z) = g(z).h(z)$ , então  $f(A) = g(A).h(A)$ ,

onde  $A$  é uma matriz admissível para cada função e  $f(A)$  denota a função de matriz que provém da função escalar  $f(z)$ . Estas exigências implicarão que a definição, quando aplicada para um polinômio  $p(z)$ , produza o usual polinômio de matriz  $p(A)$ ; e que qualquer identidade algébrica sobre funções escalares de uma variável complexa valerá para a função de matriz correspondente.

Exemplo: Se  $\text{sen}(A)$  e  $\text{cos}(A)$  podem ser definidos segundo o que foi dito acima, então

$$(\text{sen}(A))^2 + (\text{cos}(A))^2 = I.$$

Uma quinta exigência altamente desejável seria: se  $f(z) = h(g(z))$ , então  $f(A) = h(g(A))$ , para todo  $A$  admissível.

Estas exigências, se satisfeitas, assegurariam que cada identidade funcional entre funções escalares seria preservada na sua parte correspondente na função de matriz.

Exemplo:  $e^{\ln z} = z \quad \Rightarrow \quad e^{\ln A} = A.$

Muitos matemáticos desde 1883 têm se preocupado com o problema de definir uma função de matriz associada à função escalar. Como definições anteriores eram desconhecidas,

devido à falta de registros, cada um fez a sua própria definição, não se preocupando com as equivalências e semelhanças entre as definições.

Como resultado disto, desde 1880, oito definições distintas de uma função de matriz foram propostas [6]: Sylvester e Buchheim, Weyr, Cartan, Fantappiè, Giorgi, Schwerdtfeger, Cipolla, Richter.

Mostraremos que as definições são essencialmente equivalentes, a menos de generalidades, e que as exigências (I) - (IV) de Fontappiè são cumpridas pela definição mais geral (Cippola) e (V) é cumprida se  $f(z)$  e  $g(z)$  são univalentes.

## 2.2- Definições

### 2.2.1- Sylvester e Buchheim

Sylvester em 1883 definiu uma função de matriz correspondente a uma função escalar  $f(z)$  do seguinte modo:

$$f(A) = \sum_{j=1}^n \prod_{i \neq j} \frac{A - \lambda_i I}{\lambda_j - \lambda_i} f(\lambda_j), \quad (2.1)$$

onde  $A$  é uma matriz com raízes características distintas  $\lambda_1, \lambda_2, \dots, \lambda_n$ . Note que esta definição é uma extensão direta para matrizes da fórmula de interpolação de Lagrange para um polinômio  $p(z)$  de grau  $(n-1)$  que passa por  $n$  pontos distintos.

Exemplo:

$$\text{Se } A = \begin{pmatrix} 1 & 6 \\ 1 & 2 \end{pmatrix} \text{ e } f(z) = e^z, \text{ temos } \lambda_1 = 4 \text{ e } \lambda_2 = -1,$$

assim

$$f(A) = ((A+I)/(4+1)) \cdot f(4) + ((A-4I)/(-1-4)) \cdot f(-1) =$$

$$= \frac{A+I}{5} e^4 + \frac{A-4I}{-5} e^{-1} =$$

$$= \begin{pmatrix} 2/5 & 6/5 \\ 1/5 & 3/5 \end{pmatrix} e^4 + \begin{pmatrix} 3/5 & -6/5 \\ -1/5 & 2/5 \end{pmatrix} e^{-1} =$$

$$= \begin{pmatrix} (2/5)e^4 + (3/5)e^{-1} & (6/5)e^4 - (6/5)e^{-1} \\ (1/5)e^4 - (1/5)e^{-1} & (3/5)e^4 + (2/5)e^{-1} \end{pmatrix}.$$

Observamos que se  $A$  é diagonalizável, ou seja,  $A = PDP^{-1}$ , onde  $D$  é uma matriz diagonal, então  $f(A) = Pf(D)P^{-1}$ .

Buchheim em 1886 generalizou esta definição para o caso onde as raízes características de  $A$  não são necessariamente distintas. Seja  $p(z)$  um polinômio. O polinômio de grau menor ou igual a  $n-1$  que passa pelos pontos  $(a_i, p(a_i))$  com derivadas  $p'(a_i), \dots, p^{(s_i-1)}(a_i)$  e

$$n = \sum_{i=1}^t s_i \quad (2.2)$$

é definido unicamente por:

$$g_p(z) = \sum_{k=1}^t \prod_{i \neq k} (z - a_i)^{s_i} \sum_{j=0}^{s_k-1} (1/j!) g_k^{(j)}(a_k) (z - a_k)^j, \quad (2.3)$$

onde

$$g_k(z) = \frac{p(z)}{\prod_{i \neq k} (z - a_i)^{s_i}} \quad (2.4)$$



Se o grau de  $p(z)$  é menor ou igual a  $n-1$  temos  $p(z) = g_p(z)$ . Portanto  $p(A) = g_p(A)$ , se  $A$  é uma matriz  $n \times n$ .

Se o grau de  $p(z)$  é maior ou igual a  $n$  e  $A$  é uma matriz quadrada  $n \times n$  onde para qualquer  $i$ ,  $a_i$  é autovalor de  $A$  com multiplicidade  $s_i$ , temos também  $p(A) = g_p(A)$ . De fato, seja  $c(z) = \det(A - zI)$ , o polinômio característico de  $A$ . Logo  $p(z) = h(z)c(z) + g_p(z)$ , onde  $g_p(z)$  é zero ou de grau menor que  $n$ . Agora o polinômio  $p(z) - g_p(z) = h(z)c(z)$  e também suas primeiras  $(s_k - 1)$  derivadas tem valor zero para  $z = \lambda_k$ ,  $k = 1, 2, \dots, t$ , o que caracteriza a igualdade  $p(A) = g_p(A)$ , ou seja,

$$\begin{aligned} & \sum_{j=1}^t \prod_{i \neq j} (A - \lambda_i I)^{s_i} \sum_{k=0}^{s_j-1} \frac{1}{k!} \frac{d^k}{dz^k} \left[ \frac{g_p(z)}{\prod_{h \neq j} (z - \lambda_h)^{s_h}} \right]_{z=\lambda_j} = \\ & = \sum_{j=1}^t \prod_{i \neq j} (A - \lambda_i I)^{s_i} \sum_{k=0}^{s_j-1} \frac{1}{k!} \frac{d^k}{dz^k} \left[ \frac{p(z)}{\prod_{h \neq j} (z - \lambda_h)^{s_h}} \right]_{z=\lambda_j}. \quad (2.5) \end{aligned}$$

Se  $s_i$  é a multiplicidade de  $a_i$ , como raiz do polinômio minimal, então a igualdade acima ainda vale.

Agora, o termo do lado direito em (2.5) tem sentido se substituirmos  $p(z)$  por uma função analítica para os zeros repetidos do polinômio minimal de  $A$  e apenas definida para os zeros não repetidos. A definição de Buchheim é

$$f(A) = \sum_{j=1}^t \prod_{i \neq j} (A - \lambda_i I)^{s_i} \sum_{k=0}^{s_j-1} \frac{1}{k!} \frac{d^k}{dz^k} \left[ \frac{f(z)}{\prod_{h \neq j} (z - \lambda_h)^{s_h}} \right]_{z=\lambda_j} \quad (2.6)$$

onde  $s_i$  é a multiplicidade de  $\lambda_i$ , que é uma raiz do polinômio minimal de A.

Exemplo: 1) Seja

$$A = \begin{pmatrix} 5 & -6 & -6 \\ -1 & 4 & 2 \\ 5 & -6 & -4 \end{pmatrix} \quad \text{e } f(z) = e^z.$$

Temos  $\lambda_1=1$  com  $s_1=1$  e  $\lambda_2=2$  com  $s_2=2$ .

Assim,

$$f(A) =$$

$$= \prod_{i \neq 1} (A - \lambda_i I)^{s_i} \sum_{k=0}^0 \frac{1}{k!} \frac{d^k}{dz^k} \left[ \frac{e^z}{\prod_{h \neq 1} (z - \lambda_h)^{s_h}} \right]_{z=\lambda_1} (A - \lambda_1 I)^k +$$

$$+ \prod_{i \neq 2} (A - \lambda_i I)^{s_i} \sum_{k=0}^1 \frac{1}{k!} \frac{d^k}{dz^k} \left[ \frac{e^z}{\prod_{h \neq 2} (z - \lambda_h)^{s_h}} \right]_{z=\lambda_2} (A - \lambda_2 I)^k =$$

$$= (A - \lambda_2 I)^2 \frac{f(\lambda_1)}{(\lambda_1 - \lambda_2)^2} + (A - \lambda_1 I) \left( \frac{f(\lambda_2)}{(\lambda_2 - \lambda_1)} +$$

$$+ \frac{(f'(\lambda_2)(\lambda_2 - \lambda_1) - f(\lambda_2)) (A - \lambda_2 I)}{(\lambda_2 - \lambda_1)^2} \right) =$$

$$= (A - 2I)^2 \frac{e^1}{(1-2)^2} + (A - I) \left( \frac{e^2}{(2-1)} + (e^2(2-1) - e^2)(A - 2I) \right) =$$

$$= \begin{pmatrix} -3e^1 + 4e^2 & 6e^1 - 6e^2 & 6e^1 - 6e^2 \\ e^1 - e^2 & -2e^1 + 3e^2 & -2e^1 + 2e^2 \\ -3e^1 + 3e^2 & 6e^1 - 6e^2 & 6e^1 - 5e^2 \end{pmatrix}.$$

2) Seja

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -2 & 4 \end{pmatrix} \text{ e } f(z) = e^z.$$

Temos  $\lambda_1=3$  com  $s_1=1$  e  $\lambda_2=2$  com  $s_2=2$ .

Assim,

$$\begin{aligned} f(A) &= (A-2I)^2 e^3 + (A-3I) \begin{bmatrix} -e^2 I & -e^2(A-2I) \end{bmatrix} = \\ &= \begin{pmatrix} e^2 & e^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 2e^2 - e^3 & e^3 - e^2 \\ 0 & 0 & 2e^2 - 2e^3 & 2e^3 - e^2 \end{pmatrix}. \end{aligned}$$

### 2.2.2- Weyr

A definição de uma função de matriz como uma série de potências deve ter ocorrido para muitos matemáticos, mas parece que E. Weyr em 1887 foi o primeiro a publicar um artigo sobre esse assunto. A definição da série de potências é uma extensão natural da função polinomial de uma matriz.

Seja  $f(z)$  analítica para  $z = z_0$ , e

$$f(z) = f(z_0) + f'(z_0)(z-z_0) + \dots + \frac{f^{(k)}(z_0)}{k!} (z-z_0)^k + \dots \quad (2.7)$$

então podemos definir a função de matriz correspondente,

$$f(A) = f(z_0)I + f'(z_0)(A-z_0I) + \dots + \frac{f^{(k)}(z_0)}{k!} (A-z_0I)^k + \dots \quad (2.8)$$

desde que a série convirja, e a série converge se, e somente se, para cada autovalor  $\lambda$  de  $A$ ,  $(\lambda - z_0)$  está no interior ou so-

bre o círculo de convergência de  $f(z)$  (Teorema de Hensel<sup>1</sup>).

A condição de que todas as raízes características permaneçam no interior ou sobre o círculo de convergência de  $f(z)$  é claramente mais forte que o pedido na definição de Buchheim (analiticidade para os zeros repetidos do polinômio minimal e apenas definida para os zeros não repetidos).

Para verificar esta dificuldade Ferrar [10] aplicou o seguinte esquema: Seja  $t \in \mathbb{C}$  tal que as raízes características  $t\lambda_i - tz_0$  de  $tA - tz_0I$  são todas menores em valor absoluto que o raio de convergência de  $f(z)$ ,

$$f(z) = f(z_0) + f'(z_0)(z-z_0) + \dots + \frac{f^{(k)}(z_0)}{k!} (z-z_0)^k + \dots \quad (2.9)$$

Então

$$f(tA) = f(z_0)I + f'(z_0)t(A-z_0I) + \dots + \frac{f^{(k)}(z_0)}{k!} t^k(A-z_0I)^k + \dots \quad (2.10)$$

converge para um polinômio em  $t(A-z_0I)$ .

Seja  $A = PJP^{-1}$  a decomposição de Jordan de  $A$ , assim

$$\begin{aligned} f(tA) &= f(z_0)I + f'(z_0)t(PJP^{-1} - Pz_0IP^{-1}) + \dots + \\ &+ \dots + \frac{f^{(k)}(z_0)}{k!} t^k (PJP^{-1} - Pz_0IP^{-1})^k + \dots = \\ &= Pf(z_0)IP^{-1} + f'(z_0)tP(J - z_0I)P^{-1} + \dots + \\ &+ \dots + \frac{f^{(k)}(z_0)}{k!} t^k [P(J - z_0I)P^{-1}]^k + \dots = \end{aligned}$$

(1) Ver apêndice.

$$\begin{aligned}
&= Pf(z_0)IP^{-1} + Pf'(z_0)t(J - z_0I)P^{-1} + \dots + \\
&\quad + \dots + P \frac{f^{(k)}(z_0)}{k!} t^k (J - z_0I)^k P^{-1} + \dots = \\
&= P[f(z_0)I + f'(z_0)t(J - z_0I) + \dots]P^{-1}. \quad (2.11)
\end{aligned}$$

Agora  $J = \text{diag}(J_1, J_2, \dots, J_r)$ , e para cada  $J_i$  tomando  $z_0 = \lambda_i$  temos<sup>1</sup>

$$\begin{aligned}
f(tJ_i) &= f(\lambda_i)I + f'(\lambda_i)t(J_i - \lambda_i I) + \dots + \\
&\quad + \dots + \frac{f^{(v_i-1)}(\lambda_i)}{(v_i-1)!} t^{v_i-1}(J_i - \lambda_i I)^{(v_i-1)}, \quad (2.12)
\end{aligned}$$

onde  $v_i$  é a multiplicidade de  $\lambda_i$ . Assim  $f(tA) = P_A(tA)$ , onde  $P_A$  é um polinômio. Como  $P_A(tA)$  está definido para  $t=1$  temos  $f(A) = P_A(A)$  por continuidade analítica.

Exemplo: Seja  $A = \begin{pmatrix} 1 & 6 \\ 1 & 2 \end{pmatrix}$  e  $f(z) = e^z$ .

Temos  $A = \begin{pmatrix} 2/5 & 3/5 \\ 1/5 & -1/5 \end{pmatrix} \begin{pmatrix} 4 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & -2 \end{pmatrix}$  e assim

$$\begin{aligned}
f(A) = e^A &= \begin{pmatrix} 2/5 & 3/5 \\ 1/5 & -1/5 \end{pmatrix} \begin{pmatrix} e^4 & 0 \\ 0 & e^{-1} \end{pmatrix} \begin{pmatrix} 1 & 3 \\ 1 & -2 \end{pmatrix} = \\
&= \begin{pmatrix} (2/5)e^4 + (3/5)e^{-1} & (6/5)e^4 - (6/5)e^{-1} \\ (1/5)e^4 - (1/5)e^{-1} & (3/5)e^4 + (2/5)e^{-1} \end{pmatrix}
\end{aligned}$$

(1)  $(J_i - \lambda_i I)$  é nilpotente de ordem  $v_i$ .

## 2.2.3- Cartan

E. Cartan aproximadamente em 1928, propôs em uma carta para G. Giorgi que a função de matriz  $f(A)$ , correspondente a uma função escalar  $f(z)$  analítica nas raízes características  $\lambda_1, \lambda_2, \dots, \lambda_N$  de  $A$ , seria definida por analogia com o teorema da integral de Cauchy como

$$f(A) = \frac{1}{2\pi i} \int \frac{f(z)}{zI-A} dz, \quad (2.13)$$

onde a integral é tomada para cada elemento da matriz  $f(z)(zI-A)^{-1}$  em torno de um caminho fechado contendo no seu interior cada uma das raízes características de  $A$ .

Cartan não elaborou sua definição além disto. Giorgi não fez observações relativas a qualquer semelhança desta definição com a sua própria. A definição de Cartan é mais restrita que a de Buchheim, pois  $f(z)$  deve ser analítica para cada raiz característica de  $A$  e para Buchheim basta ser analítica para as raízes repetidas do polinômio minimal, e apenas definida para as não repetidas.

Exemplo: Seja  $A = \begin{pmatrix} 1 & 6 \\ 1 & 2 \end{pmatrix}$  e  $f(z) = e^z$ .

$$\text{Assim } f(A) = \frac{1}{2\pi i} \int \frac{e^z dz}{zI-A} = \frac{1}{2\pi i} \int e^z (zI-A)^{-1} dz,$$

$$\text{mas } (zI-A)^{-1} = \begin{pmatrix} \frac{z-2}{(z-4)(z+1)} & \frac{6}{(z-4)(z+1)} \\ \frac{1}{(z-4)(z+1)} & \frac{z-1}{(z-4)(z+1)} \end{pmatrix}$$

$$e \quad f(A) = \begin{pmatrix} \frac{1}{2\pi i} \int \frac{e^z(z-2)}{(z-4)(z+1)} dz & \frac{1}{2\pi i} \int \frac{6e^z}{(z-4)(z+1)} dz \\ \frac{1}{2\pi i} \int \frac{e^z}{(z-4)(z+1)} dz & \frac{1}{2\pi i} \int \frac{e^z(z-1)}{(z-4)(z+1)} dz \end{pmatrix}$$

Agora

$$\begin{aligned} \frac{1}{2\pi i} \int \frac{e^z(z-2)}{(z-4)(z+1)} dz &= \frac{1}{2\pi i} \frac{1}{5} \int \frac{e^z(z-2)}{z-4} dz - \\ &\quad - \frac{1}{2\pi i} \frac{1}{5} \int \frac{e^z(z-2)}{z+1} dz = \\ &= (2/5)e^4 + (3/5)e^{-1}. \end{aligned}$$

$$\begin{aligned} \frac{1}{2\pi i} \int \frac{6e^z}{(z-4)(z+1)} dz &= \frac{6}{2\pi i} \frac{1}{5} \int \frac{e^z}{z-4} dz - \\ &\quad - \frac{6}{2\pi i} \frac{1}{5} \int \frac{e^z}{z+1} dz = \\ &= (6/5)e^4 - (6/5)e^{-1}. \end{aligned}$$

$$\begin{aligned} \frac{1}{2\pi i} \int \frac{e^z(z-1)}{(z-4)(z+1)} dz &= \frac{1}{2\pi i} \frac{1}{5} \int \frac{e^z(z-1)}{z-4} dz - \\ &= \frac{1}{2\pi i} \frac{1}{5} \int \frac{e^z(z-1)}{z+1} dz \\ &= (3/5)e^4 + (2/5)e^{-1}. \end{aligned}$$

Logo

$$f(A) = \begin{pmatrix} (2/5)e^4 + (3/5)e^{-1} & (6/5)e^4 - (6/5)e^{-1} \\ (1/5)e^4 - (1/5)e^{-1} & (3/5)e^4 + (2/5)e^{-1} \end{pmatrix}$$

#### 2.2.4- Fantappiè

L. Fantappiè em 1928 impôs as condições I-IV e ainda a

condição V': se  $f(z)$  é analítica em um parâmetro  $t$  então os elementos de  $f(A, t)$  dependem analiticamente de  $t$ . Ele aplicou a teoria de funcionais lineares para deduzir que os elementos  $f_{rs}$  de  $f(A)$  devem ser dados pela soma dos resíduos de  $-(D_{gr}(t)/D(t))f(t)$  e suas singularidades, onde  $D_{gr}(t)$  é o cofator do elemento  $(r, s)$  de  $A-tI$  e  $D(t) = \det(A-tI)$ . Observemos que esta definição é equivalente a que sugeriu Cartan (apenas uma forma de resolver a integral).

### 2.2.5- Giorgi

G. Giorgi em 1928 propôs uma definição de matriz motivado por uma propriedade das funções polinomiais  $p(z)$ ; se  $P$  é uma matriz não-singular então

$$P^{-1}p(A)P = p(P^{-1}AP) \quad \text{ou} \quad p(A) = P p(P^{-1}AP) P^{-1}. \quad (2.14)$$

Assim, se  $X^{-1}AX = J$  é a forma de Jordan de  $A$  então

$$p(A) = X p(J) X^{-1}. \quad (2.15)$$

Agora, se esta propriedade é usada para definir uma função de matriz não-algébrica então o problema de definir  $f(A)$  recai sobre  $f(J)$ , onde  $J = X^{-1}AX$  (Forma de Jordan.).

Notemos que  $J$  é a soma direta dos blocos  $J_1, \dots, J_k$ , ou seja,

$$J = \text{diag}(J_1, J_2, \dots, J_k), \quad (2.16)$$

onde



$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \dots & 0 \\ 0 & \lambda_i & 1 & \dots & 0 \\ & & & \ddots & \\ & & & & 1 \\ 0 & & & & \lambda_i \end{pmatrix} \text{ de ordem } s_i, \quad (2.17)$$

e  $p(J) = p(J_1) + p(J_2) + \dots + p(J_k)$  para qualquer função polinomial  $p(z)$ . Usando a fórmula de Taylor em torno de  $\lambda_i$ , determinamos

$$p(J_i) = p(\lambda_i)I + p'(\lambda_i)(J_i - \lambda_i I) + \dots + \frac{p^{(s_i-1)}(\lambda_i)}{(s_i-1)!} (J_i - \lambda_i I)^{s_i-1}, \quad (2.18)$$

ou seja,

$$p(J_i) = \begin{pmatrix} p(\lambda_i) & p'(\lambda_i) & \dots & \frac{p^{(s_i-1)}(\lambda_i)}{(s_i-1)!} \\ \cdot & p(\lambda_i) & p'(\lambda_i) & \dots & \frac{p^{(s_i-2)}(\lambda_i)}{(s_i-2)!} \\ \cdot & & \cdot & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & & \cdot \\ 0 & & & & p(\lambda_i) \end{pmatrix} \quad (2.19)$$

Observemos que  $(J_i - \lambda_i I)^k = 0$  para  $k \geq s_i$ .

É natural usar a definição também para uma função não-algébrica  $f(z)$ , desde que seja analítica para os zeros repetidos do polinômio minimal e apenas definida nos zeros não repetidos (caracterizando a equivalência com a definição de Buchheim), assim

$$f(A) = P (f(J_1) + \dots + f(J_k)) P^{-1}, \quad (2.20)$$

onde  $f(J_i)$  é definida como em (2.19) trocando-se  $p$  por  $f$ .

Notemos que, como  $f(A)$  é calculada fazendo-se uso somente dos blocos de Jordan, é independente de  $P$ . Assim para cada matriz  $A$  e cada função  $f(z)$  univalente, produz-se um único valor  $f(A)$ .

### 2.2.6- *Schwerdtfeger*

H. Schwerdtfeger em 1935 redescobriu a definição de Giorgi para o caso de raízes distintas e mostrou que era equivalente a definição em séries infinitas no caso de  $f(z)$  ser inteira.

Tendo como consideração as séries de potências ele definiu a função de matriz para o caso onde  $f(z)$  é analítica nos zeros repetidos do polinômio minimal  $M(z)$ , fazendo a separação em frações parciais

$$\frac{1}{M(z)} = \sum_{k=1}^l \frac{h_k(z)}{(z-\lambda_k)^{s_k}}, \quad (2.21)$$

onde  $\lambda_1, \dots, \lambda_l$  são os zeros distintos de  $M(z)$ :

$$f(A) = \sum_{k=1}^l G_k \sum_{i=0}^{s_k-1} f^{(i)}(\lambda_k) \frac{(A-\lambda_k I)^i}{i!}, \quad (2.22)$$

onde

$$G_k = g_k(A) \text{ (covariância de Frobenius de } A) \text{ e } g_k(z) = \frac{h_k(z) M(z)}{(z-\lambda_k)^{s_k}}.$$

Esta definição é equivalente à definição de Buchheim, pois  $f(z)$  é requerida ser definida para raízes características

de  $A$  é analítica para as raízes repetidas do polinômio minimal. Schwerdtfeger não relacionou sua definição com qualquer outra anterior, mas mostrou que a relação de Cartan (2.13) é satisfeita por (2.22) e que no caso de raízes características distintas (2.22) é equivalente a fórmula de interpolação de Lagrange para matrizes.

### 2.2.7- Cipolla

M. Cipolla em 1932 reescreveu a definição de Giorgi utilizando uma função multivalente  $v(z)$  ao invés de uma univalente, produzindo um valor adicional para  $v(A)$  além daquele obtido pela definição de Giorgi.

Primeiramente decompos a matriz  $A$  na forma de Jordan

$$J = P^{-1}AP = J_1 + J_2 + \dots + J_k. \quad (2.23)$$

Um valor para  $v(J)$  é definido por

$$v(J) = v_1(J_1) + v_2(J_2) + \dots + v_k(J_k), \quad (2.24)$$

onde  $v_i(z)$  é um valor de  $v(z)$ . Os vários valores possíveis de  $v(A)$  são então dados por  $v(A) = Pv(J)P^{-1}$  onde todos os valores de  $v(J)$  e todas as matrizes  $P$ 's tais que  $P^{-1}AP = J$  são usados.

Se o mesmo valor de  $v(z)$  é usado para todos os blocos correspondentes a uma mesma raiz  $\lambda_i$ , então esta definição coincide com a de Giorgi. Assim para esta função  $v(z)$ , a definição de Giorgi e Cipolla produzem o mesmo valor e, como este valor é expresso por um polinômio em  $A$ , independe da escolha de  $P$ .

Agora, se diferentes valores de  $v(z)$  forem usadas para dois ou mais blocos  $J_i$  correspondentes a uma mesma raiz carac-

terística então os valores de  $v(J)$  obtidos não são os mesmos da definição de Giorgi e neste caso  $v(J)$  não é necessariamente um polinômio em  $J$ . Assim  $v(A)$  não é necessariamente um polinômio em  $A$ . Portanto o valor  $v(A)$  obtido de  $v(J)$  não será independente da matriz  $P$ , o que proporciona uma adicional proliferação do valor de  $v(A)$ .

Observe que o mesmo raciocínio poderia ser feito para as definições de Buchheim, Schwerdtfeger, Fantappiè, Cartan e Richter.

### 2.2.8- Richter

Richter em 1951 definiu a função de matriz  $F_n$  cujas principais propriedades eram: os elementos de  $F_n(A)$  são funções contínuas dos elementos de  $A$ ; e  $F_n(P^{-1}AP) = P^{-1}F_n(A)P$ .

Para definir uma função de matriz correspondente a uma função escalar  $f(z)$  ele usou uma aproximação por série de potências para um bloco de Jordan e assim

$$f(A) = \sum_{i=0}^l R_i \sum_{k=0}^{s_i-1} f^{(k)}(\lambda_i) \frac{(A-\lambda_i I)^k}{k!}, \quad (2.25)$$

$i=0$

onde

$$R_i = w_i(A) \prod_{i \neq l} (A-\lambda_l I)^{s_l}, \quad (2.26)$$

e os polinômios  $w_i(z)$  são polinômios satisfazendo

$$w_i(z) \prod_{i \neq l} (z-\lambda_l)^{s_l} \equiv 1 \pmod{(z-\lambda_i)^{s_i}}, \quad (2.27)$$

onde  $s_i$  é a multiplicidade de  $\lambda_i$  no polinômio minimal de  $A$ . A matriz  $f(A)$  obtida não depende da escolha de  $w_i(z)$ .

### 2.3- A equivalência das definições

Representaremos as oito definições anteriores por  $D_i$ ,  $i=1, 2, \dots, 8$  para facilitar as comparações. Dizemos que

$$D_i = D_k \quad (2.28)$$

se as definições  $D_i$  e  $D_k$ , satisfazem:

- a) para cada função  $f(z)$ , são aplicáveis exatamente ao mesmo conjunto de matrizes; e
- b) produzem o mesmo valor ou conjunto de valores para cada  $f(z)$  e  $A$  admissível.

Agora, se  $D_k$  é aplicável sempre que  $D_i$  o é, mas não o contrário, e se para cada  $f(z)$  e  $A$  admissível, para os quais  $D_i$  e  $D_k$  são aplicáveis, produzirem o(s) mesmo(s) valor(es) de  $f(A)$  dizemos que  $D_i \subset D_k$ .

*Teorema 1:*  $D_2 \subset D_3 = D_4 \subset D_1 = D_6 = D_8 = D_5 \subset D_7$ .

*Prova:* Durante as várias definições já foi discutido as condições para a aplicação de cada uma das definições. Resta saber quais os valores de  $f(A)$  são os mesmos para quaisquer duas definições aplicadas a  $f(z)$  e  $A$ . Observemos que este fato está claro para  $D_5$  e  $D_7$ .

Provemos primeiro que  $D_2 \subset D_5$ .

De  $D_2$  temos

$$f(A) = \sum_{i=0}^{\infty} \frac{1}{i!} f^{(i)}(z_0)(A-z_0I)^i. \quad (2.29)$$

As somas parciais  $w_k$ 's desta série, são polinômios em  $A$  e assim,

$$w_k(A) = P (w_k(P^{-1}AP)) P^{-1}, \quad (2.30)$$

onde  $P^{-1}AP$  é uma forma normal de Jordan de  $A$ . Agora

$$w_k(P^{-1}AP) = w_k(J_1) \oplus w_k(J_2) \oplus \dots \oplus w_k(J_l), \quad (2.31)$$

conforme visto anteriormente, e para  $k \geq \max s_l$

$$w_k(J_l) = \begin{pmatrix} w_k(\lambda_l) & w'_k(\lambda_l) & \dots & \frac{1}{(s_l-1)!} w_k^{(s_l-1)}(\lambda_l) \\ 0 & & & w_k(\lambda_l) \end{pmatrix} \quad (2.32)$$

onde  $w_k(z)$  é a soma parcial de ordem  $k$  correspondente a  $f(z)$ . É claro que o termo do lado direito de (2.29) converge se, e somente se, a série  $f(z)$  e suas derivadas  $f'$ ,  $f''$ , ...,  $f^{(s_l-1)}$  convergem em  $z=\lambda_l$ , e neste caso

$$f(A) = \lim_{k \rightarrow \infty} w_k = P^{-1}G(P^{-1}AP)P, \quad (2.33)$$

onde  $G(P^{-1}AP)$  é o valor  $f(J)$  dado por Giorgi. Portanto  $D_2 \subset D_5$ .

Observemos também que  $D_3 = D_4$  pela própria definição de  $D_4$ .

Mostremos que  $D_4 \subset D_5$ .

Seja  $P^{-1}AP$  uma forma normal de Jordan, então

$$\begin{aligned} \frac{1}{2\pi i} \int \frac{f(z)}{zI-A} dz &= \frac{1}{2\pi i} P^{-1} \left[ \int \frac{f(z)dz}{P^{-1}(zI-A)P} \right] P = \\ &= \frac{1}{2\pi i} P^{-1} \left[ \int \frac{f(z)dz}{zI - P^{-1}AP} \right] P, \quad (2.34) \end{aligned}$$

já que a integral de uma matriz é a matriz da integral de cada elemento. Agora  $zI - P^{-1}AP$  é não-singular e possui uma inversa que é soma direta das inversas das matrizes  $J_i$ , onde  $J_i$  são os

blocos de Jordan. Temos

$$(zI - J_i)^{-1} = \begin{pmatrix} (z - \lambda_i)^{-1} & (z - \lambda_i)^{-2} & \dots & (z - \lambda_i)^{-s_i} \\ & & & \\ & & & \\ 0 & & & (z - \lambda_i)^{-1} \end{pmatrix} \quad (2.35)$$

e portanto

$$\frac{1}{2\pi i} \int \frac{f(z) dz}{(zI - J_i)}, \quad (2.36)$$

calculada sobre um conjunto admissível de curvas fechadas, cada uma contendo no seu interior uma raiz característica distinta de  $A$ , é

$$\frac{1}{2\pi i} \int \frac{f(z) dz}{(z - \lambda_i)^k}, \quad (2.37)$$

que é zero para cada contorno circundando  $\lambda_j \neq \lambda_i$  e igual a  $\frac{1}{(k-1)!} f^{(k-1)}(\lambda_i)$  para o contorno circundando  $\lambda_i$ , o qual é a definição de Giorgi para  $f(J_i)$ .

Mostremos que  $D_0 = D_1$ .

Como

$$G_k(A) = \frac{h_k(A)M(A)}{(A - \lambda_k I)^{s_k}} \quad (2.38)$$

podemos escrever  $D_0$  como

$$f(A) = \sum_{k=1}^l \prod_{i \neq k} (A - \lambda_i I)^{s_i} \sum_{l=0}^{s_k-1} \frac{f^{(l)}(\lambda_k)}{l!} h_k(A) (A - \lambda_k I)^l, \quad (2.39)$$

e se expressarmos o polinômio  $h_k$ , de grau  $s_k - 1$  ou menor, como

$$\sum_{m=0}^{s_k-1} b_m (A - \lambda_k I)^m, \quad (2.40)$$

obteremos

$$f(A) = \sum_{k=1}^t \prod_{i \neq k} (A - \lambda_i I)^{s_i} \sum_{l=0}^{s_k-1} \frac{f^{(l)}(\lambda_k)}{l!} \sum_{m=0}^{s_k-1} b_m (A - \lambda_k I)^{l+m}. \quad (2.41)$$

Como  $\prod_{k=1}^t (A - \lambda_k I)^{s_k} = 0$ , consideremos em (2.41) apenas os termos

onde  $l+m < s_k$ .

Comparando (2.41) com a definição de Buchheim

$$f(A) = \sum_{i=1}^t \prod_{i \neq k} (A - \lambda_i I)^{s_i} \sum_{n=0}^{s_k-1} \frac{1}{n!} \frac{d^n}{dz^n} \left[ \frac{f(z)}{\prod_{h \neq k} (z - \lambda_h)^{s_h}} \right]_{z=\lambda_k} (A - \lambda_k I)^n \quad (2.42)$$

e concluímos que a igualdade destas definições para  $f(A)$  será estabelecida se mostrarmos que os coeficientes de  $(A - \lambda_k I)^n$  são os mesmos, ou seja,

$$\frac{1}{n!} \frac{d^n}{dz^n} \left[ \frac{f(z)}{\psi_k(z)} \right]_{z=\lambda_k} = \sum_{m=0}^k \frac{f^{(m)}(\lambda_k)}{m!} b_{k-m}, \quad (2.43)$$

onde

$$\psi_k(z) = \prod_{i \neq k} (z - \lambda_i)^{s_i}. \quad (2.44)$$



Temos

$$\frac{1}{M(z)} = \sum_{i=0}^t \frac{h_i(z)}{(z-\lambda_i)^{s_i}}, \quad (2.45)$$

assim

$$\frac{1}{\psi_k(z)} = \sum_{i=1}^t \frac{h_i(z)(z-\lambda_k)^{s_k}}{(z-\lambda_i)^{s_i}} = \theta_k(z) \quad (2.46)$$

e portanto

$$\theta_k^{(m)}(\lambda_k) = h_k^{(m)}(\lambda_k), \quad \text{para } m < s_k. \quad (2.47)$$

Agora

$$b_r = \frac{1}{r!} h_k^{(r)}(\lambda_k) = \frac{1}{r!} \theta_k^{(r)}(\lambda_k), \quad \text{para } r < s_k, \quad (2.48)$$

e o termo do lado direito de (2.43) pode ser escrito como

$$\frac{1}{n!} \sum_{m=0}^n \frac{f^{(m)}(\lambda_k)}{m!} \frac{n!}{(n-m)!} \theta_k^{(n-m)}(\lambda_k), \quad (2.49)$$

que pela fórmula de Leibniz pode ser escrito como

$$\frac{1}{n!} \frac{d^n}{dz^n} \left[ \frac{f(z)}{\psi_k(z)} \right]_{z=\lambda_k}, \quad (2.50)$$

que é o termo do lado esquerdo em (2.43).

Mostremos que  $D_8 = D_5$ .

Basta observarmos como  $h_k(z)$  foi definido:

$$\frac{1}{M(z)} = \sum_{k=1}^t \frac{h_k(z)}{(z-\lambda_k)^{s_k}}. \quad (2.51)$$

Assim,

$$\sum_{k=1}^l h_k(z) \prod_{i \neq k} (z - \lambda_i)^{s_i} = 1, \quad (2.52)$$

ou seja,  $h_k(z)$  satisfaz a congruência

$$h_k(z) \prod_{i \neq k} (z - \lambda_i)^{s_i} \equiv 1 \pmod{(z - \lambda_k)^{s_k}}. \quad (2.53)$$

Logo  $h_k(z)$  pode ser usado como  $w_k(z)$  na definição de Richter. Portanto a fórmula de Richter para  $f(A)$  é idêntica com a definição de Schwerdtfeger.

Resta mostrar que  $D_6 = D_5$ .

De  $D_6$  temos

$$f(A) = \sum_{k=1}^l G_k \sum_{n=0}^{s_k-1} f^{(n)}(\lambda_k) \frac{(A - \lambda_k I)^n}{n!}, \quad (2.54)$$

logo

$$P^{-1}f(A)P = \sum_{k=1}^l P^{-1}G_kP \sum_{n=0}^{s_k-1} f^{(n)}(\lambda_k) \frac{(J - \lambda_k I)^n}{n!}, \quad (2.55)$$

onde  $J = \text{diag}(J_1, J_2, \dots, J_r)$  são os blocos de Jordan de  $A$ .

Agora,

$$\begin{aligned} P^{-1}G_kP &= P^{-1}h_k(A)P \prod_{i \neq k} (P^{-1}AP - \lambda_i I)^{s_i} = \\ &= h_k(J) \prod_{i \neq k} (J - \lambda_i I)^{s_i} \end{aligned} \quad (2.56)$$

é uma soma direta de blocos, os quais são zero, exceto aquele correspondente a  $\lambda_k$ . Como  $\sum G_i = I$  temos  $\sum P^{-1}G_iP = I$ , ou seja,  $P^{-1}G_kP$  tem zero em toda parte, exceto na posição diagonal correspondente a  $\lambda_k$ , onde tem 1. Assim o termo número  $m$  do lado direito em (2.54) é uma soma direta das submatrizes de Giorgi,

$$\begin{pmatrix} f(\lambda_m) & f'(\lambda_m) & \dots & \frac{1}{p!} f^{(p)}(\lambda_m) \\ & & & \\ & & & \\ & & & \\ 0 & & & f(\lambda_m) \end{pmatrix}, \quad (2.57)$$

correspondendo a raiz distinta  $\lambda_m$ . Logo  $f(A)$  dada por (2.54) é idêntica a dada pela definição de Giorgi.

Observe que  $D_2 \subset D_3$  pela própria definição de  $D_2$ , e que  $D_6 = D_8$  e  $D_4 \subset D_1$  pela regra da transitividade. #

**Teorema 2:** Seja  $A$  uma matriz quadrada com as raízes características distintas  $\lambda_1, \dots, \lambda_t$  e multiplicidades no polinômio minimal de  $A$   $\mu_1, \dots, \mu_t$ , respectivamente. Seja  $g(z)$  uma função univalente para cada  $\lambda_i$  e analítica para  $\lambda_i$  tal que  $\mu_i > 1$ . Seja  $f(z)$  uma função univalente para  $z_i = g(\lambda_i)$ ,  $i = 1, \dots, t$ , e analítica para  $\lambda_i$  tal que  $\mu_i > 1$ .  
Seja  $h(z) = f(g(z))$ . Então  $h(A)$  é unicamente definida pela definição de Cipolla e  $h(A) = f(g(A))$ .

*Prova:* Seja  $1 \leq p \leq t$  e  $P^{-1}AP$  uma forma de Jordan de  $A$ . Então

$$P^{-1}f(g(A))P = f(P^{-1}g(A)P), \quad (2.58)$$

onde  $P^{-1}g(A)P$  é uma soma direta dos blocos de forma

$$K_p^{(i)} = \begin{pmatrix} g(\lambda_p) & g'(\lambda_p) & \dots & (1/(\mu_p^{(i)} - 1)!) g^{(\mu_p^{(i)} - 1)}(\lambda_p) \\ & & & \\ & & & \\ & & & \\ 0 & & & g(\lambda_p) \end{pmatrix} \quad (2.59)$$

correspondendo aos blocos  $C_p^{(i)}$  da forma de Jordan de  $A$ . Mas  $P^{-1}g(A)P$  pode ser reduzido a forma de Jordan por uma matriz  $R$  que é uma soma direta das matrizes  $R_p^{(i)}$ , que transformam os blocos correspondentes  $K_p^{(i)}$ 's na forma de Jordan. Agora usando a definição de Cipolla (Giorgi (2.2.5)) para cada bloco resultante  $R^{-1(i)} K_p^{(i)} R_p^{(i)}$ , temos

$$f(R^{-1(i)} K_p^{(i)} R_p^{(i)}) = \sum_{r=0}^{\mu_p^{(i)}-1} \frac{1}{r!} f^{(r)}(g(\lambda_p)) (R^{-1(i)} K_p^{(i)} R_p^{(i)} - g(\lambda_p)I)^r. \quad (2.60)$$

Logo,

$$\begin{aligned} R^{-1(i)} f(K_p^{(i)}) R_p^{(i)} &= \\ &= R^{-1(i)} \left[ \sum_{r=0}^{\mu_p^{(i)}-1} \frac{1}{r!} f^{(r)}(g(\lambda_p)) (K_p^{(i)} - g(\lambda_p)I)^r \right] R_p^{(i)}. \end{aligned} \quad (2.61)$$

Assim

$$f(K_p^{(i)}) = \sum_{r=0}^{\mu_p^{(i)}-1} \frac{1}{r!} f^{(r)}(g(\lambda_p)) (K_p^{(i)} - g(\lambda_p)I)^r. \quad (2.62)$$

Queremos mostrar que  $f(K_p^{(i)})$  dada por (2.62) é a função  $f(g(z))$  aplicada para os blocos  $C_p^{(i)}$  de  $P^{-1}AP$ . O termo  $r=0$  em (2.62) fornece  $f(g(\lambda_p))$  como elemento da diagonal de  $f(K_p^{(i)})$ . Eliminamos estes termos da soma e escrevemos  $K_p^{(i)} - g(\lambda_p)I$  em termos de potências da matriz

$$Q = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ & & & \ddots & \\ & & & & 1 \\ 0 & & & & 0 \end{pmatrix}, \quad (2.63)$$

a qual possui 1 ao longo da primeira super-diagonal e zero em todo o resto. Assim

$$K_p^{(i)} = g(\lambda_p)I + g'(\lambda_p)Q + \dots + \frac{1}{(\mu_p^{(i)} - 1)!} g^{(\mu_p^{(i)} - 1)}(\lambda_p) Q^{(\mu_p^{(i)} - 1)} \quad (2.64)$$

e (2.62) pode ser escrita como

$$\begin{aligned} f(K_p^{(i)}) - g(\lambda_p)I &= \\ &= \sum_{r=1}^{(\mu_p^{(i)} - 1)} \frac{1}{r!} f^{(r)}(g(\lambda_p)) [g'(\lambda_p)Q + \dots + \\ &\quad + \frac{1}{(\mu_p^{(i)} - 1)!} g^{(\mu_p^{(i)} - 1)}(\lambda_p) Q^{(\mu_p^{(i)} - 1)}] \end{aligned} \quad (2.65)$$

Isto mostra que  $f(K_p^{(i)})$  (sendo um polinômio em  $Q$ ) terá todos os termos iguais ao longo de uma dada super-diagonal.

Mostremos agora que o coeficiente de  $Q^\beta$ ,  $\beta < \mu_p^{(i)}$ , em (2.65) é igual a

$$\frac{d^\beta}{dz^\beta} [f(g(z))]_{z=\lambda_p} \quad (2.66)$$

Pelo teorema multinomial o coeficiente de  $Q^\beta$  em (2.65) é, desde

que  $Q^{\mu_p^{(i)}} = 0$ ,

$$\sum_{r=1}^{\mu_p^{(i)}} \frac{f^{(r)}(g(\lambda_p))}{r!} \left\{ \sum_{(\alpha)} \frac{r!}{\alpha_1! \alpha_2! \dots \alpha_\beta!} \prod_{i=1}^{\beta} \left[ \left( \frac{g^{(i)}(z)}{i!} \right)_{z=\lambda_p} \right]^{\alpha_i} \right\}, \quad (2.67)$$

onde a soma interna é tomada sobre todas as soluções inteiras não-negativas das equações

$$\sum_{i=1}^{\beta} \alpha_i = r \quad \text{e} \quad \sum_{i=1}^{\beta} i\alpha_i = \beta. \quad (2.68)$$

Podemos reescrever o somatório acima como

$$\frac{1}{\beta!} \sum_{r=1}^{\beta} f^{(r)}(g(\lambda_p)) \left\{ \sum_{(s)} \frac{\beta!}{\alpha_1! \alpha_2! \dots \alpha_{\beta}!} \prod_{i=1}^{\beta} \left[ \left( \frac{g^{(i)}(z)}{i!} \right)_{z=\lambda_p} \right]^{\alpha_i} \right\}, \quad (2.69)$$

que é precisamente

$$\frac{1}{\beta!} \frac{d^{\beta}}{dz^{\beta}} [f(g(z))]_{z=\lambda_p}. \quad (2.70)$$

Logo as submatrizes de  $P^{-1}f(A)P$  correspondentes a um bloco de Jordan de  $P^{-1}AP$  é o mesmo que a submatriz correspondente de  $h(P^{-1}AP)$ . Assim

$$h(A) = Ph(P^{-1}AP)P^{-1} = f(g(A)). \quad \# \quad (2.71)$$

### 3- FUNÇÃO EXPONENCIAL DE UMA MATRIZ

#### 3.1- Introdução

Um dos desafios mais frequentes em equações diferenciais ordinárias é o cálculo da função exponencial de uma matriz  $f(A,t) = e^{At}$ , que é solução do problema de valor inicial

$$\begin{cases} x'(t) = Ax(t) \\ x(0) = I. \end{cases} \quad (3.1)$$

Este desafio é devido a sensibilidade do problema, ou seja, para algumas matrizes o cálculo numérico da exponencial não é bem condicionado, no sentido em que pequenas perturbações da matriz geram grandes perturbações no resultado numérico.

#### 3.2- A sensibilidade do problema

Devido a importância da exponencial de uma matriz, vamos analisar a sensibilidade do problema e verificar para que matrizes  $A$  a exponencial é bem-condicionada.

Investigaremos assim, limitações superiores para  $\phi(t)$ , onde

$$\phi(t) = \frac{\|e^{(A+E)t} - e^{At}\|}{\|e^{At}\|}, \text{ onde } t \geq 0. \quad (3.2)$$

Observação: 1) Salvo menção contrária, usaremos sempre a 2-norma (norma euclidiana) que satisfaz as seguintes propriedades:

$\|BD\| \leq \|B\| \|D\|$ , onde  $B \in \mathbb{R}^{m \times n}$  e  $D \in \mathbb{R}^{n \times p}$  e  $\forall Q$  ortogonal  $\|Q\| = 1$ .

2) Se  $A$  e  $E$  comutarem, então  $e^{A+E} = e^A e^E$  e

$$e^{(A+E)t} - e^{At} = e^{At}(e^{Et} - I) = e^{At}Et \sum_{k=0}^{\infty} (Et)^k / (k+1)! \quad (3.3)$$

Assim

$$\begin{aligned} \varphi(t) &= \frac{\|e^{(A+E)t} - e^{At}\|}{\|e^{At}\|} \leq \|e^{Et} - I\| \leq \\ &\leq \|E\|t \left\| \sum_{k=0}^{\infty} \frac{(Et)^k}{(k+1)!} \right\| \leq \|E\|t \sum_{k=0}^{\infty} \frac{\|Et\|^k}{(k+1)!} \leq \\ &\leq \|E\|te^{\|E\|t}, \end{aligned} \quad (3.4)$$

ou seja,

$$AE = EA \Rightarrow \varphi(t) \leq \|E\|te^{\|E\|t} \quad (3.5)$$

Usando a identidade<sup>1</sup>

$$e^{(A+E)t} - e^{At} = \int_0^t e^{A(t-s)} E e^{(A+E)s} ds, \quad (3.6)$$

obtemos,

$$\begin{aligned} \varphi(t) &= \frac{\|e^{(A+E)t} - e^{At}\|}{\|e^{At}\|} \leq \\ &\leq \frac{\|E\|}{\|e^{At}\|} \int_0^t \|e^{A(t-s)}\| \|e^{(A+E)s}\| ds. \end{aligned} \quad (3.7)$$

Nosso objetivo é obter limitações superiores do integrando, para isso, veremos primeiro limitações de  $e^{At}$ .

(1) Ver apêndice.



### 3.3- Limitações de $e^{At}$

(i) Com a decomposição de Schur de  $A$ ,  $Q^*AQ = D+N$ , onde  $D = \text{diag.}(\lambda_i)$  e  $N = (n_{ij})$ ,  $n_{ij} = 0$  para  $i \geq j$ , obtemos [5]:

$$\|e^{At}\| \leq e^{\alpha(A)t} \sum_{k=0}^{n-1} \frac{\|Nt\|^k}{k!}, \quad (3.8)$$

onde  $\alpha(A) = \max \{ \text{Re}(\lambda) : \lambda \in \lambda(A) \}$ . Para tanto, observe que<sup>1</sup>:

$$\begin{aligned} e^{At} &= e^{(D+N)t} = e^{Dt} + \int_0^t e^{D(t-t_1)} N e^{(D+N)t_1} dt_1 = \\ &= e^{Dt} + \int_0^t e^{D(t-t_1)} N e^{Dt_1} + \int_0^{t_1} e^{D(t_1-t_2)} N e^{(D+N)t_2} dt_2 dt_1 = \\ &= e^{Dt} + \int_0^t e^{D(t-t_1)} N e^{Dt_1} dt_1 + \\ &\quad + \int_0^t \int_0^{t_1} e^{D(t-t_1)} N e^{D(t_1-t_2)} N e^{(D+N)t_2} dt_2 dt_1. \end{aligned} \quad (3.9)$$

Proseguindo desta forma, obtemos que:

$$\begin{aligned} e^{(D+N)t} &= e^{Dt} + \sum_{k=1}^{n-1} \int_0^t \int_0^{t_1} \dots \int_0^{t_{k-1}} e^{D(t-t_1)} N \dots N e^{Dt_k} dt_k \dots dt_1 + \\ &\quad + R_n(t), \end{aligned} \quad (3.10)$$

onde

$$R_n(t) = \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} e^{D(t-t_1)} N \dots N e^{D(t_{n-1}-t_n)} N e^{(D+N)t_n} dt_n \dots dt_1. \quad (3.11)$$

(1) Usando a equação (3.6) com  $D$  no lugar de  $A$  e  $N$  no lugar de  $E$ .

Agora, o integrando em  $R_n(t)$  é zero, pois é um produto de  $n$  matrizes  $n \times n$  triangulares estritamente superiores, a saber:

$$[ e^{D(t-t_1)} N ] \dots [ e^{D(t_{n-1}-t_n)} N ]. \text{ Logo } R_n(t) = 0.$$

Assim

$$e^{(D+N)t} = e^{Dt} + \sum_{k=1}^{n-1} \int_0^t \int_0^{t_1} \dots \int_0^{t_{k-1}} e^{D(t_1-t_2)} N \dots N e^{Dt_k} dt_k \dots dt_1. \quad (3.12)$$

Como  $\|e^{Ds}\| = e^{\alpha(A)s}$  ( $s \geq 0$ ), concluímos que

$$\begin{aligned} \|e^{At}\| &= \|e^{(D+N)t}\| = \\ &= \left\| e^{Dt} + \sum_{k=1}^{n-1} \int_0^t \int_0^{t_1} \dots \int_0^{t_{k-1}} e^{D(t_1-t_2)} N \dots N e^{Dt_k} dt_k \dots dt_1 \right\| \leq \\ &\leq \|e^{Dt}\| + \left\| \int_0^t e^{D(t-t_1)} N e^{Dt_1} dt_1 \right\| + \\ &\quad + \left\| \int_0^t \int_0^{t_1} e^{D(t-t_1)} N e^{D(t_1-t_2)} N e^{Dt_2} dt_2 dt_1 \right\| + \dots \leq \\ &= e^{\alpha(A)t} + \int_0^t e^{\alpha(A)t} \|N\| dt_1 + \int_0^t \int_0^{t_1} e^{\alpha(A)t} \|N\|^2 dt_2 dt_1 + \dots = \\ &= e^{\alpha(A)t} \left( 1 + \|N\|t + \|N\|^2 \frac{t^2}{2!} + \dots + \|N\|^{n-1} \frac{t^{n-1}}{(n-1)!} \right) = \\ &= e^{\alpha(A)t} \sum_{k=0}^{n-1} \frac{\|Nt\|^k}{k!}, \text{ como queríamos.} \end{aligned}$$

(ii) Usando a definição em série de potência, obtemos facilmente para  $t \geq 0$

$$\|e^{At}\| \leq e^{\|A\|t}. \quad (3.13)$$

(iii) Usando a forma canônica de Jordan de  $A$ , ou seja,  $X^{-1}AX = \text{diag}(J_{m_1}(\lambda_1), \dots, J_{m_p}(\lambda_p)) \equiv J$  ( $\lambda_i$ 's não necessariamente distintos), onde

$$J_k \equiv J_{m_k}(\lambda_k) = \begin{pmatrix} \lambda_k & 1 & \dots & 0 \\ 0 & \lambda_k & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \lambda_k \end{pmatrix} \in \mathbb{C}^{m_k \times m_k}, \quad k=1, 2, \dots, p \quad (3.14)$$

A matriz  $X$  não é única, mas sempre assumiremos que foi escolhida de tal forma que  $k(X)$  seja mínimo ( $k(X) = \|X\| \|X^{-1}\|$ ). Assim<sup>1</sup>:

$$e^{At} = X[\text{diag}(e^{J_1 t}, \dots, e^{J_p t})]X^{-1}, \quad [5] \quad (3.15)$$

onde

$$e^{J_k t} = e^{\lambda_k t} \begin{pmatrix} 1 & t & t^2/2 & \dots & t^r/r! \\ 0 & 1 & t & \dots & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & t \\ 0 & \dots & \dots & \dots & 1 \end{pmatrix}, \quad r = m_k - 1 \quad (3.16)$$

Agora, para qualquer  $B \in \mathbb{C}^{q \times q}$  temos  $\|B\| \leq q \cdot \max |b_{ij}|$ , assim:

$$\|e^{J_k t}\| \leq m_k \cdot |e^{\lambda_k t}| \cdot \max_{0 \leq i \leq m_k - 1} (t^i / i!), \quad (3.17)$$

portanto de (3.15):

$$\|e^{At}\| \leq m K(X) e^{\alpha(A)t} \max_{0 \leq r \leq m-1} (t^r / r!), \quad (3.18)$$

onde  $m = \max \{m_1, m_2, \dots, m_p\}$ .

<sup>1</sup> (1)  $e^{At} = e^{XJX^{-1}t} = I + XJtX^{-1} + (XJ^2t^2X^{-1})/(2!) + \dots = X(I + Jt + (J^2t^2)/(2!) + \dots)X^{-1} = Xe^{Jt}X^{-1}$

(iv) Sendo  $\mu(A) = \rho\left(\frac{A^*+A}{2}\right) = \max \left\{ \mu / \mu \in \lambda\left(\frac{A^*+A}{2}\right) \right\}$ , dita a

norma log de A, temos:

$$\|e^{At}\| \leq e^{\mu(A)t} \quad \text{para } t \geq 0. \quad [5] \quad (3.19)$$

Para provar isso, seja  $v \in \mathbb{R}^n$ ,  $v \neq 0$ ,  $\|v\| = 1$ .

Definimos  $\varphi_v(t) = \|e^{At}v\|^2 = \langle e^{At}v, e^{At}v \rangle$ .

Temos

$$\begin{aligned} \frac{d\varphi_v(t)}{dt} &= \langle Ae^{At}v, e^{At}v \rangle + \langle e^{At}v, Ae^{At}v \rangle \\ &= \langle Ae^{At}v, e^{At}v \rangle + \langle A^*e^{At}v, e^{At}v \rangle \\ &= \langle (A+A^*)e^{At}v, e^{At}v \rangle. \end{aligned} \quad (3.20)$$

Agora  $(A+A^*)$  é auto-adjunta (Hermitiana), e portanto<sup>1</sup>

$$\frac{d\varphi_v(t)}{dt} \leq \rho(A+A^*) \langle e^{At}v, e^{At}v \rangle = 2 \cdot \rho\left(\frac{A+A^*}{2}\right) \cdot \varphi_v(t) = 2 \cdot \mu(A) \varphi_v(t), \quad (3.21)$$

ou seja,

$$\frac{\varphi_v'(t)}{\varphi_v(t)} \leq 2 \cdot \mu(A), \quad (3.22)$$

e integrando de 0 a t, obtemos

$$\ln|\varphi_v(t)| \leq 2t\mu(A), \quad (3.23)$$

logo

$$\varphi_v(t) \leq e^{2t\mu(A)}, \quad (3.24)$$

ou ainda,

(1)  $A+A^*$  (auto-adjunta)  $\Rightarrow \exists P$  unitária tal que

$A+A^* = P^*DP \Rightarrow \langle P^*DPv, v \rangle = \langle DPv, Pv \rangle \leq \max \lambda(D) \langle v, v \rangle$

$$\|e^{At}v\|^2 \leq e^{2t\mu(A)}, \quad (3.25)$$

ou melhor,

$$\|e^{At}v\| \leq e^{t\mu(A)}, \quad (3.26)$$

portanto

$$\|e^{At}\| \leq e^{t\mu(A)} \quad \text{para } t \geq 0. \quad (3.27)$$

Mais ainda:  $\sup_{t > 0} \|e^{At}\| = 1 \Leftrightarrow \mu(A) \leq 0$ , pois se  $\mu(A) > 0$  então

$(A+A^*)$  é positiva definida e assim  $\varphi_v'(t) > 0$ . Portanto  $\varphi_v$  é estritamente crescente. Agora  $\varphi_v(0) = 1$ , logo  $\sup_{t \geq 0} \|e^{At}\| > 1$ . Por ou-

tro lado<sup>1</sup>, se  $\mu(A) \leq 0$  temos  $\varphi_v'(t) \leq 0$  e assim  $\varphi_v$  é decrescente.

Portanto  $1 = \varphi(0) = \sup_{t \geq 0} \|e^{At}\|$ . #

(v) Se  $A = YBY^{-1}$ , então  $e^{At} = Ye^{Bt}Y^{-1}$  e portanto

$$\|e^{At}\| \leq k(Y)e^{\mu(B)t}. \quad [5] \quad (3.28)$$

Exemplo:

$$\text{Se } A = \begin{pmatrix} -1+\delta & 4 \\ 0 & -1-\delta \end{pmatrix}, \quad \delta = 10^{-6}, \text{ então:}$$

a)  $\|e^{At}\| \leq e^{\|A\|t}$  (limitação pela série de potência).

$$\|e^{At}\| \leq e^{4.24t}, \text{ pois } \|A\| = \max_{\|x\|=1} \|Ax\| \cong 4.23$$

(1)  $\frac{\varphi_v'(t)}{\varphi_v(t)} \leq 2\mu(A)$

b)  $\|e^{At}\| \leq m k(X) e^{\alpha(A)t} \max_{0 \leq j \leq m-1} t^j/j!$ , usando a forma de Jordan

$$A = XJX^{-1},$$

$$\|e^{At}\| \leq 5 \cdot 10^6 e^{(-1+\delta)t}, \text{ pois } X = \begin{pmatrix} 1 & 2 \\ 0 & -\delta \end{pmatrix}, X^{-1} = \begin{pmatrix} 1 & 2/\delta \\ 0 & -1/\delta \end{pmatrix} \text{ e}$$

portanto  $K(X) = 5 \cdot 10^6$ , e  $\alpha(A) = -1+\delta$ .

c)  $\|e^{At}\| \leq e^{\alpha(A)t} \sum_{k=0}^{n-1} \frac{\|Nt\|^k}{k!}$ , usando a forma de Schur  $Q^*AQ = D+N$ ,

$$\|e^{At}\| \leq (1+4t)e^{(-1+\delta)t}, \text{ pois } D = \begin{pmatrix} -1+\delta & 0 \\ 0 & -1-\delta \end{pmatrix} \text{ e } N = \begin{pmatrix} 0 & 4 \\ 0 & 0 \end{pmatrix}.$$

d)  $\|e^{At}\| \leq e^{\mu(A)t}$ .

$$\|e^{At}\| \leq e^{[-1+(4+\delta^2)^{1/2}]t}, \text{ pois } \frac{A^t+A}{2} = \begin{pmatrix} -1+\delta & 2 \\ 2 & -1-\delta \end{pmatrix} \text{ e assim}$$

$$\mu(A) = -1 + (\delta^2+4)^{0.5}.$$

e)  $\|e^{At}\| \leq k(Y) e^{\mu(B)t}$ ,  $A = (YBY^{-1})$

$$\text{com } Y = \begin{pmatrix} 1 & 0 \\ 0 & 1/4 \end{pmatrix}$$

$$\|e^{At}\| \leq 4 \cdot e^{(-1+(0.25+\delta^2)^{1/2})t}, \text{ pois } K(Y) = 4 \text{ e}$$

$$\frac{B^t+B}{2} = \begin{pmatrix} -1+\delta & 1/2 \\ 1/2 & -1-\delta \end{pmatrix}, \text{ logo } \mu(B) = -1 + (1/4 + \delta^2)^{0.5}.$$

Para determinados valores de  $t$ , temos a tabela 3.1.

t	$\ e^{At}\ $	s. de pot. (a)	Jordan (b)	Schur (c)	n. log (d)	n. log (e)
0	1.	1.	1.	1.	1.	1.
1	$1.3 \times 10^{-1}$	$6.9 \times 10^1$	$1.8 \times 10^6$	$1.4 \times 10^{-1}$	$1.5 \times 10^{-2}$	$3.3 \times 10^{-1}$
10	$1.8 \times 10^{-3}$	$2.5 \times 10^{18}$	$2.3 \times 10^2$	$1.9 \times 10^{-3}$	$2.2 \times 10^4$	$2.7 \times 10^{-2}$
15	$1.8 \times 10^{-5}$	$4.1 \times 10^{27}$	1.5	$1.9 \times 10^{-5}$	$3.3 \times 10^6$	$2.2 \times 10^{-3}$
25	$1.3 \times 10^{-9}$	$1.0 \times 10^{46}$	$6.9 \times 10^{-5}$	$1.4 \times 10^{-9}$	$7.2 \times 10^{10}$	$1.4 \times 10^{-5}$
30	$1.1 \times 10^{-11}$	$1.7 \times 10^{55}$	$4.8 \times 10^{-7}$	$1.1 \times 10^{-11}$	$1.1 \times 10^{13}$	$1.2 \times 10^{-6}$

Tabela 3.1- Limitações para  $e^{At}$ 

Podemos verificar através da tabela 3.1 que algumas limitações não decaem com a mesma taxa que  $e^{At}$ . Como é bem conhecido, o comportamento assintótico de  $e^{At}$  depende do sinal de  $\alpha(A)$ :

$$\lim_{t \rightarrow \infty} e^{At} = 0 \quad \Leftrightarrow \quad \alpha(A) < 0.$$

Assim, a limitação de Jordan (3.18) e a limitação de Schur (3.8) decaem precisamente quando  $e^{At}$  decai, já não acontecendo o mesmo com a limitação de Série de Potências (3.13). A limitação usando a norma log (3.19) pode ou não ter comportamento apropriado, pois  $\mu(A)$  pode ser positivo até mesmo quando  $\alpha(A)$  é negativo, como mostra o exemplo anterior. Porém, se  $-\alpha(A) < 0$ , é sempre possível escolher  $Y$  em (3.28) tal que  $\mu(YAY^{-1}) < 0$ , como mostrou o exemplo (e). Observemos também no exemplo que a limitação de Schur é melhor que a de Jordan devido ao mau condicionamento de  $A$  ( $k(X)$  grande). Entretanto, isto não é sempre verdade, geralmente a eficácia de uma limitação em relação a outra depende de  $A$  e  $t$ . Agora, quando  $A$  é normal, isto é  $A^*A=AA^*$ , temos  $\|e^{At}\| = e^{\alpha(A)t}$ , pois se  $A$  é normal então:  $\alpha(A) = \mu(A)$  em (3.18);  $m = k(X) = 1$  em (3.18);  $N = 0$ , em (3.8), assim  $\|e^{At}\| \leq e^{\alpha(A)t}$  e, como  $\|e^{At}\| \geq \|e^{\lambda t}\| \quad \forall \lambda \in \lambda(A)$ , obtemos  $\|e^{At}\| \geq e^{\alpha(A)t}$ .

Quando  $A$  não é normal, é possível  $\|e^{At}\|$  crescer inicialmente até mesmo quando  $\alpha(A)$  é negativo. Neste caso o fator  $m \cdot k(X) \cdot \max_{0 \leq t \leq m-1} t^t / t!$  em (3.18) e o fator  $\sum_{k=0}^{n-1} \frac{\|Nt\|^k}{k!}$  em (3.8) estão acomodados na vizinhança do máximo no gráfico de  $\|e^{At}\|$  (figura 3.1).

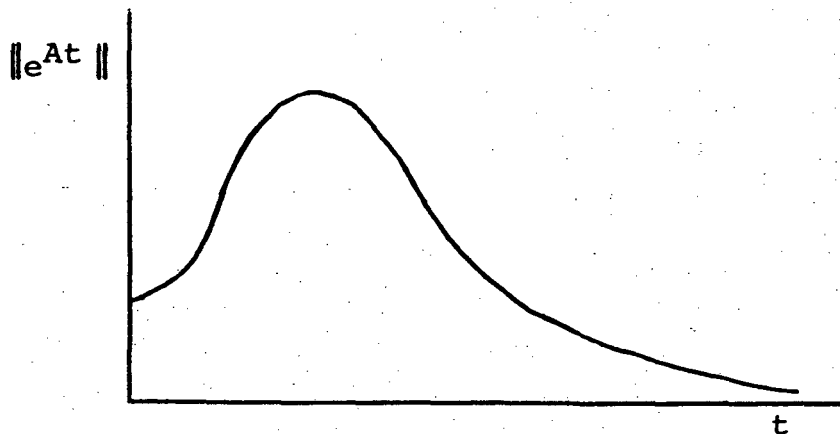


Figura 3.1- Gráfico de  $\|e^{At}\|$ .

Notemos que o crescimento de  $\|e^{At}\|$  depende do sinal de  $\mu(A)$ , no sentido que

$$\sup_{t \geq 0} \|e^{At}\| = 1 \quad \Leftrightarrow \quad \mu(A) \leq 0. \quad (3.29)$$

Vamos agora limitar o integrando em (3.7), para isso estabelecemos o seguinte lema:

*Lema 1:* Se  $\mu(t)$  é monótona crescente em  $[0, \infty)$  e  $\|e^{At}\| \leq \mu(t) e^{\beta t}$  para todo  $t \geq 0$ , então



$$\begin{aligned} \varphi(t) &= \frac{\|e^{(A+E)t} - e^{At}\|}{\|e^{At}\|} \leq \\ &\leq \|E\| t \mu(t)^2 \exp[(\beta - \alpha(A) + \|E\|\mu(t))t]. \end{aligned} \quad (3.30)$$

Prova: Seja  $c \in \mathbb{R}$  tal que  $\|e^{As}\| \leq c e^{\beta t}$  para  $\forall s \in [0, t]$  (por exemplo,  $c = \mu(t)$ ). Sejam  $U_0(t) = e^{At}$ ,

$$\begin{aligned} U_{n+1}(t) &= \int_0^t U_0(t-s) E U_n(s) ds = \\ &= \int_0^t e^{(t-s)A} E U_n(s) ds \end{aligned} \quad (3.31)$$

e provemos que:

$$a) \|U_n(t)\| \leq c^{n+1} \|E\|^n e^{\beta t} \frac{t^n}{n!} = c e^{\beta t} \left( \frac{c^n \|E\|^n t^n}{n!} \right) \quad (3.32)$$

b)  $\sum U_n$  é absolutamente convergente.

Usando indução sobre  $n$ , provemos (a). Temos para  $n = 0$

$$\|U_0(t)\| = \|e^{At}\| \leq c e^{\beta t} \leq c \|E\|^0 e^{\beta t} \frac{t^0}{0!}. \quad (3.33)$$

Suponhamos válido para  $n$ , ou seja,

$$\|U_n\| \leq c^{n+1} \|E\|^n e^{\beta t} \frac{t^n}{n!}, \quad (3.34)$$

e provemos para  $n+1$ :

$$\begin{aligned} \|U_{n+1}\| &= \left\| \int_0^t e^{(t-s)A} E U_n(s) ds \right\| \leq \\ &\leq \int_0^t \|e^{(t-s)A}\| \|E\| c^{n+1} \|E\|^n e^{\beta s} \frac{s^n}{n!} ds \leq \\ &\leq c^{n+1} \|E\|^{n+1} \int_0^t c e^{(t-s)\beta} e^{\beta s} \frac{s^n}{n!} ds = \end{aligned}$$

$$= c^{n+2} \|E\|^{n+1} e^{\beta t} \frac{t^{n+1}}{(n+1)!} \quad (3.35)$$

$\sum U_n$  é absolutamente convergente, pois

$$\|U_n\| \leq c^{n+1} \|E\|^n e^{\beta t} \frac{t^n}{n!} = c e^{\beta t} (c^n \|E\|^n \frac{t^n}{n!}), \quad (3.36)$$

logo

$$\sum \|U_n\| \leq c e^{\beta t} e^{c\|E\|t} = c e^{(\beta+c\|E\|)t} \quad (3.37)$$

Agora, temos que  $\sum U_n = U_0 + \int_0^{(t-s)A} E \sum U_n$  converge para  $v(t)$ , que é a única solução de  $v(t) = U_0(t) + \int_0^t U_0(t-s)Ev(s)ds$ ,

que é equivalente a:

$$\begin{cases} v(t)' = (A+E)v(t) \\ v(0) = I \end{cases}, \text{ ou seja,}$$

$$v(t) = e^{t(A+E)} \quad (3.38)$$

Assim, se para todo  $s \in [0, t]$   $\|e^{As}\| \leq c e^{\beta s}$ , então

$$\|e^{(A+E)s}\| \leq c e^{(\beta+c\|E\|)s}, \quad s \in [0, t]. \quad (3.39)$$

Tomando  $c = \mu(t)$  e usando a monotocidade de  $\mu(t)$  temos

$$\begin{aligned} & \int_0^t \|e^{(t-s)A}\| \|e^{(A+E)s}\| ds \leq \\ & \leq \int_0^t \mu(t-s) e^{\beta(t-s)} \mu(t) e^{(\beta+\mu(t)\|E\|)s} ds = \\ & = \int_0^t \mu(t-s) \mu(t) e^{\beta t} e^{\mu(t)\|E\|s} ds \leq \end{aligned}$$

$$\leq \mu(t)^2 e^{(\beta + \mu(t)\|E\|)t}. \quad (3.40)$$

Portanto<sup>1</sup>

$$\begin{aligned} \varphi(t) &\leq \frac{\|E\|}{\|e^{At}\|} \int_0^t \|e^{(t-s)A}\| \|e^{(A+E)s}\| ds \leq \\ &\leq \frac{\|E\|}{e^{\alpha(A)t}} t \mu(t)^2 e^{(\beta + \mu(t)\|E\|)t}. \# \end{aligned} \quad (3.41)$$

**Teorema 1:**  $\varphi(t) \leq t \|E\| e^{(\|A\| - \alpha(A) + \|E\|)t}$ . (3.42)

*Prova:* Temos  $\|e^{At}\| \leq e^{\|A\|t}$ , e aplicando o lema 1 com  $\mu(t)=1$  e  $\beta = \|A\|$  segue o resultado. #

**Teorema 2:**  $\varphi(t) \leq t \|E\| e^{(\mu(A) - \alpha(A) + \|E\|)t}$ . (3.43)

*Prova:* Temos de (3.19)  $\|e^{At}\| \leq e^{\mu(A)t}$ , e aplicando o lema 1 com  $\mu(t)=1$  e  $\beta = \mu(A)$  segue o resultado. #

**Teorema 3:** Se a decomposição de Jordan de  $A$  é dada por  $X^{-1}AX = J$ , onde  $J \equiv \text{diag}(J_{m_1}(\lambda_1), \dots, J_{m_p}(\lambda_p))$  e  $\forall i$   $J_{m_i}$  é  $m_i \times m_i$ , então

$$\varphi(t) \leq t \|E\| \mu_i(t)^2 e^{\mu_i(t)\|E\|t}, \quad (3.44)$$

onde

$$\mu_i(t) = m_i k(X) \max_{0 \leq i \leq m_i - 1} \frac{t^i}{i!}. \quad (3.45)$$

<sup>(1)</sup>  $\|e^{At}\| \geq e^{\alpha(A)t}$ , pois  $\forall \lambda \in \lambda(A)$   $\|e^{At}\| \geq e^{\lambda t}$ .

Prova: Observemos que  $\mu_i(t)$  é monótona crescente e que:

$\|e^{At}\| \leq \mu_i(t) e^{\alpha(A)t}$  de (3.18). Logo aplicando o lema 1 com  $\mu(t) = \mu_i(t)$  e  $\beta = \alpha(A)$  temos o resultado. #

**Teorema 4:** Se a decomposição de Schur de  $A$  é dada por  $Q^*AQ = D + N$ ,

então

$$\varphi(t) \leq t \|E\| \mu_s(t)^2 e^{\mu_s(t)\|E\|t}, \quad (3.46)$$

onde

$$\mu_s(t) = \sum_{k=0}^{n-1} \frac{\|Nt\|^k}{k!}. \quad (3.47)$$

Prova: Temos  $\|e^{At}\| \leq \mu_s(t) e^{\alpha(A)t}$  de (3.8), e aplicando o lema 1 com  $\mu(t) = \mu_s(t)$  e  $\beta = \alpha(A)$  segue o resultado. #

Observemos que se  $A$  é normal, ou seja,  $A^*A = AA^*$ , obtemos

$$\varphi(t) \leq t \|E\| e^{\|E\|t}. \quad (3.48)$$

Estas limitações embora não sejam necessariamente medidas exatas da sensibilidade de  $e^{At}$ , são menores quando  $A$  é normal. Isto sugere que existe uma conexão entre a normalidade de  $A$  e a sensibilidade de  $e^{At}$ . Se utilizarmos a definição

$$e^{At} = \frac{1}{2\pi i} \int_{\tau} e^{zt} (zI - A)^{-1} dz, \quad (3.49)$$

podemos encontrar uma outra forma de limitar  $\varphi(t)$ . Aqui  $\tau$  é uma curva fechada suave que circunda o espectro  $\lambda(A)$  de  $A$ . Se  $\tau$  também circunda  $\lambda(A+E)$ , então:

$$e^{(A+E)t} - e^{At} = \frac{1}{2\pi i} \int_{\tau} e^{zt} (zI - A - E)^{-1} E (zI - A)^{-1} dz. \quad (3.50)$$

Entretanto, os resultados obtidos, a partir de (3.50), não são melhores que os anteriores [5].

### 3.4- A condição da exponencial de uma matriz

No problema da inversão de uma matriz  $A$ ,  $k(A) = \|A\| \|A^{-1}\|$  é definido como a condição de uma matriz com respeito a inversão. Esse  $k(A)$  mede a sensibilidade de  $A^{-1}$  e pode ser vista na seguinte desigualdade

$$\frac{\|(A+E)^{-1} - A^{-1}\|}{\|A^{-1}\|} \leq \frac{\|E\|}{\|A\|} \frac{k(A)}{1 - \|E\| \|A^{-1}\|} \quad (\|E\| \|A^{-1}\| < 1), \quad (3.51)$$

que é facilmente verificada, pois

$$\begin{aligned} \frac{\|(A+E)^{-1} - A^{-1}\|}{\|A^{-1}\|} &\leq \frac{\|A\| \|A^{-1}\| \| (I + A^{-1}E)^{-1} - I \|}{\|A\| \|A^{-1}\|} = \\ &= \frac{k(A)}{\|A\|} \frac{\| (I + A^{-1}E)^{-1} - I \|}{\|A^{-1}\|}. \end{aligned} \quad (3.52)$$

Agora<sup>1</sup>,

$$\begin{aligned} \| (I - A^{-1}E)^{-1} - I \| &= \| I - A^{-1}E + (A^{-1}E)^2 + \dots - I \| \leq \\ &\leq \|A^{-1}E\| \|I - A^{-1}E + \dots\| \leq \\ &\leq \|A^{-1}\| \|E\| \frac{1}{1 - \|A^{-1}\| \|E\|}. \end{aligned} \quad (3.53)$$

Logo

$$\frac{\|(A+E)^{-1} - A^{-1}\|}{\|A^{-1}\|} \leq \frac{\|E\|}{\|A\|} \frac{k(A)}{1 - \|E\| \|A^{-1}\|}$$

É sempre possível escolher  $E$  tal que a limitação superior acima seja atingida. Assim, se  $k(A)$  é grande, é possível que uma pe-

(1)  $\|E\| \|A^{-1}\| < 1$

quena mudança em  $A$  induza uma mudança relativamente grande em  $A^{-1}$ . É neste sentido que  $k(A)$  mede a sensibilidade da função  $A \rightarrow A^{-1}$ .

*Definição:* A condição de exponencial de uma matriz  $A$  para um tempo  $t$  é definido por [5]:

$$V(A, t) = \lim_{\delta \rightarrow 0} V_{\delta}(A, t), \quad (3.54)$$

onde

$$V_{\delta}(A, t) = \sup_{\|E\| \leq \delta \|A\|} \frac{\|e^{(A+E)t} - e^{At}\|}{\delta \|e^{At}\|}. \quad (3.55)$$

Geometricamente,  $r = \delta \|e^{At}\| V_{\delta}(A, t)$  é o raio da menor esfera em  $\mathbb{C}^{n \times n}$  que é centrada em  $e^{At}$  e circunda a imagem do conjunto  $\{B \mid \|A-B\| \leq \delta \|A\|\}$  sob a função  $B \rightarrow e^{Bt}$ . Quando mudanças relativamente pequenas em  $A$  produzem alterações relativamente grandes em  $e^{At}$ ,  $V_{\delta}(A, t)$  é correspondentemente grande.

Definimos a derivada de uma função  $F(A)$  de  $\mathbb{C}^{n \times n}$  em  $\mathbb{C}^n$  por uma função  $DF(A)$ :  $\mathbb{C}^{n \times n} \rightarrow \mathbb{C}^n$  tal que  $\forall E$

$$F(A+E) = F(A) + DF(A)E + o(\|E\|). \quad (3.56)$$

Observemos se substituirmos

$$e^{(A+E)s} = e^{As} + \int_0^s e^{A(s-r)} E e^{(A+E)r} dr \quad (3.57)$$

em

$$e^{(A+E)t} = e^{At} + \int_0^t e^{A(t-s)} E e^{(A+E)s} ds \quad (3.58)$$

obtemos

$$e^{(A+E)t} = e^{At} + \int_0^t e^{A(t-s)} \cdot E \cdot e^{As} ds + \\ + \int_0^t e^{A(t-s)} E \int_0^s e^{A(s-r)} E e^{(A+E)r} dr ds. \quad (3.59)$$

Agora

$$\lim_{\|E\| \rightarrow 0} \frac{\left\| \int_0^t e^{A(t-s)} E \int_0^s e^{A(s-r)} E e^{(A+E)r} dr ds \right\|}{\|E\|} = 0, \quad (3.60)$$

ou seja,

$$e^{(A+E)t} = e^{At} + \int_0^t e^{A(t-s)} E e^{As} ds + o(\|E\|). \quad (3.61)$$

Assim, a derivada,  $D(e^{At})$ , da função  $F(A) = e^{At}$  é dada por

$$D(e^{At})E = \int_0^t e^{A(t-s)} E e^{As} ds. \quad (3.62)$$

Notemos que a norma do operador  $D(e^{At})$  denota a taxa de mudança de  $e^{At}$  com respeito à matriz  $A$ .

**Teorema 5:** Se  $D(e^{At})$  denota a derivada de Fréchet de  $F(A) = e^{At}$  então [5]

$$V(A, t) = \frac{\|D(e^{At})\| \|A\|}{\|e^{At}\|} \quad (3.63)$$

onde

$$\|D(e^{At})\| = \sup_{\|E\| \leq 1} \left\| \int_0^t e^{A(t-s)} E e^{As} ds \right\|. \quad (3.64)$$

*Prova:* Temos

$$e^{(A+E)t} - e^{At} = D(e^{At})E + o(\|E\|) \quad (3.65)$$

onde

$$D(e^{At})E = \int_0^t e^{A(t-s)} E e^{As} ds.$$

Assim,

$$\begin{aligned} V(A, t) &= \limsup_{\delta \rightarrow 0} \sup_{\|E\| \leq \delta \|A\|} \frac{\|D(e^{At})E + o(\|E\|)\|}{\delta \|e^{At}\|} = \\ &= \limsup_{\delta \rightarrow 0} \sup_{\|E\| \leq \delta \|A\|} \left[ \frac{\|D(e^{At})E\|}{\delta \|A\|} \frac{\|A\|}{\|e^{At}\|} + \frac{o(\delta \|A\|)}{\delta \|e^{At}\|} \right] = \\ &= \limsup_{\delta \rightarrow 0} \sup_{\|E\| \leq \delta \|A\|} \left[ \left\| D(e^{At}) \frac{E}{\delta \|A\|} \right\| \frac{\|A\|}{\|e^{At}\|} \right] = \\ &= \|D(e^{At})\| \frac{\|A\|}{\|e^{At}\|} . \# \end{aligned}$$

Corolário 1: Para qualquer  $t \geq 0$  temos  $V(A, t) \geq t \|A\|$ .

Prova: Pelo teorema temos:

$$\begin{aligned} V(A, t) &= \|D(e^{At})\| \frac{\|A\|}{\|e^{At}\|} = \\ &= \sup_{\|E\| \leq 1} \left\| \int_0^t e^{A(t-s)} E e^{As} ds \right\| \frac{\|A\|}{\|e^{At}\|} \geq \\ &\geq \left\| \int_0^t e^{A(t-s)} e^{As} ds \right\| \frac{\|A\|}{\|e^{At}\|} = \end{aligned}$$



$$= \left\| \int_0^t e^{As} ds \right\| \frac{\|A\|}{\|e^{At}\|} = t\|A\|. \quad \# \quad (3.66)$$

*Corolário 2:* Se  $A$  é normal então  $V(A, t) = t\|A\|$ .

*Prova:* Temos do corolário anterior que  $V(A, t) \geq t\|A\|$ , assim devemos mostrar que quando  $A$  é normal  $V(A, t) \leq t\|A\|$ .

Agora:

$$\begin{aligned} V(A, t) &= \|D(e^{At})\| \frac{\|A\|}{\|e^{At}\|} = \frac{\|A\|}{\|e^{At}\|} \sup_{\|E\| \leq 1} \left\| \int_0^t e^{A(t-s)} E e^{As} ds \right\| \leq \\ &\leq \frac{\|A\|}{\|e^{At}\|} \sup_{\|E\| \leq 1} \int_0^t \|e^{A(t-s)}\| \|E\| \|e^{As}\| ds \leq \\ &\leq \frac{\|A\|}{\|e^{At}\|} \int_0^t \|e^{A(t-s)}\| \|e^{As}\| ds. \end{aligned} \quad (3.67)$$

Mas quando  $A$  é normal

$$\|e^{At}\| = e^{\alpha(A)t} \quad \text{e} \quad \|e^{A(t-s)}\| = e^{\alpha(A)(t-s)}, \quad (3.68)$$

assim

$$V(A, t) \leq \frac{\|A\|}{e^{\alpha(A)t}} \int_0^t e^{\alpha(A)(t-s)} e^{\alpha(A)s} ds = t\|A\|. \quad \# \quad (3.69)$$

Concluimos destes corolários que, quando  $A$  é normal,  $V(A, t)$  é tão pequeno quanto possível. Observe, por exemplo, que se  $A$  é uma matriz estritamente triangular superior  $n \times n$ , então:

$$e^{At} = \sum_{k=0}^{n-1} \frac{(At)^k}{k!}, \quad (3.70)$$

e assim para qualquer matriz E, com  $\|E\| = 1$  e  $A^{n-1}EA^{n-1} \neq 0$ , temos:

$$\int_0^t e^{A(t-s)} E e^{As} ds = \sum_{i=0}^{n-1} \sum_{k=0}^{n-1} \frac{A^i E A^k}{(i+k+1)!} t^{i+k+1} \quad (3.71)$$

e portanto

$$\begin{aligned} V(A, t) &\geq \left\| \int_0^t e^{A(t-s)} E e^{As} ds \right\| \frac{\|A\|}{\|e^{At}\|} = \frac{\sigma(t^{2n-1})}{\sigma(t^{n-1})} \|A\| = \\ &= \sigma(t^n) \|A\|. \end{aligned} \quad (3.72)$$

Concluindo então que  $V(A, t)$  cresce com  $t^n \|A\|$  ao invés de  $t \|A\|$ , quando A é normal.

Tentaremos definir uma relação equivalente a (3.51) para o número de condição exponencial. Definimos as seguintes funções [5]:

$$\tilde{V}(A, t) = \max_{0 \leq s \leq t} V(A, s), \quad (3.73)$$

$$\sigma(A, t) = \int_0^t \|e^{-As}\| \|e^{As}\| ds. \quad (3.74)$$

A função  $\tilde{V}(A, t)$  pode ser considerada como um número de condição exponencial de A sobre o intervalo  $[0, t]$ . Agora,  $\sigma(A, t)$  é monótona crescente e  $\sigma(A, t) \geq t$ .

**Teorema 6:** Se  $\sigma(A, t) \|E\| < 1$ , então

$$\frac{\|e^{(A+E)t} - e^{At}\|}{\|e^{At}\|} \leq \frac{\|E\|}{\|A\|} \frac{\tilde{V}(A,t)}{1 - \|E\|\sigma(A,t)}. \quad (3.75)$$

Prova: Sabemos que

$$\begin{aligned} e^{(A+E)t} - e^{At} &= \int_0^t e^{A(t-s)} E e^{(A+E)s} ds, \text{ logo} \\ e^{(A+E)t} - e^{At} &= \int_0^t e^{A(t-s)} E e^{As} ds + \\ &+ \int_0^t e^{A(t-s)} E (e^{(A+E)s} - e^{As}) ds \end{aligned} \quad (3.76)$$

e assim

$$\begin{aligned} \varphi(t) = \frac{\|e^{(A+E)t} - e^{At}\|}{\|e^{At}\|} &\leq \frac{\left\| \int_0^t e^{A(t-s)} E e^{As} ds \right\|}{\|e^{At}\|} + \\ &+ \|E\| \int_0^t \|e^{-As}\| \|e^{As}\| \frac{\|e^{(A+E)s} - e^{As}\|}{\|e^{As}\|} ds \leq \\ &\leq \frac{\|E\|}{\|A\|} V(A,t) + \|E\| \int_0^t \|e^{-As}\| \|e^{As}\| \varphi(s) ds. \end{aligned} \quad (3.77)$$

Seja  $t_1 \in [0, t]$  tal que  $M = \varphi(t_1) = \max_{0 \leq s \leq t} \varphi(s)$ .

Assim

$$M \leq \frac{\|E\|}{\|A\|} V(A, t_1) + \|E\| M \sigma(A, t_1). \quad (3.78)$$

Agora  $V(A, t_1) \leq \tilde{V}(A, t)$  e  $\sigma(A, t_1) < \sigma(A, t)$ ,

assim

$$\varphi(t) \leq M \leq \frac{\|E\|}{\|A\|} \frac{\tilde{V}(A, t)}{1 - \sigma(A, t)\|E\|}, \text{ pois} \quad (3.79)$$

$$\begin{aligned}
M &\leq \frac{\|E\|}{\|A\|} \tilde{V}(A, t) + \|E\| M \sigma(A, t) = \\
&= \frac{\|E\|}{\|A\|} \tilde{V}(A, t) \left[ 1 + \frac{M \sigma(A, t) \|A\|}{\tilde{V}(A, t)} \right] \leq \\
&\leq \frac{\|E\|}{\|A\|} \tilde{V}(A, t) \left[ 1 + \sigma(A, t) \|E\| + \frac{\sigma(A, t) \|E\| M \|A\|}{\tilde{V}(A, t)} \right] = \\
&= \frac{\|E\|}{\|A\|} \tilde{V}(A, t) \left[ 1 + \sigma(A, t) \|E\| + \frac{\sigma(A, t)^2 \|E\| M \|A\|}{\tilde{V}(A, t)} \right] \leq \\
&\leq \frac{\|E\|}{\|A\|} \tilde{V}(A, t) \left[ 1 + \sigma(A, t) \|E\| + \sigma(A, t)^2 \|E\|^2 + \right. \\
&\quad \left. + \frac{\sigma(A, t)^2 \|E\|^2 \sigma(A, t) M \|A\|}{\tilde{V}(A, t)} \right] = \\
&= \frac{\|E\|}{\|A\|} \tilde{V}(A, t) \left[ 1 + \sigma(A, t) \|E\| + \sigma(A, t)^2 \|E\|^2 + \right. \\
&\quad \left. + \frac{\sigma(A, t)^3 \|E\|^2 M \|A\|}{\tilde{V}(A, t)} \right] \leq \dots \leq \\
&\leq \frac{\|E\|}{\|A\|} \frac{\tilde{V}(A, t)}{1 - \sigma(A, t) \|E\|} \cdot \#
\end{aligned}$$

Observemos que se  $A$  é normal, podemos mudar  $\tilde{V}(A, t)$  com  $V(A, t)$  e assim temos novamente menores limitações de  $\sigma(t)$ .

## 4- MÉTODOS DE APROXIMAÇÃO

## 4.1- Aproximação pela Série de Taylor

Muitas das definições apresentadas anteriormente envolvem o cálculo de autovalores. Vamos então considerar um método para o cálculo da função que a primeira vista não envolve os autovalores. A idéia é que se  $g(z)$  aproxima  $f(z)$  em  $\lambda(A)$ , então  $g(A)$  aproxima  $f(A)$ . Por exemplo o truncamento da série de Taylor,

$$f(A) = e^A \cong I + A + \frac{A^2}{2!} + \dots + \frac{A^q}{q!}. \quad (4.1)$$

Vamos analisar a diferença  $\|f(A) - g(A)\|$  usando a representação de Jordan e Schur.

*Teorema 1:* Seja  $X^{-1}AX=J$  a forma canônica de Jordan de  $A \in \mathbb{C}^{n \times n}$  com  $J = J_1 \oplus J_2 \oplus \dots \oplus J_p$  e

$$J_i = \begin{bmatrix} \lambda_i & 1 & & \dots & 0 \\ 0 & \lambda_i & 1 & & \vdots \\ & & \ddots & \ddots & \vdots \\ \vdots & & & & 1 \\ 0 & \dots & & & \lambda_i \end{bmatrix}_{m_i \times m_i} \quad (4.2)$$

sendo um bloco de Jordan  $m_i \times m_i$ . Se  $f(z)$  e  $g(z)$  são analíticas sobre um aberto contendo  $\lambda(A)$ , então,

$$\|f(A) - g(A)\| \leq K(X) \max_{\substack{1 \leq i \leq p \\ 0 \leq r \leq m_i - 1}} m_i \frac{|f^{(r)}(\lambda_i) - g^{(r)}(\lambda_i)|}{r!}. \quad (4.3)$$

*Prova:* Definindo  $h(z) = f(z) - g(z)$  temos

$$\begin{aligned} \|f(A) - g(A)\| &= \|X \text{diag}(h(J_1), \dots, h(J_p)) X^{-1}\| \leq \\ &\leq K(X) \max_{1 \leq i \leq p} \|h(J_i)\|. \end{aligned} \quad (4.4)$$

Agora

$$h(J_i) = \begin{bmatrix} h(\lambda_i) & h'(\lambda_i) & \dots & \frac{h^{(m_i-1)}(\lambda_i)}{(m_i-1)!} \\ 0 & h(\lambda_i) & & \\ & & \ddots & \\ 0 & & & h'(\lambda_i) \\ & & & h(\lambda_i) \end{bmatrix}_{m_i \times m_i}. \quad (4.5)$$

Logo<sup>1</sup>

$$\|h(J_i)\| \leq m_i \max_{0 \leq r \leq m_i-1} \frac{|h^{(r)}(\lambda_i)|}{r!}, \quad (4.6)$$

o que prova o teorema. #

**Teorema 2:** Seja  $Q^*AQ = \text{diag}(\lambda) + N$  a decomposição de Schur de  $A \in \mathbb{C}^{n \times n}$ , com  $N$  sendo estritamente triangular superior. Se  $f(z)$  e  $g(z)$  são analíticas no conjunto convexo fechado  $\Omega$  cujo interior contém  $\lambda(A)$ , então

$$\|f(A) - g(A)\|_F \leq \sum_{r=0}^{n-1} \delta_r \frac{\|N\|^r}{r!}, \quad (4.7)$$

onde

$$\delta_r = \sup_{z \in \Omega} |f^{(r)}(z) - g^{(r)}(z)|. \quad (4.8)$$

*Prova:* Ver [2].

**Teorema 3:** Se  $f(z)$  é analítica sobre um disco aberto de raio  $r$

(1) Propriedade da norma,  $\|A\|_2 \leq m \max_{i,j} |a_{ij}|$ , onde  $m$  = ordem de  $A$ .

contendo os autovalores de  $A$ ,  $f(z) = \sum_{k=0}^{\infty} \alpha_k z^k$ , então

$$\|f(A) - \sum_{k=0}^q \alpha_k A^k\| \leq \frac{n}{(q+1)!} \max_{0 \leq s \leq 1} \|A^{q+1} f^{(q+1)}(As)\|. \quad (4.9)$$

*Prova:* Definimos a matriz  $E_{q+1}(s)$  por:

$$f(As) = \sum_{k=0}^q \alpha_k (As)^k + E_{q+1}(s), \quad 0 \leq s \leq 1. \quad (4.10)$$

$F_{il}(s)$ , o elemento  $(i, l)$  de  $f(As)$ , é necessariamente analítico e assim

$$F_{il}(s) = \left[ \sum_{k=0}^q \frac{F_{il}^{(k)}(0)}{k!} s^k \right] + \frac{F_{il}^{(q+1)}(\varepsilon_{il})}{(q+1)!} s^{q+1}, \quad (4.11)$$

onde  $\varepsilon_{il}$  satisfaz  $0 \leq \varepsilon_{il} \leq s \leq 1$  (Fórmula de Taylor em torno de zero.). Comparando as potências de  $s$  em (4.10) e (4.11), concluímos que, o  $(i, l)$  elemento de  $E_{q+1}(s)$ , que chamamos  $e_{il}(s)$ , tem a forma

$$e_{il}(s) = \frac{f_{il}^{(q+1)}(\varepsilon_{il})}{(q+1)!} s^{q+1}. \quad (4.12)$$

Agora  $F_{il}^{(q+1)}(s)$  é o  $(i, l)$  elemento de  $A^{q+1} f^{(q+1)}(As)$  e portanto<sup>1</sup>

$$|e_{il}| \leq \max_{0 \leq s \leq 1} \frac{F_{il}^{(q+1)}(\varepsilon_{il})}{(q+1)!} \leq \max_{0 \leq s \leq 1} \frac{\|A^{q+1} f^{(q+1)}(As)\|}{(q+1)!} \leq$$

(1)  $\|B\|_2 \leq (mn)^{1/2} \max_{i,l} |b_{il}|$  e  $\max_{i,l} |b_{il}| \leq \|B\|_2$  para  $B_{m \times n}$ .

$$\leq \frac{n}{(q+1)!} \max_{0 \leq s \leq 1} \|A^{q+1} f^{(q+1)}(As)\|. \quad \#$$

Exemplo:

$$A = \begin{pmatrix} -49 & 24 \\ -64 & 31 \end{pmatrix} \text{ então } e^A \cong \begin{pmatrix} -0.735759 & 0.0551819 \\ -1.471518 & 1.103138 \end{pmatrix}$$

para  $q = 59$ , com erro

$$\|e^A - \sum_{k=0}^q \frac{A^k}{k!}\| \leq \frac{n}{(q+1)!} \max_{0 \leq s \leq 1} \|A^{q+1} e^{As}\| \leq 10^{-60}.$$

Assim para a função exponencial de uma matriz o erro cometido com o truncamento é

$$E_{q+1}(s) \leq \frac{n \|A^{q+1}\|}{(q+1)!} \max_{0 \leq s \leq 1} \|e^{sA}\|, \quad (4.13)$$

$$\text{agora, } \max_{0 \leq s \leq 1} \|e^{sA}\| \leq \max_{0 \leq s \leq 1} \left\| I + sA + \frac{(sA)^2}{2!} + \dots \right\| \leq$$

$$\leq \max_{0 \leq s \leq 1} \left( \|I\| + \|sA\| + \frac{\|(sA)^2\|}{2!} + \dots \right) =$$

$$= \max_{0 \leq s \leq 1} e^{\|sA\|} = \max_{0 \leq s \leq 1} e^{s\|A\|} \leq e^{\|A\|}, \quad (4.14)$$

portanto

$$E_{q+1} \leq \frac{n \|A^{q+1}\|}{(q+1)!} e^{\|A\|}. \quad (4.15)$$



#### 4.2- Aproximações de Padé

Seja  $\sum_{i=0}^{\infty} a_i z^i$  a representação em série de Taylor de  $f(z)$ ,

ou seja,

$$f(z) = \sum_{i=0}^{\infty} a_i z^i \quad (4.16)$$

Estamos interessados com a relação existente entre os coeficientes da série de Taylor e os valores da função. Estamos aproximando a função por polinômios muito grandes, o que na prática é inconveniente. As Aproximações de Padé são aproximações por frações racionais para o valor da função. A idéia é conseguir frações racionais cujas expansões em séries de Taylor coincidam com a série de Taylor da função o máximo possível. Por exemplo, seja

$$f(x) = \left( \frac{1 + 2x}{1 + x} \right)^{1/2} = 1 + \frac{1}{2} x - \frac{5}{8} x^2 + \frac{13}{16} x^3 - \dots$$

Queremos uma aproximação da forma  $\frac{a + bx}{c + dx}$ . A série não

converge para  $|x| > 1/2$ , mas  $f(x)$  é uma função suave para  $0 \leq x < \infty$ , com sua imagem percorrendo de 1 a  $2^{1/2}$ . Agora,

$$\frac{a + bx}{c + dx} = \frac{a}{c} + \frac{bc - da}{c^2} x - \frac{(bc - da)2cd}{c^4} \frac{x^2}{2!} + \dots \quad (4.17)$$

Igualando os três primeiros coeficientes das duas séries acima, obtemos

$$\frac{a}{c} = 1, \quad \frac{bc - da}{c^2} = \frac{1}{2}, \quad \frac{-(bc - da)2cd}{c^4} = \frac{-5}{4}$$

Fazendo  $a = c = 1$ , obtemos  $d = 5/4$  e  $b = 7/4$ , produzindo a

aproximação

$$\frac{1 + 7x/4}{1 + 5x/4} = 1 + \frac{1}{2}x - \frac{5}{8}x^2 + \frac{25}{32}x^3 - \frac{125}{128}x^4 + \dots$$

que tem o valor 1.4 para  $x$  no infinito. Uma outra aproximação usando os quatro primeiros coeficientes é

$$\frac{1 + 13x/4 + 41x^2/16}{1 + 11x/4 + 29x^2/16} \xrightarrow{x \rightarrow \infty} \frac{41}{29} \cong 1.413793103, \text{ que está bem}$$

mais próxima de  $2^{1/2} = 1.414213562\dots$

Definição: A  $[L/M]$  Aproximação de Padé para  $f(z) = \sum_{i=0}^{\infty} a_i z^i$  é a fração racional

$$[L/M] = \frac{P_L(z)}{Q_M(z)} = \frac{p_0 + p_1 z + \dots + p_L z^L}{1 + q_1 z + \dots + q_M z^M}, \quad (4.18)$$

onde  $P_L$  e  $Q_M$  são primos entre si e os  $L+M+1$  primeiros coeficientes da sua série de potências coincidem com  $a_0, a_1, \dots, a_{L+M}$ .

Ou seja,

$$\sum_{i=0}^{\infty} a_i z^i = \frac{P_L(z)}{Q_M(z)} + o(z^{L+M+1}) \quad (4.19)$$

Exemplo: Se  $f(z) = 1 - z/2 + z^2/3 + \dots$ , então

$$[1/0] = 1 - z/2 = f(z) + o(z^2)$$

$$[0/1] = \frac{1}{1 + z/2} = f(z) + o(z^2)$$

$$[1/1] = \frac{1 + z/6}{1 + 2z/3} = f(z) + o(z^3).$$

Da equação (4.19), multiplicando por  $Q_M(z)$ , obtemos

$$\begin{aligned}
 (1 + q_1 z + \dots + q_M z^M) (a_0 + a_1 z + \dots) &= \\
 = p_0 + p_1 z + \dots + p_L z^L + o(z^{M+L+1}). & \quad (4.20)
 \end{aligned}$$

Assim obtemos as equações para os coeficientes:

$$\left. \begin{aligned}
 a_0 &= p_0 \\
 a_1 + a_0 q_1 &= p_1 \\
 a_2 + a_1 q_1 + a_2 q_2 &= p_2 \\
 \dots &\dots \\
 a_L + a_{L-1} q_1 + \dots + a_0 q_L &= p_L
 \end{aligned} \right\} \quad (4.21)$$

$$\left. \begin{aligned}
 a_{L+1} + a_L q_1 + \dots + a_{L-M+1} q_M &= 0 \\
 \dots &\dots \\
 a_{L+M} + a_{L+M-1} q_1 + \dots + a_L q_M &= 0
 \end{aligned} \right\} \quad (4.22)$$

Se  $i < 0$ , definimos  $a_i = 0$ .

Podemos escrever (4.22) da seguinte forma

$$\begin{pmatrix}
 a_{L-M+1} & a_{L-M+2} & a_{L-M+3} & a_L \\
 a_{L-M+2} & a_{L-M+3} & a_{L-M+4} & a_{L+1} \\
 \dots & \dots & \dots & \dots \\
 a_L & a_{L+1} & a_{L+2} & a_{L+M-1}
 \end{pmatrix}
 \begin{pmatrix}
 q_M \\
 q_{M-1} \\
 \dots \\
 q_1
 \end{pmatrix}
 = - \begin{pmatrix}
 a_{L+1} \\
 a_{L+2} \\
 \dots \\
 a_{L+M}
 \end{pmatrix} \quad (4.23)$$

do qual os  $q_i$  podem ser encontrados. Os coeficientes do numerador  $p_0, p_1, \dots, p_L$  seguem imediatamente de (4.21). Assim (4.23) e (4.21), normalmente, determinam o numerador e o denominador da  $[L/M]$  Aproximação de Padé que concorda com a série de potências através da ordem  $z^{L+M}$ .

Vale lembrar que a condição  $Q_M(0)=1$  difere da definição clássica de Frobenius (1881) e Padé (1892) que requereram simplesmente que  $Q_M(z) \neq 0$ . Esta diferença pode ser vista através do seguinte exemplo: Se  $f(x) = 1 + x^2 + \dots$  e  $L = M = 1$ , podemos facilmente verificar que  $P_1(x) = Q_1(x) = x$ ,  $P_1(x)/Q_1(x) = 1$  satisfaz

$$Q_1(x) \sum_{i=0}^{\infty} a_i x^i - P_1(x) = o(x^3)$$

mas não satisfaz

$$\sum_{i=0}^{\infty} a_i x^i - \frac{P_1(x)}{Q_1(x)} = o(x^3).$$

A origem deste estudo é a série de potências  $\sum a_i z^i$ , mas nem sempre existe série de Maclaurin para uma função  $f(z)$ . Assim, dada a série de potências  $\sum_{i=0}^{\infty} a_i z^i$ , (4.23) e (4.21) mostram como as aproximações de Padé são construídas.

Uma série de potências  $\sum a_i z^i$  ou converge apenas para  $z=0$  ou existe um número  $r>0$  (podendo ser  $r$  infinito) tal que a série converge absolutamente no intervalo aberto  $(-r, r)$  e diverge fora do intervalo fechado  $[-r, r]$ . Nos extremos  $-r$  e  $r$ , a série pode convergir ou divergir, conforme o caso ( $r$  é chamado raio de convergência).

Se o raio de convergência é zero, a série de potências contém poucas informações sobre a função, mas se uma seqüência de aproximações de Padé da série converge para uma função  $g(z)$  com  $z \in D$  ( $D$  disco com raio maior que zero), então concluímos que  $g(z)$  é uma função com a dada série de potências. Agora se a dada série de potências converge para uma função em  $(-r, r)$  com  $r \in (0, \infty)$ , então uma seqüência de aproximações de Padé pode convergir para  $z \in D$  onde  $D$  é um domínio maior que  $(-r, r)$ , ou seja, estendemos o domínio de convergência.

Exemplo: Seja

$$f(z) = \left( \frac{1 + z/2}{1 + 2z} \right)^{1/2} = 1 - \frac{3z}{4} + \frac{39z^2}{32} - \dots,$$

usando a equação (4.23) calculamos o  $b_1=13/8$  da  $[1/1]$  aproximação de Padé e da equação (4.21) obtemos  $a_0=1$  e  $a_1=7/8$ , assim

$$[1/1] = \frac{1 + 7z/8}{1 + 13z/8} \quad e$$

$$\left(1 + \frac{13}{8}z\right) \left(1 - \frac{3}{4}z + \frac{39}{32}z^2 \dots\right) = 1 + \frac{7}{8}z + o(z^3)$$

Na figura 1 está comparada esta aproximação com  $f(z)$  para  $z \geq 0$ . Em particular,  $f(\infty)=0.5$  e  $[1/1](\infty)=7/13=0.54$ , dando 8% de precisão no infinito. Este exemplo mostra o grau de aproximação para uma função com raio de convergência de  $1/2$ , usando somente 3 termos da série.

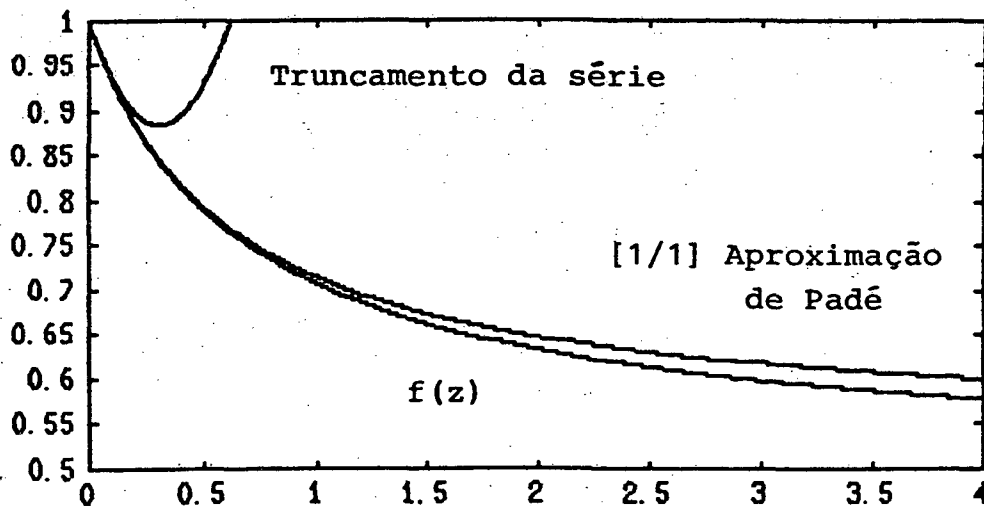


Figura 1-Comparação entre o truncamento da série com a  $[1/1]$  aproximação de Padé.

Se a regra de Cramer for usada, podemos calcular  $q_1, q_2, \dots, q_M$  de (4.23) e assim o denominador de (4.18). Logo

$$Q_M(z) = \frac{1}{\Delta} \begin{vmatrix} a_{L-M+1} & a_{L-M+2} & \dots & a_L & a_{L+1} \\ a_{L-M+2} & a_{L-M+3} & \dots & a_{L+1} & a_{L+2} \\ \dots & \dots & \dots & \dots & \dots \\ a_{L-1} & a_L & \dots & a_{L+M-2} & a_{L+M-1} \\ a_L & a_{L+1} & \dots & a_{L+M-1} & a_{L+M} \\ z^M & z^{M-1} & \dots & z & 1 \end{vmatrix}, \quad (4.24)$$

onde

$$\Delta = \begin{vmatrix} a_{L-M+1} & a_{L-M+2} & \dots & a_L \\ a_{L-M+2} & a_{L-M+3} & \dots & a_{L+1} \\ \dots & \dots & \dots & \dots \\ a_{L-1} & a_L & \dots & a_{L+M-2} \\ a_L & a_{L+1} & \dots & a_{L+M-1} \end{vmatrix}. \quad (4.25)$$

Consideremos agora

$$Q_M(z) \sum_{i=0}^{\infty} a_i z^i = \frac{1}{\Delta} \begin{vmatrix} a_{L-M+1} & a_{L-M+2} & \dots & a_{L+1} \\ a_{L-M+2} & a_{L-M+3} & \dots & a_{L+2} \\ \dots & \dots & \dots & \dots \\ a_{L-1} & a_L & \dots & a_{L+M-1} \\ a_L & a_{L+1} & \dots & a_{L+M} \\ \sum_{i=0}^{\infty} a_i z^{M+i} & \sum_{i=0}^{\infty} a_i z^{M+i-1} & \dots & \sum_{i=0}^{\infty} a_i z^i \end{vmatrix}. \quad (4.26)$$

Lembrando que  $a_i = 0$  se  $i < 0$  e subtraindo da última linha,  $z^{L+1}$  vezes a primeira linha,  $z^{L+2}$  vezes a segunda linha, ...,  $z^{L+M}$  vezes a penúltima linha, reduzimos as séries da última linha, anulando  $M$  termos. Usando os termos iniciais dessas séries, definimos:

$$P_L(z) = \frac{1}{\Delta} \begin{vmatrix} a_{L-M+1} & a_{L-M+2} & \dots & a_{L+1} \\ a_{L-M+2} & a_{L-M+3} & \dots & a_{L+2} \\ \dots & \dots & \dots & \dots \\ a_{L-1} & a_L & \dots & a_{L+M-1} \\ a_L & a_{L+1} & \dots & a_{L+M} \\ L-M & L-M+1 & \dots & L \\ \sum_{i=0}^{L-M} a_i z^{M+i} & \sum_{i=0}^{L-M+1} a_i z^{M+i-1} & \dots & \sum_{i=0}^L a_i z^i \end{vmatrix}, \quad (4.27)$$

Teorema 4: Com  $P_L(z)$  e  $Q_M(z)$  dados por (4.24) e (4.27) respectivamente,

$$Q_M(z) \sum_{i=0}^{\infty} a_i z^i - P_L(z) = o(z^{L+M+1}). \quad (4.28)$$

Prova: Note que o resto é

$$\begin{aligned}
 & Q_M(z) \sum_{i=0}^{\infty} a_i z^i - P_L(z) = \\
 & = \frac{1}{\Delta} \begin{vmatrix} a_{L-M+1} & a_{L-M+2} & \dots & a_{L+1} \\ a_{L-M+2} & a_{L-M+3} & \dots & a_{L+2} \\ \dots & \dots & \dots & \dots \\ a_{L-1} & a_L & \dots & a_{L+M-1} \\ a_L & a_{L+1} & \dots & a_{L+M} \\ \dots & \dots & \dots & \dots \\ \sum_{i=L+1}^{\infty} a_i z^{M+i} & \sum_{i=L+2}^{\infty} a_i z^{M+i-1} & \dots & \sum_{i=L+M+1}^{\infty} a_i z^i \end{vmatrix} = \\
 & = \frac{1}{\Delta} \sum_{i=1}^{\infty} z^{L+M+i} \begin{vmatrix} a_{L-M+1} & a_{L-M+2} & \dots & a_{L+1} \\ a_{L-M+2} & a_{L-M+3} & \dots & a_{L+2} \\ \dots & \dots & \dots & \dots \\ a_{L-1} & a_L & \dots & a_{L+M-1} \\ a_L & a_{L+1} & \dots & a_{L+M} \\ a_{L+1} & a_{L+i+1} & \dots & a_{L+M+i} \end{vmatrix} \quad \# \quad (4.29)
 \end{aligned}$$

A equação (4.29) é ocasionalmente usada como forma de calcular o erro na aproximação de Padé. Consideremos

$$\Delta = \begin{vmatrix} a_{L-M+1} & a_{L-M+2} & \dots & a_L \\ a_{L-M+2} & a_{L-M+3} & \dots & a_{L+1} \\ \dots & \dots & \dots & \dots \\ a_{L-1} & a_L & \dots & a_{L+M-2} \\ a_L & a_{L+1} & \dots & a_{L+M-1} \end{vmatrix} .$$

Note que se  $\Delta \neq 0$ , então as equações lineares (4.23) são não-singulares e a solução dada por (4.24) é única. Assim, podemos dividir (4.28) por  $Q_M(z)$ , produzindo

$$\sum_{i=0}^{\infty} a_i z^i - \frac{P_L(z)}{Q_M(z)} = o(z^{L+M+1}), \quad (4.30)$$

o que prova o seguinte teorema.

**Teorema 5:** [Jacobi, 1846] A  $[L/M]$  Aproximação de Padé de  $\sum a_i z^i$  é dada por

$$[L/M] = \frac{P_L(z)}{Q_M(z)}, \text{ onde } Q_M \text{ e } P_L \text{ são dadas respecti-}$$

mente por (4.24) e (4.27).

**Teorema 6:** (Unicidade). Quando existe, a  $[L/M]$  aproximação de Padé para qualquer série de potências formal  $A(z)$ , é única.

*Prova:* Assumimos que existem duas destas aproximações de Padé  $X(z)/Y(z)$  e  $U(z)/V(z)$ , onde o grau de  $X$  e  $U$  é menor ou igual a  $L$  e que o grau de  $U$  e  $V$  é menor ou igual a  $M$ . Então devemos ter, por (4.19)

$$X(z)/Y(z) - U(z)/V(z) = o(z^{L+M+1}) \quad (4.31)$$

visto que ambas aproximam a mesma série. Se multiplicarmos (4.31) por  $Y(z)V(z)$ , obtemos

$$X(z)V(z) - U(z)Y(z) = o(z^{L+M+1}). \quad (4.32)$$



Mas o lado esquerdo de (4.32) é um polinômio de grau no máximo  $L+M$ , e desta forma é identicamente nulo. Visto que nem  $Y$  nem  $V$  é identicamente nulo, concluímos que

$$X/Y = U/V . \quad (4.33)$$

Pela definição, ambos  $X$  e  $Y$ , e  $U$  e  $V$  são relativamente primos e  $Y(0)=V(0)=1.0$ , temos mostrado que as duas, supostamente aproximações de Padé distintas, são as mesmas. #

Padé (1892) ajustou as aproximações em uma ordem semi-infinita ou tabela. Pela tábua de Padé entendemos o arranjo:

M \ L	0	1	2	3
0	[0/0]	[1/0]	[2/0]	[3/0]
1	[0/1]	[1/1]	[2/1]	[3/1]
2	[0/2]	[1/2]	[2/2]	[3/2]
3	[0/3]	[1/3]	[2/3]	[3/3]
4	[0/4]	[1/4]	[2/4]	[3/4]

Tabela 4.1- Tábua de Padé

As somas parciais da série de Taylor ocupam a primeira linha da tabela 4.1.

#### 4.3- Aproximação de Padé para a função exponencial

Por Taylor,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \dots \quad (4.34)$$

e desejamos calcular  $Q_M(z)$ , da  $[L/M]$  aproximação de Padé,

$$Q_M(z) = \frac{1}{C_M} \begin{vmatrix} \frac{1}{(L-M+1)!} & \frac{1}{(L-M+2)!} & \cdots & \frac{1}{L!} & \frac{1}{(L+1)!} \\ \frac{1}{(L-M+2)!} & \frac{1}{(L-M+3)!} & \cdots & \frac{1}{(L+1)!} & \frac{1}{(L+2)!} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{L!} & \frac{1}{(L+1)!} & \cdots & \frac{1}{(L-M+1)!} & \frac{1}{(L-M)!} \\ z^M & z^{M-1} & \cdots & z & 1 \end{vmatrix}, \quad (4.35)$$

onde

$$C_M = \begin{vmatrix} \frac{1}{(L-M+1)!} & \frac{1}{(L-M+2)!} & \cdots & \frac{1}{L!} \\ \frac{1}{(L-M+2)!} & \frac{1}{(L-M+3)!} & \cdots & \frac{1}{(L+1)!} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{1}{L!} & \frac{1}{(L+1)!} & \cdots & \frac{1}{(L+M-1)!} \end{vmatrix}. \quad (4.36)$$

Removendo o denominador de cada linha, usando as propriedades de determinante e a definição:

$$p = \prod_{i=1}^n \frac{1}{(L+i-1)!}, \quad (4.37)$$

obtemos

$$C_M = p \begin{vmatrix} \frac{L!}{(L-M+1)!} & \frac{L!}{(L-M+2)!} & \cdots & L & 1 \\ \frac{(L+1)!}{(L-M+2)!} & \frac{(L+1)!}{(L-M+3)!} & \cdots & L+1 & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{(L+M-1)!}{L!} & \frac{(L+M-1)!}{(L+1)!} & \cdots & L+M-1 & 1 \end{vmatrix}, \quad (4.38)$$

subtraindo a  $(M-1)$ -ésima linha da  $M$ -ésima, a  $(M-2)$ -ésima linha da  $(M-1)$ -ésima, ..., a primeira linha da segunda, e usando a identidade

$$\frac{r!}{s!} - \frac{(r-1)!}{(s-1)!} = (r-s) \frac{(r-1)!}{s!}, \quad (4.39)$$

obtemos

$$C_M = p^{(M-1)!} \begin{vmatrix} \frac{L!}{(L-M+1)!} & \frac{L!}{(L-M+2)!} & \dots & L & 1 \\ \frac{L!}{(L-M+2)!} & \frac{L!}{(L-M+3)!} & \dots & 1 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \frac{(L+M-2)!}{L!} & \frac{(L+M-2)!}{(L+1)!} & \dots & 1 & 0 \end{vmatrix} =$$

$$= p(-1)^{M-1} (M-1)! \begin{vmatrix} \frac{L!}{(L-M+2)!} & \frac{L!}{(L-M+3)!} & \dots & 1 \\ \frac{(L+1)!}{(L-M+3)!} & \frac{(L+1)!}{(L-M+4)!} & \dots & 1 \\ \dots & \dots & \dots & \dots \\ \frac{(L+M-2)!}{L!} & \frac{(L+M-2)!}{(L+1)!} & \dots & 1 \end{vmatrix}. \quad (4.40)$$

Agora, (4.40) é um determinante  $(M-1) \times (M-1)$  com a mesma forma de (4.38), apenas com  $M-1$  no lugar de  $M$ , assim com o mesmo procedimento chegamos a<sup>1</sup>

$$C_M = p \prod_{i=1}^M (-1)^{i-1} (i-1)! =$$

$$= (-1)^{M(M-1)/2} \prod_{i=1}^M \frac{(i-1)!}{(L+i-1)!}. \quad (4.41)$$

(1) Usando a identidade (4.39) na coluna 1 obtemos  $r-s=M-1$ , na coluna 2 obtemos  $r-s=M-2$ ,...

Consideremos o coeficiente de  $(-z)^k$  em  $Q_M(z)$ ,

$$(-1)^k q_k = \frac{1}{C_M} \begin{vmatrix} \frac{1}{(L-M+1)!} & \frac{1}{(L-M+2)!} & \cdots & \frac{1}{(L-k+1)!} & \cdots & \frac{1}{(L+1)!} \\ \frac{1}{(L-M+2)!} & \frac{1}{(L-M+3)!} & \cdots & \frac{1}{(L-k+2)!} & \cdots & \frac{1}{(L+2)!} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{1}{L!} & \frac{1}{(L+1)!} & \cdots & \frac{1}{(L+M-k)!} & \cdots & \frac{1}{(L+M)!} \end{vmatrix} \quad (4.42)$$

onde a coluna rodeada por  $| \quad |$  é retirada. Usando o mesmo procedimento anterior, definimos

$$\tilde{p} = \prod_{i=0}^M \frac{1}{(L+i)!} \quad (4.43)$$

e assim

$$(-1)^k q_k = \frac{\tilde{p}}{C_M} \begin{vmatrix} \frac{(L+1)!}{(L-M+1)!} & \cdots & \frac{(L+1)!}{(L-k+1)!} & \cdots & 1 \\ \frac{(L+2)!}{(L-M+2)!} & \cdots & \frac{(L+2)!}{(L-k+2)!} & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{(L+M)!}{L!} & \cdots & \frac{(L+M)!}{(L+M-k)!} & \cdots & 1 \end{vmatrix} \quad (4.44)$$

Usando as operações entre linhas e a identidade (4.39) obtemos

$$(-1)^k q_k = \frac{(-1)^M \tilde{p} M!}{C_M k} \begin{vmatrix} \frac{(L+1)!}{(L-M+2)!} & \cdots & \frac{(L+1)!}{(L-k+2)!} & \cdots & 1 \\ \frac{(L+2)!}{(L-M+3)!} & \cdots & \frac{(L+2)!}{(L-k+3)!} & \cdots & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \frac{(L+M-1)!}{L!} & \cdots & \frac{(L+M-1)!}{(L+M-k)!} & \cdots & 1 \end{vmatrix} \quad (4.45)$$

Prosseguindo assim, obtemos

$$\begin{aligned}
 (-1)^k q_k &= \pm \frac{\tilde{p}}{C_M k!} \prod_{i=1}^k (M-i+1)! \\
 &= \pm \frac{\tilde{p}}{C_M k!} \frac{(L+1)!}{(L-M+k+1)!} \cdots \frac{(L+1)!}{L!} \cdots 1 \left| \begin{array}{ccc} \frac{(L+1)!}{(L-M+k+1)!} & \cdots & \frac{(L+1)!}{L!} \cdots 1 \\ \frac{(L+2)!}{(L-M+k+2)!} & \cdots & \frac{(L+2)!}{(L+1)!} \cdots 1 \\ \cdots & \cdots & \cdots \\ \frac{(L+M-k)!}{L!} & \cdots & \frac{(L+M-k)!}{(L+M-k-1)!} \cdots 1 \end{array} \right| = \\
 &= \pm \frac{\tilde{p}}{C_M k!} \frac{(L+M-k)!}{L!} \prod_{i=1}^k (M-i+1)! \left| \begin{array}{ccc} \frac{L!}{(L-M+k+1)!} & \cdots & 1 \\ \frac{(L+1)!}{(L-M+k+2)!} & \cdots & 1 \\ \cdots & \cdots & \cdots \\ \frac{(L+M-k-1)!}{L!} & \cdots & 1 \end{array} \right| = \\
 &= \pm \frac{1}{C_M} \left\{ \prod_{i=1}^M \frac{1}{(L+1)!} \right\} \frac{(L+M-k)!}{L! k!} \left\{ \prod_{i=1}^k (M-i+1)! \right\} \prod_{i=1}^{M-k-1} i! = \\
 &= \pm \frac{1}{C_M} \frac{(L+M-k)!}{L! k! (M-k)!} \prod_{i=1}^M \frac{i!}{(L+i)!} \quad (4.46)
 \end{aligned}$$

O sinal do lado direito de (4.46) é o mesmo de (4.41), porque os determinantes (4.36) e (4.42) têm a mesma dimensão, e sofrem as mesmas operações. Logo

$$(-1)^k q_k = (-1)^{M(M-1)/2} \frac{1}{C_M} \frac{(L+M-k)!}{L! k! (M-k)!} \prod_{i=1}^M \frac{i!}{(L+i)!} \quad (4.47)$$

e observando que para  $k = 0$ , temos exatamente (4.41). Portanto

$$q_k = (-1)^k \frac{C_M}{C_M} \frac{(L+M-k)!}{(L+M)!} \frac{M!}{(M-k)!} \frac{1}{k!} \quad (4.48)$$

e

$$Q_M(z) = \sum_{k=0}^M \frac{(L+M-k)!}{(L+M)!} \frac{M!}{(M-k)!} \frac{(-z)^k}{k!}. \quad (4.49)$$

Da mesma forma chegamos a

$$P_L(z) = \sum_{k=0}^L \frac{(L+M-k)!}{(L+M)!} \frac{L!}{(L-k)!} \frac{z^k}{k!}. \quad (4.50)$$

Portanto a  $[L/M]$  aproximação de Padé para  $\exp(z)$  é

$$[L/M] = \frac{\sum_{k=0}^L \frac{(L+M-k)! L! z^k}{(L+M)! (L-k)! k!}}{\sum_{k=0}^M \frac{(L+M-k)! M! (-z)^k}{(L+M)! (M-k)! k!}}. \quad (4.51)$$

ou seja,

$$[L/M] = \frac{P_L(z)}{Q_M(z)}. \quad (4.52)$$

Segue algumas aproximações para a função  $e^z$ , disposta na tabela de Padé (Tabela 4.2).

M \ L	0	1	2
0	$\frac{1}{1}$	$\frac{1+x}{1}$	$\frac{2+2x+x^2}{2}$
1	$\frac{1}{1-x}$	$\frac{2+x}{2-x}$	$\frac{6+4x+x^2}{6-2x}$
2	$\frac{2}{2-2x+x^2}$	$\frac{6+2x}{6-4x+x^2}$	$\frac{12+6x+x^2}{12-6x+x^2}$
3	$\frac{6}{6-6x^2+3x^2-x^3}$	$\frac{24+6x}{24-18x+6x^2-x^3}$	$\frac{60+24x+3x^2}{60-36x+9x^2-x^3}$

Tabela 4.2- Aproximações para  $e^x$ .

Notando que:

$$P_L(z) = \frac{1}{(L+M)!} \int_0^{\infty} t^M (t+z)^L e^{-t} dt, \quad (4.53)$$

pois

$$\begin{aligned} \frac{1}{(L+M)!} \int_0^{\infty} t^M (t+z)^L e^{-t} dt &= \frac{1}{(L+M)!} \sum_{i=0}^L \binom{L}{i} z^i \int_0^{\infty} t^{M+L-i} e^{-t} dt = \\ &= \frac{1}{(L+M)!} \sum_{i=0}^L \frac{L!}{i!(L-i)!} z^i \int_0^{\infty} t^{M+L-i} e^{-t} dt, \end{aligned} \quad (4.54)$$

tomando  $u = t^{M+L-i}$  temos  $du = (M+L-i)t^{M+L-i-1} dt$ , e  $dv = e^{-t} dt$  donde  $v = -e^{-t}$ . Assim a integral fica

$$\int_0^{\infty} t^{M+L-i} e^{-t} dt = \lim_{r \rightarrow \infty} \left[ -t^{M+L-i} e^{-t} \right]_0^r + (L+M-i) \int_0^{\infty} t^{M+L-i-1} e^{-t} dt =$$

$$\begin{aligned}
&= (M+L-i) \int_0^{\infty} t^{M+L-i-1} e^{-t} dt = \dots \\
&= (M+L-i)!, \tag{4.55}
\end{aligned}$$

portanto

$$\frac{1}{(L+M)!} \int_0^{\infty} t^M (t+z)^L e^{-t} dt = \sum_{i=0}^L \frac{(L+M-i)! L! z^i}{(L+M)! i! (L-i)!} \tag{4.56}$$

Da mesma forma

$$Q_M(z) = \frac{1}{(L+M)!} \int_0^{\infty} (t-z)^M t^L e^{-t} dt \tag{4.57}$$

Podemos então escrever a seguinte igualdade

$$\begin{aligned}
e^z Q_M(z) - P_L(z) &= \frac{1}{(L+M)!} \left\{ \int_0^{\infty} (t-z)^M t^L e^{-t+z} dt - \int_0^{\infty} t^M (t+z)^L e^t dt \right\} = \\
&= \frac{1}{(L+M)!} \int_0^z t^L (t-z)^M e^{-t} dt = \\
&= (-1)^L \frac{z^{L+M+1}}{(L+M)!} \int_0^1 \mu^L (1-\mu)^M e^{(1-\mu)z} d\mu \tag{4.58}
\end{aligned}$$

e assim



$$e^A = \frac{P_L(A)}{Q_M(A)} + \frac{(-1)^L}{(L+M)!} A^{L+M+1} Q_M^{-1}(A) \int_0^1 \mu^L (1-\mu)^M e^{A(1-\mu)} d\mu, \quad (4.59)$$

de onde podemos concluir que o erro na aproximação de Padé depende da norma de  $A$ . Mas este problema pode ser contornado usando a igualdade  $e^A = (e^{A/m})^m$ , a idéia é dividir  $A$  por  $m$  tal

que  $\frac{P_L(A/m)}{Q_M(A/m)}$  seja uma aproximação apropriada para  $e^{A/m}$  e depois

obter  $e^A$ . Observe o teorema.

*Teorema 7.* Se  $\frac{\|A\|}{2j} \leq \frac{1}{2}$  [1], então

$$\left[ \frac{P_L(A/2j)}{Q_M} \right]^{2j} = e^{A+E} \quad (4.60)$$

onde

$$\frac{\|E\|}{\|A\|} \leq 8 \left( \frac{\|A\|}{2j} \right)^{L+M} \frac{L!M!}{(L+M)!(L+M+1)!} \leq \left( \frac{1}{2} \right)^{L+M-3} \frac{L!M!}{(L+M)!(L+M+1)!} \quad (4.61)$$

Prova: Notemos primeiro que se  $\|A\| \leq \frac{1}{2}$  então  $\frac{P_L}{Q_M}(A) = e^{A+F}$ , onde

$$\|F\| \leq 8 \|A\|^{L+M+1} \frac{L!M!}{(L+M)!(L+M+1)!} \quad (4.62)$$

Para  $L$  e  $M$  admissíveis consideremos a igualdade

$$\frac{P_L(A)}{Q_M(A)} = e^A + \frac{(-1)^M}{(L+M)!} A^{L+M+1} Q_M^{-1}(A) \int_0^1 e^{(1-\mu)A} \mu^L (1-\mu)^M d\mu, \quad (4.63)$$

e assim

$$e^{-A} \frac{P_L(A)}{Q_M(A)} = I + H \quad (4.64)$$

onde<sup>1</sup>

$$\begin{aligned} \|H\| &\leq \frac{1}{(L+M)!} \|A\|^{L+M+1} \frac{L+M}{L} \int_0^1 e^{\mu L} (1-\mu)^M d\mu \leq \\ &\leq 4 \|A\|^{L+M+1} \frac{L!M!}{(L+M)!(L+M+1)!} \end{aligned} \quad (4.65)$$

Agora,  $\|A\| \leq \frac{1}{2}$  e assim  $\|H\| \leq 1$  donde

$$\begin{aligned} \|\log(I+H)\| &= \left\| \sum_{k=1}^{\infty} (-1)^{k+1} \frac{H^k}{k} \right\| \leq \sum_{k=1}^{\infty} \frac{\|H\|^k}{k} \leq \|H\| \sum_{k=0}^{\infty} \|H\|^k = \frac{\|H\|}{1-\|H\|} \leq \\ &\leq 8 \|A\|^{L+M+1} \frac{L!M!}{(L+M)!(L+M+1)!} \end{aligned} \quad (4.66)$$

Colocando  $F = \log(I+H)$ , resulta que  $e^{-A} \frac{P_L(A)}{Q_M(A)} = I + H = e^F$ , e assim

$$\frac{P_L(A)}{Q_M(A)} = e^A e^F = e^{A+F}, \text{ já que } F \text{ e } A \text{ comutam.}$$

$$(1) \quad Q_M(A) = I + G, \text{ com } G = \sum_{j=1}^M \frac{(L+M-j)!M!}{(L+M)!(M-j)!} \frac{(-A)^j}{j!},$$

$$\text{agora } \frac{(L+M-j)!M!}{(L+M)!(M-j)!} \leq \left[ \frac{M}{L+M} \right]^j, \text{ donde } \|G\| \leq \frac{M}{L+M} \text{ e assim}$$

$$\|Q_M^{-1}\| \leq \|(I+G)^{-1}\| \leq \frac{1}{(1-\|G\|)} \leq \frac{L+M}{L}.$$

Portanto se  $E=2^j F$ , vem

$$\left[ \frac{P_L(A/2^j)}{Q_M} \right]^{2^j} = \left[ e^{A/2^j + F} \right]^{2^j} = e^{A+E} \quad \# \quad (4.67)$$

Usando a aproximação diagonal fica

- Se  $\frac{\|A\|}{2^j} \leq \frac{1}{2}$  então

$$\left[ \frac{P_M(A/2^j)}{Q_M} \right]^{2^j} = e^{A+E}, \quad (4.68)$$

onde

$$\frac{\|E\|}{\|A\|} \leq 8 \left( \frac{\|A\|}{2^j} \right)^{2M} \frac{(M!)^2}{(2M!(2M+1)!)} \leq \left( \frac{1}{2} \right)^{2M-3} \frac{(M!)^2}{(2M!(2M+1)!)} \quad (4.69)$$

Para a aproximação usando série de Taylor, teremos

- Se  $\frac{\|A\|}{2^j} \leq \frac{1}{2}$  então

$$\left[ T_k(A/2^j) \right]^{2^j} = e^{A+E}, \quad (4.70)$$

onde

$$\frac{\|E\|}{\|A\|} \leq 8 \left( \frac{\|A\|}{2^j} \right)^{2k} \frac{1}{k+1} \leq \left( \frac{1}{2} \right)^{k-3} \frac{1}{k+1} e \quad T_k(A) = \sum_{i=0}^k \frac{A^i}{i!} \quad (4.71)$$

## 5- TESTES NUMÉRICOS

### 5.1- Introdução

Nosso objetivo agora é comparar, na prática, o que acontece com as aproximações apresentadas anteriormente. Neste sentido, apresentaremos os resultados obtidos, com auxílio de um computador CONVEX-C210 (UFSC), e faremos uma comparação entre os erros absoluto ou relativo com os limitantes de erros usados. Notaremos que esses limitantes nem sempre fornecem um critério ideal de parada na execução do método.

Para isso, construímos matrizes  $A$ , partindo da decomposição  $A = PDP^{-1}$ . Primeiramente, obtemos a matriz  $P$  randomicamente, usando a subrotina  $RANVCISSED(n,v)$ , disponível na biblioteca VECLIB do CONVEX-C210, [24]. Nesta subrotina fornecemos a semente  $issed(1234)$ , declarada como inteiro de dupla precisão, e  $n$ . Obtemos como retorno o vetor  $v$ , declarado como real dupla precisão, contendo  $n$  números pseudo-aleatórios no intervalo  $[0,1)$ .

Obtida, então, a matriz  $P$ , calculamos a matriz  $P^{-1}$ , usando também uma subrotina,  $DGEDI$ , disponível na biblioteca LINPACK [24], incluída no VECLIB do CONVEX-C210.

E finalmente, fornecendo a matriz diagonal  $D$ , obtemos a matriz  $A$  que será usada.

Assim,  $e^A = Pe^{DP^{-1}}$ . Essa será a "real" exponencial de  $A$  para as nossas comparações.

### 5.2- Aproximações por Taylor

Para a aproximação pela série de Taylor, usamos como critério de parada:

1) o limitante do erro absoluto, obtido de acordo com o teorema 3 do capítulo 4,

$$n \left\| \frac{A^{k+1}}{(k+1)!} \right\|_{\infty} e^{\|A\|_{\infty}} \quad (5.1)$$

para uma aproximação de ordem  $k$ , onde  $n$  é a ordem da matriz

2) a norma do incremento de cada passo,  $\|T_{k+1} - T_k\|_{\infty}$ .

Observação: Em [1], [5], os limitantes são calculados com a 2-norma. Mas, eles podem ser estendidos a norma- $\infty$ .

Exemplo 1: Usando uma matriz gerada randomicamente de ordem 3.

a) Com autovalores distintos  $\langle 1, 2, 3 \rangle$ . ( $\|A\|_{\infty} = 5.83$ )

n° de iterações	$\left\  \frac{A^{k+1}}{(k+1)!} \right\ _{\infty} e^{\ A\ _{\infty}}$	$\ T_{k+1} - T_k\ _{\infty}$	erro absoluto
11	101.88	9.89E-2	3.59E-2
15	0.94	3.39E-4	8.33E-5
17	1.31E-2	1.27E-5	2.69E-6
20	6.10E-5	5.92E-8	1.03E-8

Tabela 5.1

b) Com autovalores distintos  $\langle 5, 5.1, 5.2 \rangle$ . ( $\|A\|_{\infty} = 5$ )

n° de iterações	$\left\  \frac{A^{k+1}}{(k+1)!} \right\ _{\infty} e^{\ A\ _{\infty}}$	$\ T_{k+1} - T_k\ _{\infty}$	erro absoluto
11	1200.43	2.69	2.03
15	31.24	7.01E-2	3.36E-2
17	3.25	7.30E-3	2.95E-3
20	7.00E-2	1.57E-4	5.13E-5
25	4.29E-5	9.63E-8	2.38E-8

Tabela 5.2

c) Com autovalores não distintos (5, 5, 5.4). (  $\|A\|_{\infty} = 6.53$  )

n° de iterações	$\left\  \frac{A^{k+1}}{(k+1)!} \right\ _{\infty} n e \ A\ _{\infty}$	$\ T_{k+1} - T_k\ _{\infty}$	erro absoluto
11	34445.04	16.66	16.28
15	1595.08	0.77	0.44
17	211.87	0.10	4.84E-2
20	1.75	8.51E-4	2.97E-4
30	1.17E-5	5.66E-9	1.30E-9

Tabela 5.3

Exemplo 2: Usando uma matriz gerada randomicamente de ordem 5.

a) Com autovalores distintos (1, 1.5, 2, 2.5, 3).

(  $\|A\|_{\infty} = 21.09$  )

n° de iterações	$\left\  \frac{A^{k+1}}{(k+1)!} \right\ _{\infty} n e \ A\ _{\infty}$	$\ T_{k+1} - T_k\ _{\infty}$	erro absoluto
11	2600747051.07	0.35	0.13
15	8918791.74	1.22E-3	3.01E-4
17	335428.80	4.61E-5	9.76E-6
20	1560.99	2.14E-7	3.75E-8
25	7.44E-2	1.02E-11	3.84E-12
30	1.27E-6	1.74E-16	2.49E-12

Tabela 5.4

b) Com autovalores distintos próximos (3, 3.00001, 3.000002, 3.000003, 3.0000222). (  $\|A\|_{\infty} = 3.00$  )

n° de iterações	$\left\  \frac{A^{k+1}}{(k+1)!} \right\ _{\infty} n e \ A\ _{\infty}$	$\ T_{k+1} - T_k\ _{\infty}$	erro absoluto
1	1.63	1.62E-2	5.87E-3
15	5.51E-3	5.49E-5	1.34E-5
17	2.07E-4	2.06E-6	4.35E-7
20	9.61E-7	9.57E-9	1.67E-9
25	4.58E-11	4.56E-13	3.49E-13

Tabela 5.5

c) Com autovalores não distintos (4, 4, 4, 3, 3).

(  $\|A\|_{\infty} = 25.54$  )

n° de iterações	$\left\  \frac{A^{k+1}}{(k+1)!} \right\ _{\infty} n e \ A\ _{\infty}$	$\ T_{k+1} - T_k\ _{\infty}$	erro absoluto
11	382337950722.90	6.16	3.36
15	42326036197.87	6.82E-2	2.42E-2
17	2843120593.65	4.58E-3	1.38E-3
20	31471796.64	5.07E-5	1.25E-5
36	1.59E-6	2.57E-18	5.97E-12

Tabela 5.6

Exemplo 3: Usando uma matriz gerada randomicamente de ordem 20.

a) Com autovalores distintos (0.5, 1, 1.5, 2, 2.5, 2.7, 3, 3.2, 3.4, 3.5, 4, 4.5, 4.8, 5, 5.1, 5.2, 5.4, 5.5, 5.9, 6.5).

(  $\|A\|_{\infty} = 90.32$  )

n° de iterações	$\left\  \frac{A^{k+1}}{(k+1)!} \right\ _{\infty} n e \ A\ _{\infty}$	$\ T_{k+1} - T_k\ _{\infty}$	erro absoluto
25	1.89E 37	5.61E-4	1.94E-4
30	1.55E 34	4.62E-7	1.26E-7
35	5.45E 30	1.61E-10	1.83E-10
75	1.61E-6	4.79E-47	1.98E-12

Tabela 5.7

b) Com autovalores distintos próximos ( 2, 2.0000001, 2.0001, 2.00059, 2.000000101, 2.0000012, 2.000310001, 2.000030002, 2.00000003, 2.000000132, 2.000013, 2.0000300021, 2.000100020003, 2.001, 2,015, 2.015678, 2.000123, 2.01000012, 2.02000101, 2.000300020001). (  $\|A\|_{\infty} = 2.46$  )

n° de iterações	$\left\  \frac{A^{k+1}}{(k+1)!} \right\ _{\infty} n e^{\ A\ _{\infty}}$	$\ T_{k+1} - T_k\ _{\infty}$	erro absoluto
11	0.22	9.59E-4	2.27E-4
15	1.28E-3	5.47E-6	9.65E-7
17	3.62E-6	1.54E-8	2.16E-9
20	5.80E-9	2.47E-11	2.99E-12

Tabela 5.8

c) Com autovalores não distintos (1.5, 1.5, 1.5, 1.5, 1.5, 1.5, 2.3, 2.3, 2.3, 2.3, 2.3, 2.3, 2.3, 2.3, 2.5, 2.5, 2.5, 2.5, 2.5, 2.5). (  $\|A\|_{\infty} = 27.45$  )

n° de iterações	$\left\  \frac{A^{k+1}}{(k+1)!} \right\ _{\infty} n e^{\ A\ _{\infty}}$	$\ T_{k+1} - T_k\ _{\infty}$	erro absoluto
11	992368836532.56	5.93E-2	1.72E-2
17	44730734.79	2.67E-6	4.59E-7
20	122546.35	7.32E-9	1.04E-9
31	1.39E-6	8.32E-20	1.13E-12

Tabela 5.9

### 5.3- Taylor com balanceamento (Scaling)

Como podemos observar, a magnitude da norma influencia diretamente na rapidez da convergência. Assim usamos a igualdade  $e^A = (e^{A/m})^m$  para contornar este problema como foi visto no final da capítulo 4.



Se  $\frac{\|A\|}{2J} < 0.5$  então

$$\frac{\|e^A - [T_k(A/2J)]^2\|_\infty}{\|e^A\|_\infty} \leq \frac{1}{k+1} \|A\|_\infty (0.5)^{k-3} e^{\|A\|_\infty} (0.5)^{k-3} \frac{1}{1/(k+1)} \quad (5.2)$$

Exemplo 4: Usando os mesmos dados do exemplo 1, obtemos:

a)

n° de iterações	limitante do erro relativo	$\ T_{k+1} - T_k\ _\infty$	erro relativo	erro absoluto
5	1.046	2.46E-4	4.34E-5	3.09E-3
7	0.110	3.36E-7	3.73E-8	2.66E-6
10	9.20E-3	4.74E-12	3.46E-13	2.47E-11
15	1.90E-4	4.70E-21	2.87E-15	2.05E-13
19	9.37E-6	7.94E-29	2.87E-15	2.05E-13

Tabela 5.10

b)

n° de iterações	limitante do erro relativo	$\ T_{k+1} - T_k\ _\infty$	erro relativo	erro absoluto
5	0.82	3.97E-4	3.06E-4	4.55E-2
7	9.76E-2	1.29E-6	7.03E-7	1.04E-4
8	4.06E-2	5.77E-8	2.73E-8	4.06E-6
10	7.87E-3	7.84E-11	2.95E-11	4.38E-9
15	1.62E-4	9.73E-19	5.07E-15	7.53E-13
19	8.03E-6	2.26E-25	5.07E-15	7.53E-13

Tabela 5.11

c)

n° de iterações	limitante do erro relativo	$\ T_{k+1} - T_k\ _{\infty}$	erro relativo	erro absoluto
5	1.25	9.47E-4	6.49E-4	2.78E-1
7	0.13	4.20E-6	2.02E-6	8.68E-4
9	2.32E-2	9.65E-9	3.56E-9	1.52E-6
12	2.13E-3	5.05E-13	1.16E-13	4.99E-11
15	2.12E-4	1.07E-17	5.17E-15	2.21E-12
20	4.98E-6	4.06E-26	5.17E-15	2.21E-12

Tabela 5.12

Exemplo 5: Usando os mesmos dados do exemplo 2, obtemos

a)

n° de iterações	limitante do erro relativo	$\ T_{k+1} - T_k\ _{\infty}$	erro relativo	erro absoluto
5	17.39	3.59E-6	1.77E-7	4.95E-5
6	2.11	3.65E-8	1.40E-9	3.93E-7
10	3.40E-2	6.55E-17	1.10E-14	3.09E-12
21	7.66E-6	2.42E-44	1.10E-14	3.09E-12

Tabela 5.13

b)

n° de iterações	limitante do erro relativo	$\ T_{k+1} - T_k\ _{\infty}$	erro relativo	erro absoluto
5	0.40	8.24E-4	3.62E-4	7.28E-3
7	5.65E-2	3.86E-6	1.19E-5	2.39E-5
10	4.71E-3	4.04E-10	8.63E-10	1.73E-9
12	9.77E-4	5.17E-13	1.08E-13	2.18E-12
15	9.76E-5	1.24E-17	1.68E-14	3.39E-13
19	4.81E-6	3.36E-24	1.68E-14	3.39E-13

Tabela 5.14

c)

n° de iterações	limitante do erro relativo	$\ T_{k+1} - T_k\ _{\infty}$	erro relativo	erro absoluto
5	32.85	9.99E-6	6.81E-7	5.44E-4
6	3.08	1.38E-7	7.25E-9	5.78E-6
8	0.24	1.45E-11	5.27E-13	4.21E-10
10	4.15E-2	8.39E-16	7.49E-15	6.34E-12
21	9.28E-6	7.64E-42	7.49E-15	6.34E-12

Tabela 5.15

Exemplo 6: Usando os mesmos dados do exemplo 3, obtemos

a)

n° de iterações	limitante do erro relativo	$\ T_{k+1} - T_k\ _{\infty}$	erro relativo	erro absoluto
5	75592.41	2.04E-7	2.24E-8	1.42E-4
7	8.09	3.59E-12	3.83E-13	2.43E-9
23	7.48E-6	7.60E-56	7.49E-14	4.75E-10

Tabela 5.16

b)

n° de iterações	limitante do erro relativo	$\ T_{k+1} - T_k\ _{\infty}$	erro relativo	erro absoluto
5	0.31	3.15E-4	9.41E-5	1.01E-3
7	4.59E-2	8.79E-7	1.64E-7	1.78E-6
8	1.96E-2	3.21E-8	5.53E-9	5.99E-8
10	3.86E-3	3.31E-11	4.40E-12	4.77E-11
15	8.01E-5	1.88E-19	2.07E-14	2.24E-13
18	8.34E-6	8.51E-25	2.07E-14	2.24E-13

Tabela 5.17

c)

n° de iterações	limitante do erro relativo	$\ T_{k+1} - T_k\ _{\infty}$	erro relativo	erro absoluto
5	42.73	2.08E-6	6.51E-8	1.27E-5
6	3.59	1.68E-8	4.21E-10	8.26E-8
10	4.47E-2	1.29E-17	1.51E-14	2.96E-11
21	9.97E-6	6.92E-46	1.51E-14	2.96E-11

Tabela 5.18

## 5.4- Padé

Para as aproximações de Padé, o cálculo de um limitante para o erro absoluto ou relativo leva em conta informações não somente sobre os autovalores, mas também sobre  $v(A)$ , onde  $v(A) = \min_T k(T)$  e  $A = TJT^{-1}$  (Forma de Jordan). Assim se  $A$  é uma matriz qualquer, o cálculo de um limitante depende de  $v(A)$  e o cálculo de  $v(A)$  não é obvio. Agora se  $A$  é simétrica e se usarmos a norma definida por  $\|A\| = \sum_{i,j} |a_{ij}|$ , Wragg & Davies [12] mostram que:

$$\left\| \exp(A) - \frac{P_q(A)}{Q_q(A)} \right\| \leq \frac{n (q!)^2 \Delta^{2q+1} \exp(\Delta)}{(2q)!(2q+1)!Q_q(\Delta)}, \quad (5.3)$$

desde que os autovalores de  $A$  estejam no interior de  $[-\Delta, \Delta]$ .

Exemplo 7: Usando a matriz  $A = \begin{pmatrix} -7 & 0 & 6 \\ 0 & 5 & 0 \\ 6 & 0 & 2 \end{pmatrix}$  onde  $\|A\|_{\infty} = 13$ .

n° de iterações	usando o limitante (5.3)	erro absoluto
6	237939983.06	0.14
10	55890810.41	4.58E-8
20	1.53E-6	3.89E-11

Tabela 5.19

Exemplo 8: Usando a matriz  $A = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & 2 \\ 4 & 2 & 3 \end{pmatrix}$  onde  $\|A\|_{\infty} = 9$ .

n° de iterações	usando o limitante (5.3)	erro absoluto
10	42.56	1.72E-2
12	5.76E-2	1.40E-5
14	4.20E-5	5.87E-9
15	9.23E-7	3.56E-10

Tabela 5.20

Exemplo 9: Usando a matriz  $A = \begin{pmatrix} 0.5 & -1 & 0.3 \\ -1 & 1 & 0.2 \\ 0.3 & 0.2 & 0.1 \end{pmatrix}$  onde  $\|A\|_{\infty} = 2.2$ .

n° de iterações	usando o limitante (5.3)	erro absoluto
3	0.22	7.63E-3
5	5.05E-5	7.34E-7
7	2.59E-9	1.62E-11

Tabela 5.21

### 5.5- Padé com balanceamento (Scaling)

Utilizamos a igualdade  $e^A = (e^{A/m})^m$ , com a finalidade de diminuir a norma da matriz que aparece nos cálculos das aproximações de Padé. Obtemos desta forma, uma rapidez significativa na convergência dos resultados. Usamos neste caso o teorema 7 do capítulo 4 para definirmos um limitante do erro relativo:

$$\text{se } \frac{\|A\|_\infty}{2^j} < \frac{1}{2} \text{ então } \frac{\|e^A - \frac{P_q(A)}{Q_q(A)}\|_\infty}{\|e^A\|} \leq \|E\|_\infty e^{\|E\|_\infty} \quad (5.4)$$

onde

$$\|E\|_\infty \leq (0.5)^{2q-3} \frac{(q!)^2 \|A\|_\infty}{(2q)!(2q+1)!} \quad (5.5)$$

Exemplo 10: Usando os mesmos dados do exemplo 1, obtemos

a)

n° de iterações	usando o limitante (5.4)	erro relativo	erro absoluto
3	7.24E-6	2.20E-9	1.57E-7
4	7.18E-9	3.08E-13	2.20E-11

Tabela 5.22

b)

n° de iterações	usando o limitante (5.4)	erro relativo	erro absoluto
3	6.20E-6	4.64E-8	6.89E-6
4	6.15E-9	1.79E-11	2.66E-9

Tabela 5.23

c)

n° de iterações	usando o limitante (5.4)	erro relativo	erro absoluto
3	8.10E-6	1.39E-7	5.96E-5
4	8.04E-9	7.03E-11	3.01E-8

Tabela 5.24

Exemplo 11: Usando os dados do exemplo 2, obtemos

a)

n° de iterações	usando o limitante (5.4)	erro relativo	erro absoluto
3	2.61E-5	4.83E-13	1.35E-10
4	2.59E-8	2.89E-15	8.09E-13

Tabela 5.25

b)

n° de iterações	usando o limitante (5.4)	erro relativo	erro absoluto
3	3.72E-6	8.32E-8	1.67E-6
4	3.69E-9	4.64E-11	9.32E-10

Tabela 5.26

c)

n° de iterações	usando o limitante (5.4)	erro relativo	erro absoluto
3	3.16E-5	3.48E-12	2.77E-9
4	3.14E-8	8.29E-15	6.62E-12

Tabela 5.27

Exemplo 12: Usando os dados do exemplo 3, obtemos

a)

n° de iterações	usando o limitante (5.4)	erro relativo	erro absoluto
3	1.12E-4	2.03E-13	1.29E-9
4	1.11E-7	1.78E-13	1.13E-9

Tabela 5.28

b)

n° de iterações	usando o limitante (5.4)	erro relativo	erro absoluto
3	3.05E-6	1.04E-8	1.13E-7
4	3.02E-9	3.01E-12	3.26E-11

Tabela 5.29

c)

n° de iterações	usando o limitante (5.4)	erro relativo	erro absoluto
3	3.40E-5	1.36E-13	2.68E-11
4	3.37E-8	4.20E-14	8.25E-12

Tabela 5.30

### 5.6- Runge Kutta de ordem 4

Aplicamos também o método de Runge-Kutta de quarta ordem [9] para calcular a exponencial de uma matriz, porém, como podemos observar nas tabelas 5.31, 5.32 e 5.33, os resultados obtidos não foram animadores. Para termos de comparação, usamos um limitante de erro absoluto [1,19].

Exemplo 13: Usando os mesmos dados de exemplo 1, obtemos

a)

n° de iterações	limitante $\ A\ _{\infty}^5 e^{\ A\ _{\infty} h^5}$	erro absoluto
20	0.57	1.31E-3
30	8.13E-2	2.70E-4
60	2.75E-3	1.76E-5
100	2.21E-4	2.32E-6

Tabela 5.31



b)

n° de iterações	limitante $\ A\ _{\infty}^5 e^{\ A\ _{\infty} h^5}$	erro absoluto
20	0.11	1.96E-2
30	1.62E-2	4.15E-3
60	5.49E-4	2.78E-4
100	4.41E-5	3.71E-5

Tabela 5.32

c)

n° de iterações	limitante $\ A\ _{\infty}^5 e^{\ A\ _{\infty} h^5}$	erro absoluto
20	2.01	1.21E-1
30	2.86E-1	2.59E-2
60	9.72E-3	1.75E-3
100	7.81E-4	2.35E-4

Tabela 5.33

### 5.7- Comentários

Podemos observar nas tabelas 5.1 a 5.9, 5.19 a 5.21 e 5.31 a 5.33 apresentadas acima, os limitantes dos erros diferem dos erros reais principalmente quando a norma da matriz é um número grande. Este problema foi contornado fazendo-se uso da igualdade  $e^A = (e^{A/m})^m$ , onde dividimos a matriz A por uma potência de 2, com a finalidade de diminuir a norma da matriz resultante. Escolhemos potência j tal que  $\| \frac{A}{2^j} \| < 0.5$ , o que foi utilizado na construção das tabelas 5.10 a 5.18 e 5.20 a 5.30.

Observemos que as aproximações racionais de Padé convergem com um número de iterações consideravelmente menor que o simples uso do truncamento da série de Taylor ou do método Runge Kutta, utilizado comumente em equações

diferenciais.

Para o exemplo apresentado na introdução, obtemos para a exponencial de A usando Padé com Scaling e  $t=1$ , a seguinte matriz  $e^{At}$ :

$$\begin{pmatrix} 0.215060185905783 & 0.185179115395620 & 0.797249026691702E-1 \\ 0.185179115395620 & 0.294785088574953 & 0.185179115395620 \\ 0.797249026691702E-1 & 0.185179115395620 & 0.215060185905783 \end{pmatrix}$$

com 5 iterações e limitante do erro relativo igual a  $0.31E-11$ , e para  $t=10$ , obtemos a seguinte matriz  $e^{At}$ :

$$\begin{pmatrix} 0.71433587384836E-3 & 0.10102220234292E-2 & 0.71433381269474E-3 \\ 0.10102220234292E-2 & 0.14286696865431E-2 & 0.10102220234292E-2 \\ 0.71433381269474E-3 & 0.10102220234292E-2 & 0.71433381269474E-3 \end{pmatrix}$$

com 5 iterações e limitante do erro relativo igual a  $0.31E-10$ .

Usando Taylor com Scaling obtemos para  $t=1$ , a seguinte matriz  $e^{At}$ :

$$\begin{pmatrix} 0.215060185905783 & 0.185179115395620 & 0.797249026691703E-1 \\ 0.185179115395620 & 0.294785088574953 & 0.185179115395620 \\ 0.797249026691703E-1 & 0.185179115395620 & 0.215060185905783 \end{pmatrix}$$

com 19 iterações e limitante do erro relativo igual a  $0.64E-5$  e para  $t=10$ , obtemos a seguinte matriz  $e^{At}$ :

$$\begin{pmatrix} 0.71433587384838E-3 & 0.10102220234292E-2 & 0.71433381269477E-3 \\ 0.10102220234292E-2 & 0.14286696865431E-2 & 0.10102220234292E-2 \\ 0.71433381269477E-3 & 0.10102220234292E-2 & 0.71433381269477E-3 \end{pmatrix}$$

com 22 iterações e limitante do erro relativo igual a  $0.69E-5$ .

Para calcular potências de uma matriz,  $B^p$ , onde B é uma matriz  $n \times n$  e p um número natural, utilizamos o algoritmo

abaixo [2],

Seja  $P = \sum_{k=0}^t R_k 2^k$ , onde  $R_t \neq 0$ , a decomposição de  $P$  em fatores de 2.

$$k = 0$$

Enquanto  $R_k \neq 0$

$$B = B^2$$

$$k = k + 1$$

$$F = B$$

Para  $j = k+1, t$

$$B = B^2$$

Se  $R_j \neq 0$

$$F = F * B$$

(obtemos como resultado final  $F = B^P$ ).

## 6- CONCLUSÃO

Como podemos observar, as várias definições apresentadas para função de uma matriz  $f(A)$ , associada à  $f(z)$ , com  $z \in \mathbb{C}$ , são equivalentes a menos de generalidades e satisfazem algumas propriedades importantes.

Lembremos que, na definição de  $f(A)$ , usamos os autovalores de  $A$ , o que dificulta o cálculo de  $f(A)$ . Esta dificuldade ficou bem caracterizada no momento em que analisamos a sensibilidade do problema de definir a exponencial de uma matriz,  $e^A$ . O problema é bem condicionado quando a matriz é normal.

Tanto nas aproximações de Taylor como nas aproximações de Padé para a função  $e^A$  foram calculados limitantes dos erros absolutos ou relativos. Verificamos que a norma da matriz está diretamente relacionada com as aproximações. Assim procuramos contornar este problema usando a igualdade  $(e^{A/m})^m = e^A$ , ou seja, reduzindo a norma da matriz. Com isto as aproximações foram mais eficientes, no sentido de que a rapidez de convergência aumentou consideravelmente.

## 7- REFERÊNCIAS BIBLIOGRÁFICAS

- [1] C. B. MOLER & C. F. VAN LOAN. Nineteen dubious ways to compute the matrix exponential of a matrix: SIAM Review, Vol. 20, n° 4, October 1978.
- [2] G. H. GOLUB & C. F. VAN LOAN. Matrix computations, second edition, 1989, the Johns Hopkins University Press, London.
- [3] F. R. GANTMACHER. The theory of matrices, Vol. 1, Chelsea Publishine Comoany, 1959, New York, N.Y..
- [4] P. P. PETRUSHEV & V. A. POPOV. Rational approximation of real functions. Encyclopedia of mathematics and its applications, Vol. 28, Edited by G. C. rota, Cambridge University Press, 1987, New York, N.Y..
- [5] C. F. VAN LOAN. The sensitivity of the matrix exponential, SIAM Numer. Anal, 14 (1977), pp. 971-981.
- [6] R. F. RINEHART. The equivalence of definitions of a matric function, Amer. Math Monthly, 62 (1955), pp. 395-414.
- [7] W. FAIR & Y. L. LUKE. Pade approximations to the operator exponential, Numer. math, 14 (1970), pp. 379-382.
- [8] T. KATO. Perturbation theory for linear operators, Chap. 9, Springer-Veklag, New York, 1966.
- [9] J. M. ORTEGA & W. G. POOLE JR.. An introduction to numerical methods for differential equations. Pitman Publishing Inc., 1981, Massachusetts.
- [10] W. L. FERRAR. Finite Matrices, Clarendon Press, 1951.
- [11] A. WRAGG & C. DAVIS. Computation of the exponential of a matrix, I: Theoretical Considerations, J. Ins. Maths Applics (1973) 11, 369-375.
- [12] A. WRAGG & C. DAVIS. Computation of the exponential of a matrix, II: Practical Considerations, J. Inst. Maths

- Applics (1975) 15, 273-278.
- [13] A. GONÇALVES. Introdução à Álgebra, Rio de Janeiro, Coleção Projeto Euclides, CNPq, 1979.
- [14] E. L. LIMA. Curso de Análise, Vol. 1 e Vol. 2, Rio de Janeiro, Coleção Projeto Euclides, CNPq, 1976.
- [15] E. L. LIMA. Espaços Métricos, Rio de Janeiro, Coleção Projeto Euclides, CNPq, 1977.
- [16] K. HOFFMAN & R. KUNZE. Algebra Linear, Editora Polígono, São Paulo, 1971.
- [17] L. V. AHLFORS. Complex Analysis, McGraw-Hill Book Company, New York (1953).
- [18] G. STRANG. Linear Algebra and its Applications, Third Edition, Harcourt Brace Jovanovich, 1988.
- [19] A. CARL-ERICK FROBERG. introduction to Numerical Analysis, Wesley Publishing Company, inc., 1966, Massachusetts, Palo Alto, London.
- [20] W. H. PRESS; B. P. FLANNERY; S. A. TEUKOLSKY; W. T. VETTERLING. Numerical Recipes, The Art of Scientific Computing, Cambridge University Press, Cambridge, New York, 1989.
- [21] J. M. SOTOMAYOR T.. Lições de Equações Diferenciais Ordinárias, Rio de Janeiro, Coleção Projeto Euclides, CNPq, 1979.
- [22] J. DONGARRA; J. R. BUNCH; C. B. MOLER; G. W. STEWART. Linpack Users Guide, SIAM Publications, Philadelphia, 1978.
- [23] B. T. SMITH; J. M. BOYLE; Y. IKEBE; V. C. KLEMA; C. B. MOLER. Matrix Eigensystem Routines: EISPACK Guide, 2nd Ed., Springer-Verlang, New York, 1970.
- [24] CONVEX VECLIB, User's Guide, Fifth Edition, Convex Computer Corporation, Richardson, Texas, 1990.

## APÊNDICE

**Teorema 1** : Seja  $A$  uma matriz quadrada. Sejam  $\lambda_1, \dots, \lambda_t$  os autovalores distintos de  $A$ . Seja, para cada  $n \in \mathbb{N}$ ,

$$f_n(z) = \sum_{k=1}^t [\alpha_{k1} + \alpha_{k2}(z-\lambda_k) + \dots + \alpha_{ks_k}(z-\lambda_k)^{s_k-1}] \psi_k(z), \quad (1)$$

onde

$$\alpha_{k1} = \frac{f_n(\lambda_k)}{\psi_k(\lambda_k)},$$

$$\alpha_{k2} = f_n(\lambda_k) \left[ \frac{1}{\psi_k(z)} \right]_{z=\lambda_k}' + f_n'(\lambda_k) \frac{1}{\psi_k(\lambda_k)},$$

$\vdots$

$$\alpha_{kj} = \frac{1}{(j-1)!} \left[ \frac{f_n(z)}{\psi_k(z)} \right]_{z=\lambda_k}^{(j-1)} \quad \text{com } j=1, 2, \dots, s_k, \quad (2)$$

$$k=1, 2, \dots, t, \quad \psi_k(z) = \frac{\psi(z)}{(z-\lambda_k)^{s_k}}, \quad \text{onde } \psi(z) \text{ é o}$$

polinômio minimal da  $A$ .

Então a sequência de matrizes  $f_n(A)$  converge quando  $n \rightarrow \infty$  se, e somente se, a sequência  $f_n(z)$  converge no espectro de  $A$ , isto é, os limites

$$\lim_{n \rightarrow \infty} f_n(A) \quad \text{e} \quad \lim_{n \rightarrow \infty} f_n(\lambda(A)) \quad (3)$$

sempre existem simultaneamente. Além disso, a equação

$$\lim_{n \rightarrow \infty} f_n(\lambda(A)) = f(\lambda(A)) \quad (4)$$

implica que

$$\lim_{n \rightarrow \infty} f_n(A) = f(A) \quad (5)$$

e reciprocamente.

*Prova:* Observemos primeiro que podemos escrever (1) da seguinte forma

$$f_n(z) = \sum_{k=1}^l \left[ f_n(\lambda_k) Z_{k1}(z) + f'_n(\lambda_k) Z_{k2}(z) + \dots + f_n^{(s_k-1)}(\lambda_k) Z_{ks_k}(z) \right], \quad (6)$$

onde

$$\begin{aligned} Z_{k1} &= \frac{1}{\psi_k(\lambda_k)} + (z-\lambda_k) \left( \frac{1}{\psi_k(\lambda_k)} \right)'_{z=\lambda_k} + (z-\lambda_k)^2 \left( \frac{1}{\psi_k(\lambda_k)} \right)''_{z=\lambda_k} + \dots, \\ Z_{k2} &= \frac{(z-\lambda_k)}{\psi_k(\lambda_k)} + (z-\lambda_k)^2 \left( \frac{1}{\psi_k(z)} \right)'_{z=\lambda_k} + \dots, \\ &\vdots \\ Z_{ks_k} &= \frac{(z-\lambda_k)^{s_k-1}}{\psi_k(\lambda_k)} + (z-\lambda_k)^{s_k} \left( \frac{1}{\psi_k(z)} \right)'_{z=\lambda_k} + \dots \end{aligned} \quad (7)$$

onde os  $Z_{kj}$  são polinômios cujos os coeficientes são construídos utilizando os  $\psi_k$  e os autovalores, logo independente de  $f_n$ .

1) Se os valores de  $f_n(z)$  convergem no espectro de  $A$  quando  $n \rightarrow \infty$  para um valor limite, então da fórmula anterior obtemos:



$$f_n(A) = \sum_{k=1}^t \left[ f_n(\lambda_k) Z_{k1}(A) + f'_n(\lambda_k) Z_{k2}(A) + \dots + f_n^{(s_k-1)}(\lambda_k) Z_{ks_k}(A) \right], \quad (8)$$

Segue que o limite  $\lim_{n \rightarrow \infty} f_n(A)$  existe, pois as matrizes  $Z_{kj}(A)$  ( $k=1, 2, \dots, t; j=1, 2, \dots, s_k$ ) são completamente determinadas e não dependem de  $f_n(z)$ .  $f_n$  é avaliado somente sobre o espectro e os  $Z_{kj}$  são linearmente independentes. Assim de (4) deduzimos (5).

2) Suponha, reciprocamente, que  $\lim_{n \rightarrow \infty} f_n(A)$  existe. Sendo as matrizes  $Z$  linearmente independentes, podemos expressar de (8), os  $m$  ( $m = \sum s_k$ ) valores de  $f_n(z)$  sobre o espectro de  $A$  (como uma forma linear) pelos  $m$  elementos da matriz  $f_n(A)$ . Assim a existência do limite  $\lim_{n \rightarrow \infty} f_n(\lambda(A))$  segue. #

Observação: Para maiores detalhes consulte [3].

*Teorema de Abel:* Seja  $\sum a_n x^n$  uma série de potências cujo raio de convergência  $r$  é finito e positivo.

Se  $\sum a_n r^n$  converge, então  $\sum a_n x^n$  converge uniformemente no intervalo  $[0, r]$ . Em particular,

$$\lim_{x \rightarrow r^-} (\sum a_n x^n) = \sum a_n r^n. \quad (9)$$

Prova : Ver [14]

*Teorema de Hensel:* A série de potências

$$f(A) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(z_0)(A-z_0I)^k \quad (10)$$

converge se, e somente se, toda raiz característica  $\lambda_i$  de  $A$  permanece dentro ou sobre o círculo de convergência de  $f(z)$ , e para toda raiz  $\lambda_i$  de multiplicidade  $s_i$  do polinômio minimal de  $A$ , a  $(s_i-1)$ -ésima derivada da série acima converge em  $\lambda_i$  para  $f^{(s_i-1)}(\lambda_i)$ .

*Prova:* Seja  $\sum \frac{1}{k!} f^{(k)}(z_0)(A-z_0I)^k$  convergente. Se  $A = PJP^{-1}$  então  $J = P^{-1}AP$  e a hipótese é equivalente a

$$P^{-1} \sum \frac{1}{k!} f^{(k)}(z_0)(A-z_0I)^k P \quad (11)$$

convergir, isto é,

$$\sum \frac{1}{k!} f^{(k)}(z_0)(J-z_0I)^k. \quad (12)$$

Agora  $J = \text{diag}(J_1, J_2, \dots, J_v)$ , cada  $J_i$  associado a  $\lambda_i$ , com

$$J_i = \begin{bmatrix} J_{i1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_{ir_i} \end{bmatrix}, \text{ onde } J_{ir_i} = \begin{bmatrix} \lambda_i & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \dots & \lambda_i & 1 \\ 0 & \dots & \dots & \lambda_i \end{bmatrix}. \quad (13)$$

Para qualquer  $i$  existe pelo menos um bloco de ordem  $r_i$ , onde  $r_i$  é o grau de  $(\lambda - \lambda_i)$  no polinômio minimal, e é a maior ordem possível entre os blocos  $J_{ij}$ . O bloco correspondente a esse em

$$\sum \frac{1}{k!} f^{(k)}(z_0)(J-z_0I)^k \quad (14)$$

é

$$\begin{pmatrix} f(\lambda_i) & f'(\lambda_i) & \dots & \frac{1}{(r_i-1)!} f^{(r_i-1)}(\lambda_i) \\ \vdots & & & \\ 0 & \dots & & f(\lambda_i) \end{pmatrix}, \quad (15)$$

onde

$$f(\lambda_i) = \sum \frac{1}{k!} f^{(k)}(z_0)(\lambda_i - z_0)^k$$

$f'(\lambda_i)$  = a derivada da série acima avaliada em  $\lambda_i$

.....

$f^{(r_i-1)}(\lambda_i)$  = a  $(r_i-1)$ -ésima derivada da série acima avaliada em  $\lambda_i$ .

Assim,  $\lambda_i$  está dentro ou sobre o raio de convergência de  $f$ ,  $f'$ , ...,  $f^{(r_i)}$ . Por outro lado, se  $\sum \frac{1}{k!} f^{(k)}(z_0)(\lambda_i - z_0)^k$  converge para  $|\lambda_i - z_0| \leq r$  e  $f(z) = \sum \frac{1}{k!} f^{(k)}(z_0)(z - z_0)^k$  converge para  $|z - z_0| \leq r$  então pelo teorema de Abel

$$f(\lambda_i) = \sum \frac{1}{k!} f^{(k)}(z_0)(\lambda_i - z_0)^k \quad (17)$$

converge, bem como suas derivadas até ordem  $(s_i-1)$ . Então para qualquer  $i$  temos que as somas parciais e suas derivadas até a ordem  $(s_i-1)$  convergem em  $\lambda_i$  ( $z_0 = \lambda_i$ ), e pelo teorema 1 segue que

$$f(A) = \sum_{k=0}^{\infty} f^{(k)}(z_0)(A - z_0 I)^k \quad \#$$

Teorema 2. Se A e E são matrizes quadradas e  $t \geq 0$  então

$$e^{(A+E)t} - e^{At} = \int_0^t e^{A(t-s)} E e^{(A+E)s} ds, \quad (18)$$

Prova: Aplicando variação dos parâmetros sobre

$$\begin{cases} x' = Ax + E e^{(A+E)t} \\ x(0) = I \end{cases} \quad (19)$$

Se  $\phi(t)$  é matriz fundamental de

$$\begin{cases} x' = Ax \\ x(0) = I, \end{cases} \quad (20)$$

e  $\psi(t)$  é uma solução de (19), então

$$\psi'(t) = A\psi(t) + E e^{(A+E)t} \quad \text{e} \quad \phi(t) = e^{At}. \quad (21)$$

Fazendo  $\psi(t) = \phi(t)c(t)$ , temos

$$\psi'(t) = \phi'(t)c(t) + \phi(t)c'(t). \quad (22)$$

Assim

$$A\phi(t)c(t) + \phi(t)c'(t) = A\phi(t)c(t) + E e^{(A+E)t} e \quad (23)$$

$$c'(t) = \phi^{-1}(t) E e^{(A+E)t}. \quad (24)$$

Logo

$$c(t) = I + \int_0^t \phi^{-1}(s) E e^{(A+E)s} ds, \quad (c(0) = I). \quad (25)$$

Portanto

$$\psi(t) = \phi(t) + \phi(t) \int_0^t e^{-As} E e^{(A+E)s} ds, \quad (26)$$

ou melhor,

$$\psi(t) = e^{At} + \int_0^t e^{A(t-s)} E e^{(A+E)s} ds. \quad (27)$$

Agora,

$$Y(t) = e^{(A+E)t} \quad (28)$$

é solução de

$$\begin{cases} x'(t) = Ax + Ee^{(A+E)t} \\ x(0) = I \end{cases} \quad (29)$$

também, pois

$$Y'(t) = (A+E)e^{(A+E)t} = Ae^{(A+E)t} + Ee^{(A+E)t} = AY + Ee^{(A+E)t} \quad (30)$$

Logo, por unicidade de soluções segue que

$$e^{(A+E)t} = e^{At} + \int_0^t e^{A(t+s)} E e^{(A+E)s} ds. \quad \# \quad (31)$$