

UNIVERSIDADE ESTADUAL DE CAMPINAS  
INSTITUTO DE MATEMÁTICA, ESTATÍSTICA E  
COMPUTAÇÃO CIENTÍFICA  
DEPARTAMENTO DE MATEMÁTICA APLICADA

**Reconstrução de Funções a partir de Coeficientes de  
Fourier e de Momentos Ortogonais: Novos Métodos \***

*Autor:* YIN JIAHONG

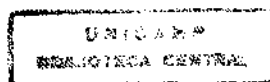
*Orientador:* PROF. DR. ALVARO RODOLFO DE PIERRO

Tese apresentada ao Instituto de  
Matemática Estatística e Com-  
putação Científica da Universidade  
Estadual de Campinas, como parte  
dos pré-requisitos para obtenção  
do título de Doutorado em  
Matemática Aplicada.

**UNICAMP - Campinas**  
**Setembro - 1998**

---

\*Pesquisa financiada pela FAPESP, processo 96/00837-0.



98.2.254

# Reconstrução de Funções a partir de Coeficientes de Fourier e de Momentos Ortogonais: Novos Métodos

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pelo Sr. **Yin Jiahong** e aprovada pela Comissão Julgadora.

Campinas, 03 de Setembro de 1998.



---

Prof. Dr. Alvaro Rodolfo De Pierro  
Orientador

Tese apresentada ao Instituto de Matemática, Estatística e Computação Científica, UNICAMP, como requisito parcial para a obtenção do Título de DOUTOR em Matemática Aplicada.

**Tese de Doutorado defendida e aprovada em 03 setembro de 1998**

**Pela Banca Examinadora composta pelos Profs. Drs.**



---

**Prof (a). Dr (a). ALVARO RODOLFO DE PIERRO**



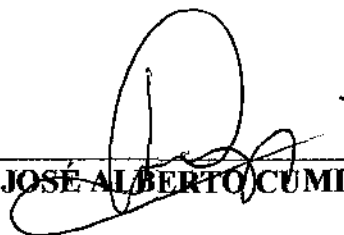
---

**Prof (a). Dr (a). NIR COHEN**



---

**Prof (a). Dr (a). LÚCIO TUNES DOS SANTOS**



---

**Prof (a). Dr (a). JOSÉ ALBERTO CUMINATO**



---

**Prof (a). Dr (a). MAX HENRIQUE MACHADO COSTA**

## Agradecimentos

Agradeço, sinceramente, ao Prof. Álvaro Rodolfo De Pierro não apenas pela orientação científica, mas também pela ajuda amigável recebida no dia-a-dia. Sinto-me muito honrado por ter sido orientado por ele. Durante o período de Doutorado, ele sempre me encorajou a superar todas as dificuldades encontradas no decorrer do trabalho. Sem sua ajuda contínua e sincera, certamente não haveria terminado este trabalho. Aos Professores Musheng Wei e Nir Cohen pelas valiosas contribuições dadas a esta tese deixo meu mais caro agradecimento. Não poderia deixar de agradecer aos Professores Ana, Boldrini, Lúcio e Sonia, com os quais fiz disciplinas, por todas as coisas novas que aprendi e foram vitais na minha pesquisa. Aos amigos Bia, Biloti, Clarice, Diomar, Ernesto, Luiz e Michel pela amizade e ajuda nos momentos difíceis no Brasil. Às secretárias da Matemática Aplicada Fátima e Luciana por toda ajuda recebida. À FAPESP, Capes e Unicamp por terem possibilitado este curso de Pós-graduação. Aos meus pais, Yin Zhizhuan e Xu Shushan, minha esposa, Tang Guirong, minha filha, Yin Feifei, pelo apoio e amor.

## A todos os amigos

Onde há ódio, levar o amor.

Onde há duvida, levar a fé.

Onde há erro, levar a verdade.

Onde há desespero, levar a esperança.

Onde há tristeza, levar a alegria.

Porque é de ti que recebemos lições para a toda a vida!

São Francisco de Assis

## Resumo

Nosso principal objetivo neste trabalho é desenvolver métodos para reconstruir valores discretos de uma função  $f(x)$  com suporte compacto a partir de um número finito de seus coeficientes finitos de Fourier, evitando o chamado fenômeno de Gibbs. Primeiramente, estabelecemos algumas relações aproximadas entre a transformada de Fourier discreta e os coeficientes de Fourier da função  $f(x)$ . Em seguida, usando estas relações, apresentamos alguns algoritmos para a reconstrução de funções. Comparados com os métodos de filtragem mais usados podemos reduzir fortemente o fenômeno de Gibbs na função reconstruída. Como uma aplicação direta destas relações, apresentamos um algoritmo eficiente para calcular os coeficientes de Fourier de  $f(x)$ . Obtemos as estimativas do erro aproximado dos coeficientes de Fourier e a função reconstruída. Discutimos também a reconstrução de um sinal com ruído a partir dos momentos ortogonais e chegamos a uma melhor estimativa do erro. Algumas simulações numéricas ilustram as vantagens de nossos novos métodos.

## Abstract

Our main goal in this dissertation is to develop methods for the reconstruction of the discrete values of a compactly supported function  $f(x)$  from its finite Fourier coefficients, avoiding the so-called Gibbs phenomenon. First, we establish some approximated relations between the discrete Fourier transform and the Fourier coefficients of  $f(x)$ . Then, using these relations, we present several algorithms for reconstructing the function. Compared with the traditional filtering methods, we can greatly reduce the Gibbs phenomenon in the reconstructed function. As a direct application of these relations, we present an efficient algorithm to calculate the Fourier coefficients of  $f(x)$ . We obtain the error estimations of the approximated Fourier coefficients and the reconstructed function. We also discuss the reconstruction of a noisy signal based on its orthogonal moments and get a better error estimation. Several numerical simulations illustrate the advantages of our new approach.

# Conteúdo

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Reconstrução de Uma Função com Suporte Compacto a Partir dos Coeficientes de Fourier . . . . .	1
1.2	Cálculo dos Coeficientes de Fourier Complexos . . . . .	7
1.3	Análise da Reconstrução de um Sinal com Ruído Baseada nos Momentos Ortogonais . . . . .	9
<b>2</b>	<b>O Fenômeno de Gibbs</b>	<b>11</b>
2.1	Sobre o Fenômeno de Gibbs . . . . .	11
2.2	Filtros no Espaço de Fourier . . . . .	15
2.3	O Método de Fourier-Gegenbauer . . . . .	23
<b>3</b>	<b>Uma Nova Relação entre os Coeficientes de uma Série de Fourier e a Transformada de Fourier Discreta.</b>	<b>28</b>
3.1	A Série de Fourier, a Transformada de Fourier e a Transformada de Fourier Discreta . . . . .	28
3.2	Uma Nova Relação entre $\hat{f}_k$ e $\tilde{f}_k$ . . . . .	31
<b>4</b>	<b>Cálculo dos Coeficientes de Fourier Complexos</b>	<b>44</b>
4.1	Algoritmos Eficientes para Calcular os Coeficientes de Fourier Complexos . .	44
4.2	Exemplos Numéricos . . . . .	47



<b>5</b>	<b>A Reconstrução de <math>\{f_j, j = 0, 1, \dots, N-1\}</math> a Partir dos Coeficientes de Fourier <math>\{\hat{f}_k, k = -N/2, \dots, N/2 - 1\}</math></b>	<b>52</b>
5.1	Algoritmos para a Reconstrução de Funções . . . . .	52
5.2	Estimativa do Erro da Função Reconstruída . . . . .	57
5.3	Exemplos Numéricos . . . . .	61
<b>6</b>	<b>Análise da Reconstrução de um Sinal com Ruído Baseado nos Momentos Ortogonais</b>	<b>81</b>
6.1	Resultados Preliminares . . . . .	81
6.2	Algoritmos de Reconstrução . . . . .	84
6.3	Erro do Sinal Reconstruído . . . . .	85
6.4	Exemplos Numéricos . . . . .	91
<b>7</b>	<b>Conclusões</b>	<b>97</b>

# Capítulo 1

## Introdução

Nosso principal objetivo neste trabalho é introduzir novos e mais eficientes métodos para reconstruir os valores discretos de uma função com suporte compacto  $f(x)$  a partir de finitos coeficientes de Fourier. Para isto, estabelecemos primeiramente algumas relações aproximadas entre a transformada de Fourier discreta e os coeficientes de Fourier da função  $f(x)$ . Em seguida, usando estas relações, deduzimos algoritmos para a reconstrução. Comparando os métodos de filtragem conhecidos com nossos métodos, conseguimos reduzir significativamente a influência do fenômeno de Gibbs na função reconstruída. Como uma aplicação direta destas relações, apresentamos um algoritmo eficiente para calcular os coeficientes de Fourier de  $f(x)$ . Para os métodos anteriores obtemos as estimativas do erro na aproximação dos coeficientes de Fourier e da função reconstruída. Trabalhamos também no problema da reconstrução de um sinal com ruído a partir dos momentos ortogonais e obtemos uma melhor estimativa do erro cometido. Em todos os casos, verificamos as vantagens de nossos novos métodos com simulações numéricas.

### 1.1 Reconstrução de Uma Função com Suporte Compacto a Partir dos Coeficientes de Fourier

É muito conhecido que uma função  $f(x)$  com suporte compacto em  $[0,1]$  pode se expressar como uma série de Fourier por

$$f(x) \simeq \sum_{k=-\infty}^{+\infty} \hat{f}_k e^{i2k\pi x}, \quad 0 \leq x \leq 1, \quad (1.1)$$

onde os coeficientes de Fourier  $\hat{f}_k$  são definidos como

$$\hat{f}_k = \int_0^1 f(x) e^{-i2k\pi x} dx, \quad k = 0, \pm 1, \pm 2, \dots \quad (1.2)$$

Lembremos agora os resultados sobre a convergência da série de Fourier. Daqui em diante, uma função definida em  $[0,1]$ , que é estendida em  $\mathbb{R}$  periodicamente, é chamada periódica se  $f(0)$  e  $f(1)$  existem e são iguais. Se  $f(x)$  é contínua, periódica e de variação limitada em  $[0,1]$ , então, a série de Fourier em (1.1) é uniformemente convergente a  $f(x)$ . Se  $f(x)$  é de variação limitada em  $[0,1]$ , ela converge a  $\frac{1}{2}[f(x_0^-) + f(x_0^+)]$  em cada ponto para todo  $x \in [0,1]$  ( $f(0^-) = f(1^-)$ ), onde  $f(x_0^-)$  e  $f(x_0^+)$  denotam respectivamente os limites da função a esquerda e a direita do ponto  $x_0$ .

Em muitas aplicações práticas, encontraremos o problema: dados  $N$  coeficientes de Fourier  $\hat{f}_k$  de uma função desconhecida definida em  $[0,1]$  (assumimos que  $N$  é par), para  $-N/2 \leq k \leq N/2 - 1$ ; como podemos reconstruir exatamente os valores da função? Um método direto é construir uma soma de Fourier clássica como (1.1):

$$f_N(x) = \sum_{k=-N/2}^{N/2-1} \hat{f}_k e^{i2k\pi x}, \quad 0 \leq x \leq 1, \quad (1.3)$$

o que é diferente da discussão teórica da série de Fourier truncada:

$$f_N(x) = \sum_{k=-N/2}^{N/2} \hat{f}_k e^{i2k\pi x}, \quad 0 \leq x \leq 1, \quad (1.4)$$

mas (1.3) corresponde diretamente à forma como a aproximação é calculada na prática. Este é um bom caminho para reconstruir os valores da função se  $f(x)$  é suficientemente suave e periódica. De fato, se  $f(x)$  é analítica e periódica, é conhecido que a série de Fourier converge exponencialmente, isto é:

$$\max_{0 \leq x \leq 1} |f(x) - f_N(x)| \leq e^{-\alpha N}, \quad \alpha > 0.$$

Contudo, se  $f(x)$  é descontínua ou não periódica,  $f_N(x)$  não é uma boa aproximação de  $f(x)$ . Neste caso, a série de Fourier (1.3) converge muito lentamente (convergência sublinear) dentro do intervalo como função de  $N$ , com taxa  $O(N^{-1})$ , e não converge uniformemente em intervalos que contém as descontinuidades, sendo aí onde as grandes oscilações ocorrem.

A incapacidade para recuperar valores de uma função descontínua ou não periódica a partir dos coeficientes de Fourier é o chamado fenômeno de Gibbs. Isso parece significar que é inerentemente impossível obter as informações locais exatas (os valores da função) a partir do conhecimento de propriedades totais (os coeficientes de Fourier) de uma função suave por partes. Isto é mais uma consequência do fato de as funções harmônicas terem o suporte estendido a toda a reta.

A conclusão anterior tem consequências importantes, do ponto de vista numérico: muitos fenômenos físicos são representados por uma função suave por partes. Por exemplo, os problemas em engenharia aeronáutica e do espaço envolvem escoamento de fluidos que incluem ondas de choque, isto é, há descontinuidades no campo da pressão. Na previsão numérica do tempo considera-se a face do globo que tem declives grandes em regiões montanhosas [12]. No cálculo numérico de equações diferenciais parciais pelo método de Fourier-Galerkin [28], se há descontinuidade numa solução exata da equação, então, a solução de Fourier-Galerkin é caracterizada pelas fortes oscilações globais. Em consequência, a convergência do erro piora numa região suave longe da descontinuidade. Para resolver equações não lineares, o problema é mais complexo porque aquelas oscilações podem induzir instabilidade [20]. O fenômeno de Gibbs parece limitar os métodos baseados na aproximação global para tais problemas. Outro exemplo é a compressão de dados. É essencial ter uma boa representação de dados para decidir como compactá-los, isto é, qual é a informação que deve ser mantida. É importante saber se a representação de dados de Fourier é adequada para sinais não suaves.

Muitas metodologias foram pesquisadas para resolver este problema. Dentre elas destacam-

se os métodos de filtragem [30] e o método de Fourier-Gegenbauer (ou FG) [12], [13]. Há dois tipos de métodos de filtragem: a aproximação no espaço de Fourier (modificando os coeficientes da expansão) e a aproximação no espaço físico. Ambos os métodos são bem-sucedidos a respeito da precisão, mas longe da descontinuidade apenas. O método de Fourier-Gegenbauer é bem-sucedido na remoção completa do fenômeno de Gibbs, mas ainda existem algumas dificuldades práticas para sua implementação. A quantidade de cálculos em alguns casos é tão grande que não podem ser efetuados pelos computadores atuais.

Problemas semelhantes também aparecem em outro método de transformada de Fourier (FT). Por exemplo, em reconstrução de imagens a partir de um conjunto de integrais sobre uma reta de densidade do objeto, obtemos as amostras discretas da transformada de Fourier pelo teorema do corte (ou teorema de projeção) de Fourier [17]. Precisamos reconstruir os valores de densidade discreta do objeto. É claro que podemos reconstruir exatamente uma função com suporte compacto a partir de todos os valores da transformada de Fourier. Mas, quando conhecemos apenas um conjunto discreto de valores da transformada de Fourier, a questão é como recuperamos alguns valores discretos da função original? Por simplicidade, discutimos o caso em que a função é de uma variável e com o suporte em  $[0,1]$ . Mas a extensão a dimensões maiores é imediata. A transformada de Fourier e a transformada de Fourier inversa são respectivamente definidas por

$$\hat{f}(\omega) = \int_{-\infty}^{+\infty} f(x)e^{-i2\pi x\omega} dx = \int_0^1 f(x)e^{-i2\pi x\omega} dx \quad (1.5)$$

e

$$f(x) = \int_{-\infty}^{+\infty} \hat{f}(\omega)e^{i2\pi x\omega} d\omega, \text{ a.e.} \quad (1.6)$$

Como  $f(x)$  é uma função com o suporte compacto em  $[0,1]$ ,  $\hat{f}(\omega)$  é suficientemente suave (Teorema de Paley-Wiener, ver [3]). Notemos que, neste caso, o suporte de  $\hat{f}(\omega)$  é  $\mathbb{R}$ .

Sejam

$$0 = x_0 < x_1 < \dots < x_{N-1} < x_N = 1, \quad \Delta x_j = x_{j+1} - x_j = \Delta x = \frac{1}{N} \quad (1.7)$$

e

$$f_j = f(\eta_j), \quad j = 0, \dots, N-1 \text{ com } \eta_j = x_j + \delta, \quad (1.8)$$

onde  $0 \leq \delta \leq \Delta x$  é uma constante.

Como podemos utilizar a transformada de Fourier discreta para aproximar a transformada de Fourier contínua? Em [5], mostra-se que podemos desenvolver a transformada de Fourier como um caso particular da transformada de Fourier contínua, usando uma amostra da função no domínio do tempo e da frequência. Se  $\{\hat{f}(k), k = -N/2, \dots, N/2-1\}$  em (1.5) são dados, então,

$$f_j \approx \sum_{k=-N/2}^{N/2-1} \hat{f}(k) e^{i2k\pi x_j}, \quad j = 0, \dots, N-1, \quad (1.9)$$

onde, notemos que  $\hat{f}(k) = \hat{f}_k$ . Da fórmula acima, observamos que este problema é um caso particular da aproximação da série finita de Fourier. Então, aparece o fenômeno de Gibbs.

No que segue, através de aproximação da integral, consideramos este problema novamente.

Para obter os valores aproximados de  $\{f_j, j = 0, \dots, N-1\}$ , suponhamos que, se  $|\omega|$  é suficientemente grande, i.e., existe uma constante positiva  $W$  tal que se  $|\omega| \geq W$ , o intervalo correspondente na integral (1.6) pode ser ignorado. Então, temos

$$f(x) \approx \int_{-W}^W \hat{f}(\omega) e^{i2\pi x \omega} d\omega, \quad 0 \leq x \leq 1. \quad (1.10)$$

Aproximando a integral (1.10) pela soma de Riemann, temos [8]

$$f_j \approx \sum_{k=0}^{M-1} \hat{f}(\omega_k) e^{i2\pi \eta_j \omega_k} \Delta \omega_k, \quad j = 0, \dots, N-1, \quad (1.11)$$

onde

$$-W = \omega_0 < \omega_1 < \dots < \omega_{M-1} < \omega_M = W, \text{ e } \Delta\omega_k = \omega_{k+1} - \omega_k = \Delta\omega = \frac{2W}{M}.$$

Pelo fato de que a integral (1.10) é aproximada pela soma de Riemann (1.11), podemos obter uma boa aproximação somente para  $\Delta\omega$  pequeno, logo,  $M$  tem que ser grande para a constante fixa  $W$ . Então, precisamos uma amostra muito grande de  $\hat{f}(\omega_k)$ . Por outro lado, preferimos utilizar a transformada rápida de Fourier para reduzir a quantidade de cálculos e isto significa que tomamos a frequência máxima  $W = \frac{1}{2\Delta x}$  e escolhemos  $M = N$ . No caso em que o suporte de  $f(x)$  é  $[0, 1]$ ,  $\Delta\omega = \frac{2W}{M} = 1$ , então, isto produz um erro grande quando utilizamos a soma de Riemann em (1.11). Também, observamos que  $\omega_k = k - N/2$ ,  $k = 0, \dots, N-1$ , e  $\hat{f}(\omega_k) = \hat{f}_k$ . Neste caso, (1.11) é mesmo como (1.3), então, este problema é equivalente à aproximação com a série de Fourier finita e o fenômeno de Gibbs ocorre. Se tomamos  $\Delta\omega = O(\Delta x) = O(\frac{1}{N})$  e  $W = \frac{1}{2\Delta x}$ , a ordem dos cálculos será  $O(N^3)$  e isso é proibitivamente grande para  $N$  grande, especialmente no caso de dimensões maiores. É claro que, se o suporte de uma função  $f(x)$  é  $[a, b]$  com a distância  $A = b - a$  grande, o erro na soma de Riemann em (1.11) pode ser pequeno, mas a aproximação (1.10) será pior ainda, pois  $W = \frac{1}{2\Delta x}$  é pequeno. Alguns tipos de filtro (ou função de janela) são utilizados para reduzir o erro [30].

No nosso trabalho, consideramos o fenômeno de Gibbs para funções que aparecem comumente na prática. Portanto faremos as seguintes suposições, não muito fortes, sobre a função:

**Suposições.** A função  $f(x)$  satisfaz

1.  $f(x)$  é definida para  $x \in R$ ,  $f \in L^2(R)$ ;  $f(x) = 0$  para  $x \notin [0, 1]$  e  $\sup_{x \in [0, 1]} |f(x)| \leq C$ , onde  $C > 0$  é uma constante.
2.  $f(x)$  tem um número finito de pontos de descontinuidade  $z_1, \dots, z_l$  conhecidos à priori.

3. Para  $x \notin Z = \{z_1, \dots, z_l\}$ ,  $f'(x)$  existe e  $\sup_{x \notin Z} |f'(x)| \leq C$ .

As Suposições 1-3 são perfeitamente aceitáveis para funções que aparecem nos problemas práticos. Se o suporte de  $f(x)$  é no intervalo  $[a, b] \neq [0, 1]$ , podemos fazer a transformação  $\xi = \frac{x-a}{b-a}$  e mudar o suporte para  $[0, 1]$ . Também, notemos que os pontos de descontinuidade, quando não são dados do problema, podem ser localizados a partir dos coeficientes de Fourier, e então, podemos assumir que os pontos de descontinuidade sempre são dados. No Teorema 3.3 do Capítulo 3, precisamos a seguinte suposição:

4. Para  $x \notin Z = \{z_1, \dots, z_l\}$ , existe  $f''(x)$  e  $\sup_{x \notin Z} |f''(x)| \leq C$ .

Nossa abordagem para reduzir o fenômeno de Gibbs pertence ao conjunto de métodos de filtragem. Mas, obtemos os novos filtros usando um ponto de vista completamente novo e diferente dos métodos tradicionais. Deduzimos uma nova relação entre a transformada de Fourier discreta de valores discretos de  $f(x)$  e os seus coeficientes de Fourier. Usando esta relação, obtemos um novo filtro e apresentamos um algoritmo eficiente para reconstruir uma função original a partir dos coeficientes de Fourier (ou de amostras discretas da transformada de Fourier). Assim, podemos utilizar a FFT e obtemos uma precisão muito maior que a obtida com os outros métodos na literatura.

## 1.2 Cálculo dos Coeficientes de Fourier Complexos

Em muitos problemas práticos, temos que calcular os coeficientes de Fourier de uma função real e não periódica  $f(x)$  ( $x \in [a, b]$ ). Como não se conhece uma fórmula fechada para calcular os coeficientes de Fourier temos que aproximá-los de algum modo. É por causa desta aproximação que muitos métodos numéricos que usam a série de Fourier perdem uma boa parte da sua precisão.

Existem muitos métodos práticos para calcular os coeficientes de Fourier. A transformada



rápida de Fourier (FFT) é extensamente utilizada, por causa do seu cálculo simples e rápido [5]. Neste caso, a transformada de Fourier é usada como uma aproximação dos coeficientes de Fourier. Mas, esta aproximação para a integral induz um erro fundamental. Por outro lado, usando este método, podemos calcular apenas  $N$  coeficientes de Fourier a partir de  $N$  amostras. M. A. Slonim *et al.*, apresentaram um método descontínuo (DM) em [25], onde os coeficientes de Fourier são calculados como uma função de valores dos saltos nos pontos de descontinuidade (e derivadas para algumas funções). Este método é muito eficiente para uma função escada, exponencial e senoidal por partes, que sempre aparecem em sistemas elétricos. Mas não é bom para uma função suave por partes. De fato, podemos ter a fórmula exata para calcular os coeficientes de Fourier destas funções elementares. E. Neuman [23] apresentou o algoritmo para calcular os coeficientes de Fourier usando funções Spline naturais. Aproximou a função  $f(x)$  pelas funções Splines de grau ímpar  $s(x)$  com nós arbitrários e grau  $(2q-1)$ ,  $1 \leq q$ , então, os coeficientes exatos de Fourier de  $s(x)$  servem como aproximação dos coeficientes de Fourier de  $f(x)$ . Ele apresentou uma estimativa do erro desta aproximação no caso em que a função é suficientemente suave e os nós da função Spline natural  $s(x)$  são equidistantes. Este método funciona bem para uma função suficientemente suave e os coeficientes de Fourier  $\hat{f}_k$  podem ser calculados para qualquer inteiro  $k$ . Mas a quantidade do cálculo é enorme. M.T. Abuelma'atti [1] apresentou um algoritmo simples baseado apenas em operações matemáticas básicas. As amostras da função não precisam ser equidistantes e podemos calcular todos os coeficientes. Este método consiste na interpolação da função entre as amostras por uma função linear por partes, calculando depois a inclinação de cada segmento. Mas, este método não funciona bem para uma função descontínua. Por outro lado, é muito difícil estender estes métodos para dimensões maiores (exceto quando é possível usar a FFT).

O método apresentado no Capítulo 4 é deduzido em um modo natural a partir de definição dos coeficientes de Fourier. Mostramos que este método consegue uma melhor exatidão que outros existentes e com menos cálculos.

### 1.3 Análise da Reconstrução de um Sinal com Ruído Baseada nos Momentos Ortogonais

Vários tipos de momentos têm sido regularmente usados em reconhecimento de modelos, análise de imagens, compressão de dados e redução de ruído [24] [4] [27]. Existem momentos geométricos, complexos, Fourier-Melin, radiais e ortogonais. As suas propriedades podem ser encontradas na literatura. Mas, os problemas fundamentais sobre a robustez com relação ao ruído e à digitalização raramente são discutidos. Quando reconstruímos um sinal com ruído como podemos recuperar 'exatamente' o sinal original a partir de um conjunto finito de momentos? É claro que, quanto maior a ordem dos momentos, maior informação sobre o sinal eles fornecem, mas, também é maior a degradação que sofrem por causa do ruído. Em [24], discute-se a reconstrução de imagens a partir de dados discretos com ruído pelo método dos momentos e apresenta-se a estimativa do erro mínimo quadrático que consiste numa relação entre a ordem de momentos, a suavidade de imagem, a taxa da amostragem e as características do ruído. Mas a estimativa do erro é muito grosseira, especialmente para um número grande (de momentos)  $N$ . No Capítulo 6, melhoramos o resultado apresentado em [24]. O principal resultado é uma estimativa correta do erro 'tight', e mostramos que o erro da reconstrução para um sinal sem ruído ou com ruído não tende a infinito se a discretização do sinal é fixa.

Por simplicidade, no Capítulo 6, discutimos somente o problema de recuperação de sinal em uma dimensão, mas todos os resultados podem ser facilmente estendidos para dimensões maiores. Como em [24], somente escolhemos polinômios de Legendre para momentos ortogonais. Usando os dados discretos do sinal com ruído, apresentamos três algoritmos de reconstrução e obtemos a estimativa do erro de reconstrução correspondente. Esta estimativa revela correta e claramente uma relação entre a ordem dos momentos, a suavidade de sinal, a taxa da amostragem e as características do ruído. Também mostramos algumas simulações que são consistentes com nossa análise. Por outro lado, somente consideramos

um ruído aditivo.

Esta tese está organizada como segue. No Capítulo 2 revisamos alguns métodos de filtragem e outros trabalhos sobre o fenômeno de Gibbs. No Capítulo 3, deduzimos algumas novas relações aproximadas entre os coeficientes de Fourier  $\hat{f}_k$  da função  $f(x)$  (ou a transformada de Fourier com frequência  $k$ ) e a transformada de Fourier discreta  $\tilde{f}_k$  das amostras discretas  $\{f_j, j = 0, 1, \dots, N-1\}$ , se consideramos que  $f(\cdot)$  é definida em seu suporte  $[0, 1]$ , e a partir dela, deduzimos um novo filtro. No Capítulo 4, apresentamos um algoritmo eficiente para calcular os coeficientes de Fourier. No Capítulo 5, apresentamos nossos algoritmos para a reconstrução de  $\{f_j, j = 0, 1, \dots, N-1\}$  a partir dos coeficientes de Fourier (ou dos valores da transformada de Fourier)  $\{\hat{f}_k, k = -N/2, \dots, N/2 - 1\}$ . Nos Capítulos 4 e 5, alguns experimentos numéricos ilustram as vantagens dos novos métodos. No Capítulo 6 discutimos a reconstrução de um sinal com ruído baseada nos momentos ortogonais, e deduzimos uma melhor estimativa do erro. No Capítulo 7, concluímos nosso trabalho com algumas observações.

# Capítulo 2

## O Fenômeno de Gibbs

Neste capítulo, descrevemos os principais métodos existentes para superar o fenômeno de Gibbs e explicamos porque este fenômeno aparece na aproximação de uma função por uma série de Fourier finita. Apresentamos uma recapitulação sobre os métodos de filtragem para reduzir este fenômeno, que estão relacionados com nosso trabalho. Também, revisamos alguns resultados sobre o método de Fourier-Gegenbauer que, teoricamente, supera completamente o fenômeno de Gibbs, porém com um custo computacional elevado.

### 2.1 Sobre o Fenômeno de Gibbs

Primeiramente, vejamos um exemplo:

$$f(x) = \begin{cases} 1, & 1/4 < x \leq 3/4, \\ 0, & 0 < x \leq 1/4, \quad 3/4 < x \leq 1. \end{cases} \quad (2.1)$$

Os coeficientes de Fourier desta função são:

$$\hat{f}(k) = \begin{cases} 1/2, & k = 0, \\ 0, & k \neq 0, \text{ par}, \\ \frac{(-1)^{(k-1)/2}}{k\pi}, & k \neq 0, \text{ ímpar}. \end{cases} \quad (2.2)$$

O gráfico da série de Fourier truncada desta função aparece na Fig. 2.1. Há algumas oscilações se comparada com a função original. Quando  $N$  aumenta, a amplitude das oscilações

aumenta numa vizinhança da descontinuidade apesar de tender pontualmente a um limite finito (que não é o valor da função original na descontinuidade). A convergência da série não é uniforme. Este é o fenômeno de Gibbs.

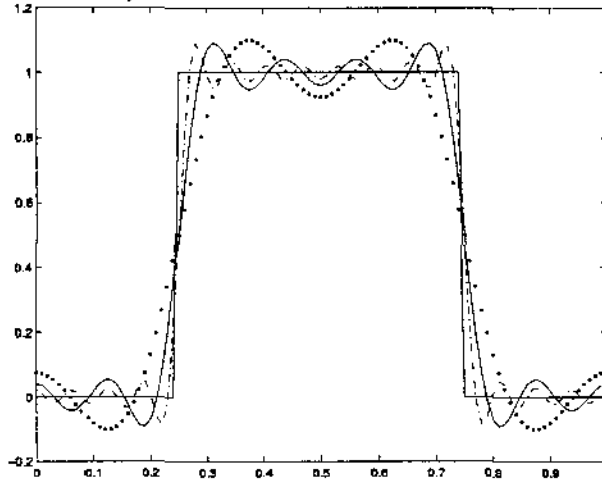


Figura 2.1: Aproximação da série de Fourier truncada da onda quadrada com diferentes valores de  $N$ :  $N = 8$  ( $\dots$ ),  $N = 16$  ( $\_$ ) e  $N = 32$  ( $-.-$ ).

Este comportamento pode ser explicado facilmente em termos da representação da integral singular da série de Fourier truncada. De (1.2) e (1.4), temos

$$\begin{aligned} f_N(x) &= \sum_{k=-N/2}^{N/2} \hat{f}_k e^{i2k\pi x} \\ &= \sum_{k=-N/2}^{N/2} \int_0^1 f(y) e^{-i2k\pi y} dy e^{i2k\pi x} = \int_0^1 \left[ \sum_{k=-N/2}^{N/2} e^{i2k\pi(x-y)} \right] f(y) dy. \end{aligned}$$

Então, a expressão integral de  $f_N(x)$  é como segue

$$f_N(x) = \int_0^1 D_N(x-y) f(y) dy, \tag{2.3}$$

onde  $D_N(\xi)$  é o núcleo de Dirichlet

$$\begin{aligned}
D_N(\xi) &= 1 + 2 \sum_{k=1}^{N/2} \cos(2k\pi\xi) \\
&= \begin{cases} \frac{\sin((N+1)\pi\xi)}{\sin(\pi\xi)}, & \xi \notin Z, \\ N+1, & \xi \in Z. \end{cases} \quad (2.4)
\end{aligned}$$

O gráfico do núcleo de Dirichlet no intervalo  $[-1/2, 1/2]$  é mostrado na Fig. 2.2. O núcleo de Dirichlet pode ser considerado como a projeção ortogonal da distribuição Delta no espaço dos polinômios trigonométricos de grau  $N/2$ , com o produto interno canônico em  $L^2$ .  $D_N(\xi)$  é uma função par que muda de sinal nos pontos  $\{\xi_j = j/(N+1)\}$  e satisfaz:

$$\int_0^1 D_N(\xi) d\xi = 1, \quad (2.5)$$

o que pode ser obtido, fazendo  $f = 1$  em (2.3). Além disso, quando  $N \rightarrow \infty$ ,  $D_N$  tende uniformemente a zero em cada intervalo fechado (excluindo os pontos singulares  $\xi \in Z$ , onde  $Z$  é o conjunto dos inteiros). Isto implica que, para todo  $\delta > 0$  e todo  $\epsilon > 0$ , existe um inteiro  $N(\delta, \epsilon) > 0$  tal que

$$|D_N(\xi)| < \epsilon, \text{ se } N > N(\epsilon, \delta) \text{ e } \delta \leq \xi \leq 1 - \delta. \quad (2.6)$$

Voltemos agora à onda quadrada mostrada na Fig. 2.1. Por simplicidade, mudaremos a origem para o ponto de descontinuidade, isto é, consideraremos uma função periódica

$$\phi(x) = \begin{cases} 1, & 0 \leq x < 1/2, \\ 0, & 1/2 \leq x < 1. \end{cases} \quad (2.7)$$

A série de Fourier truncada de  $\phi$  é

$$\phi_N(x) = \int_{x-1/2}^x D_N(y) dy = \int_0^x D_N(y) dy + \int_{-1/2}^0 D_N(y) dy + \int_{x-1/2}^{-1/2} D_N(y) dy.$$

Se  $x$  não está próximo de  $1/2$ , por (2.6), a última integral no lado direito é arbitrariamente pequena, quando  $N$  é suficientemente grande. A segunda integral, por (2.5), é igual a  $1/2$ , então,

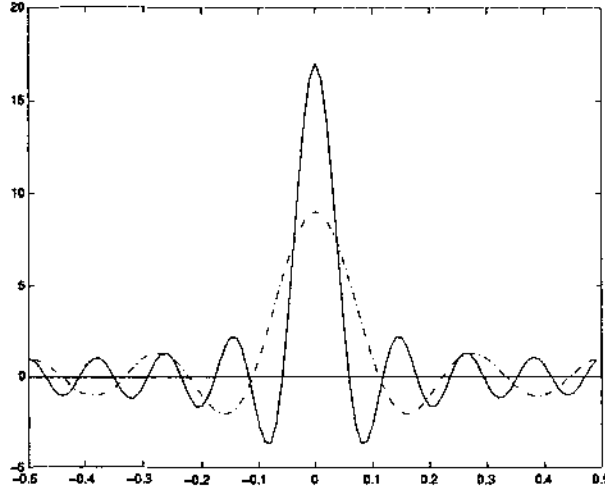


Figura 2.2: Núcleo de Dirichlet para  $N = 8$  (- -) e  $N = 16$  (—).

$$\phi_N(x) \simeq 1/2 + \int_0^x D_N(y)dy, \text{ se } N \rightarrow \infty.$$

Esta fórmula explica o fenômeno de Gibbs para a onda quadrada. Se  $x > 0$  está suficientemente longe de 0, então,  $\int_0^x D_N(y)dy \simeq \int_0^{1/2} D_N(y)dy = 1/2$  por (2.5) e (2.6), portanto  $\phi_N(x)$  está próximo de 1. Mas, a função  $x \rightarrow \int_0^x D_N(y)dy$  tem os máximos e os mínimos alternados nos pontos onde  $D_N$  torna-se nula,  $\xi_j = j/(N+1)$ ; isto explica o comportamento oscilatório. O máximo absoluto ocorre em  $\xi_1 = 1/(N+1)$ , onde, para  $N$  suficientemente grande,

$$\int_0^{1/(N+1)} D_N(y)dy \simeq \frac{1}{\pi} \int_0^\pi \frac{\sin t}{t} dt = 0.58949\dots$$

Então, a seqüência  $\{\phi_N(\xi_j)\}$  tende a  $1.08949\dots > 1 = \phi(0^+)$ , se  $N \rightarrow \infty$ . Equivalentemente

$$\lim_{N \rightarrow \infty, x \rightarrow 0^+} \sup \phi_N(x) > \phi(0^+).$$

Igualmente, para  $x$  negativo, obtemos

$$\lim_{N \rightarrow \infty, x \rightarrow 0^-} \sup \phi_N(x) < \phi(0^-).$$

Esta é a caracterização matemática do fenômeno de Gibbs.

Se  $f(x)$  é uma função de variação limitada em  $[0,1]$ , e há uma descontinuidade tipo salto em  $x = x_0$ , obtemos

$$f_N(x) \simeq \frac{1}{2}[f(x_0^+) + f(x_0^-)] + [f(x_0^+) - f(x_0^-)] \int_0^{x-x_0} D_N(y) dy, \quad N \rightarrow \infty. \quad (2.8)$$

Isto mostra que a seqüência  $\{f_N(x)\}$  sofre o fenômeno de Gibbs em  $x = x_0$  com a mesma estrutura do caso da onda quadrada [7].

## 2.2 Filtros no Espaço de Fourier

A filtragem no espaço de Fourier está relacionada com nosso trabalho. Consideraremos a projeção de Fourier-Galerkin (1.3) de uma função suave (ou analítica) por partes  $f(x)$ ,  $0 \leq x \leq 1$ . A razão da convergência lenta (e não uniforme) de  $f_N(x)$  para  $f(x)$  deve-se a dois fatos: o decaimento lento dos coeficientes de Fourier  $\hat{f}_k$  e a natureza global da série de Fourier, onde os coeficientes de Fourier são determinados no intervalo total (incluindo os pontos de descontinuidade). Aumentando a taxa da decaimento dos coeficientes de Fourier, sem perder a exatidão, pode-se reduzir o fenômeno de Gibbs. Esta é a idéia do método de filtragem no espaço de Fourier. Mas, é óbvio que o decrescimento nas freqüências altas provocará perda de informação no caso de uma função descontínua. Então, este método não funciona bem numa vizinhança da descontinuidade. Descreveremos o método como segue.

Assumiremos que são dados  $N$  coeficientes de Fourier de uma função suave (ou analítica) por partes. Sem perda de generalidade, podemos considerar apenas um ponto de descontinuidade  $x = \xi$ . Gostaríamos de recuperar os valores de  $f(x)$ ,  $0 \leq x \leq 1$ , multiplicando os coeficientes de Fourier pelo fator  $\sigma(\frac{2k}{N})$ , tal que a soma modificada



$$f_N^\sigma(x) = \sum_{k=-N/2}^{N/2-1} \hat{f}_k \sigma\left(\frac{2k}{N}\right) e^{i2k\pi x}, \quad (2.9)$$

convirja mais rapidamente que a soma original (1.3).

Como na fórmula (2.3), também, podemos rescrever (2.9) como uma convolução no espaço físico:

$$f_N^\sigma(x) = \int_0^1 \Psi_N(x-y) f(y) dy, \quad (2.10)$$

onde a função do núcleo  $\Psi_N(\xi)$  é a representação do filtro  $\sigma\left(\frac{2k}{N}\right)$  no espaço físico.

$$\Psi_N(\xi) = \sum_{k=-N/2}^{N/2-1} \sigma\left(\frac{2k}{N}\right) e^{i2k\pi\xi}, \quad 0 \leq \xi \leq 1. \quad (2.11)$$

Para entender por que o método de filtragem é eficiente para reduzir o fenômeno de Gibbs, introduziremos a definição geral do filtro dada por Harvè H. Vandeven [30], que caracteriza o comportamento de  $\sigma(\eta)$  como uma função do seu argumento.

**Definição 2.1** (Filtro no espaço de Fourier de ordem  $p$ ) Uma função real e suave  $\sigma : R \rightarrow R$  é chamada um filtro de ordem  $p$  se

1.  $\sigma(0) = 1, \sigma^{(l)}(0) = 0; \quad 1 \leq l \leq p-1,$
2.  $\sigma(\eta) = 0, \quad 1 \leq |\eta|,$
3.  $\sigma \in C^{p-1}.$

Desta definição, concluímos que o filtro não modifica as frequências baixas, mas apenas as altas. Notemos que, pelas condições 2 e 3, temos que

$$\sigma^{(l)}(1) = 0, \quad 0 \leq l \leq p-1.$$

Mas, somente omitir as frequências altas não é suficiente. É essencial que  $\sigma(\eta)$  seja uma função suave de  $\eta$ . Cortar as frequências altas sem modificar o resto (isto é, usando  $\sigma(\eta) = 1,$

para  $|\eta| \leq \eta_0$  e  $\sigma(\eta) = 0$  para  $|\eta| > \eta_0$ ) não melhora a convergência, pois somente fica a soma original com os poucos termos. Por outro lado, se  $f(x)$  é uma função periódica ( $f \in C^{p-1}$ ), então, multiplicando os seus coeficientes de Fourier pelo filtro que satisfaz a condição 1, não muda a ordem da exatidão. Isto, é as condições 2 e 3 são fundamentais, se  $f(x)$  é apenas suave por partes.

Há muitos exemplos de filtros que tem sido usados durante anos. Mencionaremos alguns no que segue.

1. Em 1900 Fejér sugeriu utilizar a soma média parcial em vez da soma original. Isto é equivalente a um filtro de primeira ordem,

$$\sigma_1(\eta) = 1 - |\eta|. \quad (2.12)$$

2. O filtro de Lanczos é formalmente um filtro de primeira ordem,

$$\sigma_2(\eta) = \frac{\sin(\pi\eta)}{\pi\eta}. \quad (2.13)$$

Mas, notemos que em  $\eta = 0$ , ele satisfaz as condições do filtro de segunda ordem. Em [16], mostra-se que este filtro é obtido somente por médio dos valores da função.

3. O filtro de segunda ordem é o filtro de Raised-Cosine,

$$\sigma_3(\eta) = \frac{1}{2}[1 + \cos(\pi\eta)]. \quad (2.14)$$

4. O filtro de Raised-Cosine afiado é dado por

$$\sigma_4(\eta) = \sigma_3^4(\eta)[35 - 84\sigma_3(\eta) + 70\sigma_3^2(\eta) - 20\sigma_3^3(\eta)]. \quad (2.15)$$

Este é um filtro de quarta ordem.

5. Um filtro exponencial de ordem  $p$  (para  $p$  par) é dado por

$$\sigma_5(\eta) = e^{-\alpha\eta^p}. \quad (2.16)$$

Notemos que, formalmente, o filtro exponencial não conforma-se como na definição de filtro pois  $\sigma_5(1) = e^{-\alpha}$ . Na prática, podemos escolher  $\alpha$  tal que  $e^{-\alpha}$  está dentro do erro do computador específico. Este filtro é sempre usado por causa da sua simplicidade e dos bons resultados numéricos.

6. H. Vandeven [30] sugeriu um filtro de ordem  $p$  como segue:

$$\sigma_6(\eta) = 1 - \frac{(2p-1)!}{(p-1)!^2} \int_0^{\eta} [t(1-t)]^{p-1} dt. \quad (2.17)$$

Este é essencialmente o filtro induzido pelas wavelets de Daubechies [12].

Estes filtros podem ser caracterizados pelo núcleo  $\Psi_N(\xi)$ . De fato, a representação (2.10) permite descrever as formas do método com filtro mais geral que (2.11). O núcleo  $\Psi_N(\xi)$  não precisa possuir a forma particular de (2.11). A única condição é que  $\Psi_N(\xi)$  seja um polinômio aproximante da função Delta, isto é, um polinômio trigonométrico de grau  $N/2$  tal que

$$\int_0^1 \Psi_N(\xi) d\xi = 1$$

e tal que para  $\delta > 0$  e todo  $\epsilon > 0$  existe um inteiro  $N(\delta, \epsilon) > 0$  tal que

$$|\Psi_N(\xi)| < \epsilon, \text{ se } N > N(\delta, \epsilon) \text{ e } \delta \leq \xi < 1 - \delta. \quad (2.18)$$

Sujeitos a estas suposições, podemos repetir os argumentos usados para induzir (2.8), e obtemos a fórmula assintótica:

$$f_N^\sigma(x) \simeq \frac{1}{2}[f(x_0^+) + f(x_0^-)] + [f(x_0^+) - f(x_0^-)] \int_0^{x-x_0} \Psi_N(y) dy \quad (2.19)$$

nos pontos próximos dos pontos de descontinuidade de  $f(x)$ . Então, o comportamento de  $f_N^\alpha(x)$  depende da função

$$\phi_N(z) = \int_0^z \Psi_N(y) dy$$

na vizinhança da origem. Há o fenômeno de Gibbs, se existe uma seqüência  $\{z_n > 0\}$ , com  $z_n \rightarrow 0$ , quando  $N \rightarrow \infty$ , em que  $1/2 < \alpha \leq \phi_N(z_N)$  (para algum  $\alpha$  independente de  $N$ ), neste caso

$$\lim_{N \rightarrow \infty} f_N^\alpha(z_N) > f(x_0^+).$$

No que segue, escrevemos os núcleos de alguns filtros e descrevemos seus gráficos na Fig. 2.3, comparados com o núcleo de Dirichlet. Notemos que  $\Psi_N(\xi)$  é uma função periódica, e descrevemos seu gráfico no intervalo  $[-0.5, 0.5]$ .

1. O núcleo  $\Psi_N^1$  gerado pela soma de Cesáro é chamado núcleo de Fejér. A sua expressão analítica é dada por

$$\begin{aligned} \Psi_N^1(\xi) &= 1 + 2 \sum_{k=1}^{N/2} \left(1 - \frac{k}{N/2+1}\right) \cos 2k\pi\xi \\ &= \begin{cases} \frac{1}{N/2+1} \frac{\sin[(N+1)\pi\xi]}{\sin(\pi\xi)}, & \xi \neq j, \quad j \in \mathbb{Z}, \\ N/2 + 1, & \xi = j, \quad j \in \mathbb{Z}. \end{cases} \end{aligned}$$

2. O núcleo  $\Psi_N^2$  correspondente ao filtro de Lanczos é dado por

$$\Psi_N^2(\xi) = 1 + \sum_{k=1}^{N/2} \frac{\sin[k(2\pi\xi + 2\pi/N)] + \sin[k(2\pi\xi - 2\pi/N)]}{2k\pi/N}.$$

3. O núcleo  $\Psi_N^3(\xi)$  associado com o filtro de Raised-Cosine é

$$\Psi_N^3(\xi) = \frac{1}{4} [D_N(2\pi\xi - 2\pi/N) + 2D_N(2\pi\xi) + D_N(2\pi\xi + 2\pi/N)].$$

4. O núcleo  $\Psi_N^6$  correspondente a  $\sigma_6$  é descrito na Fig. 2.3 (d)

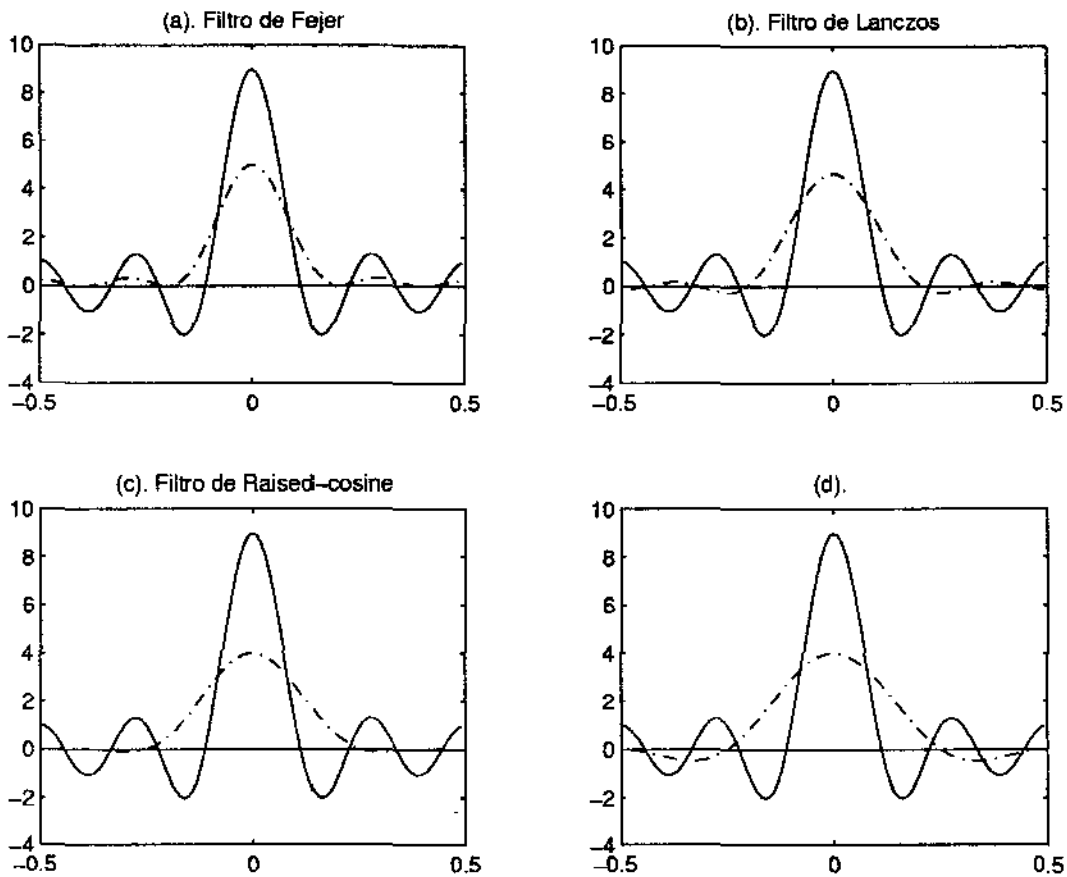


Figura 2.3: Comparação do núcleo de Dirichlet dos filtros para  $N=8$ . O núcleo de Dirichlet é denotado pela curva sólida e os núcleos dos filtros pela curva tracejada. (d) mostra o núcleo do filtro  $\sigma_6$  com  $p = 10$ .

Em seguida, revemos uma estimativa do erro, usando o método dos filtros. Para quaisquer filtros de ordem  $p$ , H. Vandeven provou o seguinte resultado [30]:

**Teorema 2.1** Seja  $f(x)$  uma função por partes ( $f(x) \in C^p$ ) com um ponto de descontinuidade  $z$ . Seja  $\sigma(\frac{2k}{N})$  um filtro de ordem  $p$  que satisfaz a Definição 2.1. E mais, seja  $x$  em  $[0, 1]$  e  $d(x) = \min_{k=-1,0,1} |x - z + k|$ . Então

$$|f(x) - f_N^\sigma(x)| \leq CN^{1-p}d(x)^{1-p}K(f) + CN^{\frac{1}{2}-p}\|f^{(p)}\|_{L^2} \quad (2.20)$$

onde

$$K(f) = \sum_{l=0}^{p-1} d(x)^l (f^{(l)}(z^+) - f^{(l)}(z^-)) \int_{-\infty}^{\infty} |G_l^{(p-l)}(\eta)| d\eta$$

e

$$G_l(\eta) = \frac{\sigma(\eta) - 1}{\eta^l}.$$

O teorema acima demonstra que o processo filtrado funciona bem longe da descontinuidade, pois todos os termos no lado direito em (2.20) podem ser  $O(N^{1-p})$ . Então, a ordem  $p$  é alcançada para uma função suave por partes. Para qualquer função analítica por partes, H. Vandeven [30] demonstrou que, se admitida a ordem  $p$  do filtro  $\sigma_6(\eta)$ , então, uma exatidão exponencial pode ser obtida nos pontos de descontinuidade.

**Teorema 2.2** Seja  $f(x)$  uma função analítica por partes com um ponto de descontinuidade  $z$ . Seja  $\sigma(\frac{2k}{N})$  um filtro  $\sigma_6(\frac{k}{N})$  de ordem  $p = cN^{\epsilon/4}$ , onde  $c$  é uma constante positiva independente de  $N$ . Então, a seguinte estimativa é válida:

$$\sup_{x \in [0,1], d(x) > N^{-1+\epsilon}} |f(x) - f_N^{\sigma_6}(x)| \leq N^\beta (CN^{-\epsilon/2})^{N^{\epsilon/4}}. \quad (2.21)$$

para uma constante positiva  $C$  independente de  $N$  e uma constante positiva  $\beta$  independentes de ambos de  $f$  e  $N$ .

Uma das vantagens do método de filtragem no espaço de Fourier é que não precisa cálculos adicionais. O cálculo da soma parcial de Fourier ou da soma filtrada pode ser feito com  $N \log(N)$  operações.

Todos os filtros mencionados acima são filtros bilaterais, isto é, a região nos dois lados em torno da descontinuidade tem que ser excluída na estimativa do erro. Isto é esperado pois utilizam núcleos simétricos  $\Psi_N(\xi)$  em (2.11) como o resultado de  $\sigma(\frac{2k}{N}) = \sigma(-\frac{2k}{N})$ . Em geral, é inaceitável para resolver uma equação diferencial parcial não-linear, pois aquelas oscilações talvez eventualmente poluem a região suave e/ou produzem instabilidade não-linear. Podemos utilizar outros tipos de filtros que só satisfazem as condições parciais na Definição 2.1. Por exemplo, Wei Cai *et al.* [6] demonstraram a existência de um filtro complexo  $\sigma(\frac{2k}{N})$ . Eles apresentaram um filtro unilateral,

$$\sigma_7\left(\frac{2k}{N}\right) = \sigma_6\left(\frac{2k}{N}\right) \sum_{l=1}^m C_m^l (-1)^{l+1} e^{ilk\Delta}, \quad (2.22)$$

onde

$$\Delta = \frac{1}{N^{1-\varepsilon}}, \quad m = N^{\varepsilon/4},$$

tal que para qualquer função analítica  $f(x)$  em  $[0,1]$ , tal que  $f(0) \neq f(1)$ , a soma de Fourier filtrada (2.9) satisfaz

$$\max_{0 \leq x \leq x_R} |f(x) - f_N^{\sigma_7}(x)| \leq N^\beta (CN^{\varepsilon/2})^{-N^{\varepsilon/4}}, \quad (2.23)$$

onde

$$x_R = 1 - \frac{2}{N^{1-(5/4)\varepsilon}}$$

e  $C$  e  $\beta$  são as constantes independentes de  $N$ . O filtro  $\sigma_7(\frac{2k}{N})$  é baseado na seguinte idéia simples: como a soma de Fourier filtrada bilateralmente  $f_N^{\sigma_6}$  é uma boa aproximação para  $f(x)$  na região  $\Delta$  longe de descontinuidade, podemos utilizar a extrapolação a partir dos pontos dentro desta região para outros pontos. Este filtro funciona bem assintoticamente, mas um inteiro suficientemente grande  $N$  é necessário para que  $m$  seja de tamanho razoável. Eles também apresentaram outros filtros unilaterais usando o método de quadrados mínimos.

Outros tipos de filtros são aqueles no espaço físico, que são construídos para localizar as informações que determinam os coeficientes de Fourier  $\hat{f}_k$ . Acham uma função  $\Psi(x, y)$  tal que

$$\int_0^1 \Psi(x, y) f_N(y) dy \sim f(x).$$

Notemos que se  $\Psi(x, y) = \Psi(x - y)$ , voltamos à fórmula (2.10). D. Gottlieb e M. Tadmor [14] apresentaram um filtro deste tipo. Mas, como o método de filtragem no espaço de Fourier, não consegue uma boa precisão próximo da descontinuidade.

## 2.3 O Método de Fourier-Gegenbauer

Descreveremos brevemente os trabalhos recentes sobre o fenômeno de Gibbs. Em [12], [13], D. Gottlieb e C. W. Shu demonstraram que os primeiros  $(2N + 1)$  coeficientes de Fourier  $\hat{u}(k)$  de uma função analítica mas não-periódica  $u(x)$  contêm informações suficientes para construir uma função interpolante com a exatidão espectral. Seu conceito básico é re-expandir a soma de Fourier que converge lentamente na série de Gegenbauer que converge bem mais rápido (a convergência exponencial desta série do polinômio foi provada nos termos limitados). Eles apresentaram um procedimento para eliminar o fenômeno de Gibbs completamente, isto é, para obter uma exatidão exponencial na norma do máximo em qualquer intervalo de analiticidade, usando as séries de Fourier e de Gegenbauer de uma função descontínua mas analítica por partes. Aplicaram este método em vários exemplos. Descrevemos o principal resultado no que segue. Como o polinômio de Gegenbauer é definido em  $[-1, 1]$ , nesta seção, por conveniência, assumimos que  $u(x)$  é uma função analítica por partes mas não-periódica em  $[-1, 1]$ , e sua soma de Fourier truncada é

$$u_N(x) = \sum_{k=-N}^N \hat{u}_k e^{i\pi k x}. \quad (2.24)$$



**Definição 2.2** Os polinômios de Gegenbauer  $C_n^\lambda(x)$  são explicitamente dados por

$$(1-x^2)^{\lambda-\frac{1}{2}}C_n^\lambda(x) = \frac{(-1)^n}{2^n n!} G(\lambda, n) \frac{d^n}{dx^n} [(1-x^2)^{n+\lambda-\frac{1}{2}}], x \in [-1, 1], \quad (2.25)$$

onde  $G(\lambda, n)$  é definido por

$$G(\lambda, n) = \frac{\Gamma(\lambda + \frac{1}{2})\Gamma(n + 2\lambda)}{\Gamma(2\lambda)\Gamma(n + \lambda + \frac{1}{2})}.$$

Assumindo que os primeiros  $(2N + 1)$  coeficientes de Fourier  $\hat{u}(k)$  são dados, estamos interessados em recuperar os primeiros  $m$  coeficientes na expansão de Gegenbauer de  $u(x)$  (com exatidão exponencial na norma do máximo):

$$u(x) = \sum_{l=0}^{\infty} \hat{u}^\lambda(l) C_l^\lambda(x), \quad (2.26)$$

onde os coeficientes de Gegenbauer são definidos por

$$\hat{u}^\lambda(l) = \frac{1}{h_l^\lambda} \int_{-1}^1 (1-x^2)^{\lambda-\frac{1}{2}} u(x) C_l^\lambda(x) dx \quad (2.27)$$

e

$$h_l^\lambda = \pi^{\frac{1}{2}} C_l^\lambda(1) \frac{\Gamma(\lambda + \frac{1}{2})}{\Gamma(\lambda)\Gamma(l + \lambda)}.$$

Desde que não conhecemos a função exata  $u(x)$ , mas somente a sua série de Fourier truncada (2.24), podemos ter apenas uma aproximação de  $\hat{u}^\lambda(l)$ , denotada por  $\hat{g}_N^\lambda(l)$ , dada por

$$\hat{g}_N^\lambda(l) = \frac{1}{h_l^\lambda} \int_{-1}^1 (1-x^2)^{\lambda-\frac{1}{2}} u_N(x) C_l^\lambda(x) dx. \quad (2.28)$$

Agora um fato notável: que os coeficientes aproximados de  $\hat{g}_N^\lambda(l)$  podem ser expressados explicitamente em termos dos coeficientes de Fourier  $\hat{u}(k)$  como segue:

$$\hat{g}_N^\lambda(l) = \delta_{0l}\hat{u}(0) + \Gamma(\lambda)l^l(l+\lambda) \sum_{0 < |k| \leq N} J_{l+\lambda}(\pi k) \left(\frac{2}{\pi k}\right)^\lambda \hat{u}(k), \quad (2.29)$$

onde  $\Gamma(\lambda)$  e  $J_n(x)$  são as funções Gamma e de Bessel respectivamente. A correspondente expansão de Gegenbauer, baseada nos coeficientes aproximados  $\hat{g}_N^\lambda(l)$ , será

$$u_{m,N}^\lambda(x) = \sum_{l=0}^m \hat{g}_N^\lambda(l) C_l^\lambda(x). \quad (2.30)$$

as equações (2.29) e (2.30) são conhecidos como a aproximação de FG de  $f(x)$ .

A diferença entre a soma parcial de Gegenbauer com  $m$  termos da função  $u(x)$

$$u_m^\lambda(x) = \sum_{l=0}^m \hat{u}^\lambda(l) C_l^\lambda(x) \quad (2.31)$$

e aquela da série de Fourier truncada  $u_N(x)$  é chamada de erro truncado,

$$TE(x, \lambda, m, N) = |u_m^\lambda(x) - u_{m,N}^\lambda(x)| = \left| \sum_{l=0}^m [\hat{u}^\lambda(l) - \hat{g}_N^\lambda(l)] C_l^\lambda(x) \right|. \quad (2.32)$$

Isto mede o erro na expansão de Gegenbauer finita por causa do truncamento de Fourier.

O erro total da aproximação de FG,

$$E(x, \lambda, m, N) = |u(x) - u_{m,N}^\lambda(x)|, \quad (2.33)$$

pode ser separado em dois componentes como segue:

$$\begin{aligned} |u(x) - u_{m,N}^\lambda(x)| &= |u(x) - u_m^\lambda(x) + u_m^\lambda(x) - u_{m,N}^\lambda(x)| \\ &\leq |u(x) - u_m^\lambda(x)| + |u_m^\lambda(x) - u_{m,N}^\lambda(x)|. \end{aligned}$$

O segundo termo é o erro de truncamento (2.32). O primeiro termo

$$RE(x, \lambda, m, N) = \left| \sum_{l=0}^{\infty} \hat{u}^\lambda(l) C_l^\lambda(x) - \sum_{l=0}^m \hat{u}^\lambda(l) C_l^\lambda(x) \right|, \quad (2.34)$$

por causa do truncamento da série de Gegenbauer, chama-se erro de regularização.

Agora, descreveremos os resultados básicos sobre a convergência do método de FG [12], [28]. Estes resultados são válidos para uma função analítica  $u(x)$ ,  $x \in [-1, 1]$ , tal que suas derivadas satisfazem a estimativa:

$$\max_{-1 \leq x \leq 1} \left| \frac{d^k u(x)}{dx^k} \right| \leq C(\rho) \frac{k!}{\rho^k}, \quad (2.35)$$

onde  $1 \leq \rho$  e  $C(\rho)$  são constantes, independentes de  $k$ .

1. Ambos os erros: o de truncamento e de regularização, definidos em (2.32) e (2.34), alcançam seus máximos na fronteira  $x = \pm 1$ .

2. Na fronteira  $x = 1$ , o erro de truncamento satisfaz

$$TE(1, \lambda, m, N) \leq A \frac{(m + \lambda)\Gamma(m + 2\lambda)\Gamma(\lambda)}{(m - 1)!\Gamma(2\lambda)} \left(\frac{2}{\pi N}\right)^{\lambda-1}.$$

3. Se  $\lambda = \gamma m$  e  $m = \beta N$  onde  $\gamma, \beta$  são constantes positivas, então, o erro de truncamento decai exponencialmente com  $N$  como segue:

$$TE(1, \lambda, m, N) \leq AN^2 q^N, \quad q = \left[ \frac{\beta^\gamma (1 + 2\gamma)^{1+2\gamma}}{(2\pi e \gamma)^\gamma} \right]^\beta.$$

4. Na fronteira  $x = 1$  o erro de regularização satisfaz

$$RE(1, \lambda, m, N) \leq A(\lambda, \rho) m^\lambda q^m, \quad q = \frac{1}{2\rho},$$

onde  $A(\lambda, \rho)$  é independente de  $m$ , o erro decai exponencialmente com  $m$  pois  $q < 1$ .

5. Se  $\lambda = \gamma m$ , então

$$TE(1, \lambda, m, N) \leq Aq^m, \quad q = \frac{(1 + 2\gamma)^{1+2\gamma}}{\rho 2^{1+2\gamma} \gamma^\gamma (1 + \gamma)^{(1+\gamma)}}.$$

É fácil verificar que  $q$  é uma função crescente de  $\gamma$  e  $q < \frac{1}{\rho} \leq 1$  para todo  $\gamma > 0$ . Então, a convergência espectral é mais rápida quando  $\gamma$  é pequeno.

Os resultados acima demonstram a convergência espectral da aproximação de FG no limite  $m \rightarrow \infty$ ,  $N \rightarrow \infty$ . Em [28] L. Vozovoi apontou que o mínimo número de termos necessário

para obter um erro pequeno por causa do fato que  $q^{-m}$  (ou  $q^{-N}$ ) pode ser consideravelmente grande, se os coeficientes  $C(\rho)$ , e  $A$  nas estimativas acima são grandes. Na prática, o cálculo da série de Gegenbauer é limitada a  $m, \lambda \sim 100$  por causa das funções Gamma e de Bessel em (2.29). E ainda, o cálculo dos termos de ordem alta é sujeito ao erro da máquina. L. Vozovoi *et al.* aplicaram o método de FG a EDP [29], e apresentaram algumas técnicas como a subtração do polinômio para acelerar a convergência. Sobre o método de FG temos as seguintes observações:

1. O fenômeno de Gibbs pode ser completamente evitado, do ponto de vista teórico, usando o método de FG. Os resultados de D. Gottlieb *et al.* são os melhores sobre como remover o fenômeno de Gibbs.
2. A quantidade de cálculos envolvidos no método de FG pode ser muito grande para algumas funções. Em alguns casos pode ser proibitivo por causa do número grande  $m$  e da computação da função de Bessel.
3. Em dimensões maiores o esforço computacional pode ser enorme.
4. Alguns métodos de FG modificados para acelerar a convergência são eficientes para EDP, mas de difícil aplicação em outras áreas.

Nosso trabalho está somente relacionado com o método de filtragem no espaço de Fourier. Nos capítulos seguintes apresentaremos alguns filtros práticos para uma função em  $C^1$  ou  $C^2$ . Para estas funções, não podemos escolher o filtro de ordem superior a 2. Comparado com os outros filtros com a mesma ordem  $p$ ,  $p \leq 2$ , nosso filtro é melhor em dois aspectos: a exatidão da reconstrução e os bons resultados numa vizinhança das descontinuidades.

## Capítulo 3

# Uma Nova Relação entre os Coeficientes de uma Série de Fourier e a Transformada de Fourier Discreta.

O objetivo de nosso trabalho neste capítulo é deduzir uma relação aproximada entre a transformada de Fourier discreta  $\{\tilde{f}_k\}$  e os coeficientes de Fourier  $\{\hat{f}_k\}$  de uma função  $f(x)$ , tal que possamos calcular  $\{f_j, j = 0, \dots, N - 1\}$  eficiente e exatamente usando a transformada rápida de Fourier (FFT) (ver [5]). Baseados nesta relação, conseguimos obter um novo filtro no espaço de Fourier para reduzir o fenômeno de Gibbs. Em primeiro lugar, realçaremos alguns resultados e motivações das série de Fourier, a transformada de Fourier e a transformada de Fourier discreta. Depois, estabeleceremos nossa nova relação.

### 3.1 A Série de Fourier, a Transformada de Fourier e a Transformada de Fourier Discreta

No Capítulo 1, já mencionamos alguns conceitos sobre a série de Fourier de uma função  $f(x)$  com o suporte compacto contido no intervalo  $[0,1]$ :

$$\begin{cases} f(x) \simeq \sum_{k=-\infty}^{+\infty} \hat{f}_k e^{i2k\pi x} \\ \hat{f}_k = \int_0^1 f(x) e^{-i2k\pi x} dx \end{cases}$$

e a transformada de Fourier contínua e a fórmula inversa:

$$\begin{cases} f(x) = \int_{-\infty}^{+\infty} \hat{f}(\omega) e^{i2\pi x \omega} \\ \hat{f}(\omega) = \int_0^1 f(x) e^{-i2\pi x \omega} dx \end{cases}$$

É claro que a relação entre os coeficientes de Fourier  $\hat{f}_k$  e a transformada de Fourier  $\hat{f}(k)$  com frequência  $k$  é como segue:

$$\hat{f}_k \equiv \hat{f}(k) = \int_0^1 f(x) e^{-i2k\pi x} dx, \quad k = 0, \pm 1, \pm 2, \dots \quad (3.1)$$

Nos problemas práticos, somente tomamos a aproximação da série de Fourier finita (1.3), ou tomamos os valores discretos da transformada de Fourier contínua. Além disso, precisamos obter apenas os valores discretos da função. Embora nós temos duas expressões (1.1) e (1.6) de uma função com o suporte compacto para diferentes problemas, podemos ter a mesma fórmula do cálculo de (1.3) e (1.9) para  $x_j, j = 0, \dots, N - 1$  (quando o suporte da função é  $[0,1]$ ). Neste caso, portanto, a aproximação da série de Fourier da função é equivalente à transformada de Fourier contínua para o cálculo, e em ambos casos ocorrerá o fenômeno de Gibbs.

Recordemos a definição da transformada de Fourier discreta. Dada uma função  $f(x) \in L^2(\mathbb{R})$ , com o suporte em  $[0,1]$  e um inteiro par  $N > 0$ , sejam  $\{x_j, j = 0, \dots, N\}$  e  $\{f_j, j = 0, \dots, N - 1\}$  definidos em (1.7) e (1.8) respectivamente. Então, a transformada de Fourier discreta de  $\{f_j, j = 0, \dots, N - 1\}$  é

$$\tilde{f}_k = \frac{1}{N} \sum_{j=0}^{N-1} f_j e^{-i2k\pi \frac{j}{N}}, \quad -N/2 \leq k \leq N/2 - 1 \quad (3.2)$$

e a transformada inversa é:

$$f_j = \sum_{k=-N/2}^{N/2-1} \tilde{f}_k e^{i2k\pi \frac{j}{N}}, \quad j = 0, \dots, N-1. \quad (3.3)$$

A transformada de Fourier discreta e a fórmula inversa são as projeções exatas entre  $N$  números complexos  $\{f_j, j = 0, \dots, N-1\}$  e  $N$  números complexos  $\{\tilde{f}_k, k = -N/2, \dots, N/2-1\}$ . Podemos utilizar a transformada rápida de Fourier (FFT) e a transformada rápida inversa de Fourier (IFFT) para calculá-los respectivamente. Também, podemos escolher  $N$  ímpar inteiro. Mas, escolheremos os valores da função  $f(x)$  nos pontos  $2^m$  (algum inteiro  $m$ ), para poder utilizar a FFT. Neste caso, os valores de  $\{f_j, j = 0, \dots, N-1\}$  em (3.3) serão complexos, e temos que omitir a parte imaginária.

Qual é a relação entre a transformada de Fourier discreta e a transformada de Fourier contínua? A referência [5] explica claramente que a transformada de Fourier discreta pode ser deduzida da transformada de Fourier contínua. Baseados nisto, podemos utilizar a transformada de Fourier discreta para calcular aproximadamente a transformada de Fourier contínua. Qual é o erro desta aproximação? Comparando (1.9) (ou (1.3)) com (3.3), podemos facilmente obter (1.9) apenas substituindo  $\tilde{f}_k$  por  $\hat{f}_k$  em (3.3). Em geral, não é possível obter uma melhor aproximação de  $\{f_j, j = 0, \dots, N-1\}$  em (1.9), a menos que  $\tilde{f}_k \equiv \hat{f}_k, k = -N/2, \dots, N/2-1$ . Quando nós temos alguma relação mais exata entre  $\tilde{f}_k$  e  $\hat{f}_k$ , podemos calcular bem  $\{f_j, j = 0, \dots, N-1\}$  a partir da aproximação  $\{\tilde{f}_k, k = -N/2, \dots, N/2-1\}$ , o que é nossa principal motivação.

Há uma relação exata entre  $\tilde{f}_k$  e  $\hat{f}_k$  [7]:

$$\tilde{f}_k = \hat{f}_k + \sum_{m=-\infty, m \neq 0}^{+\infty} \hat{f}_{k+Nm}, \quad k = -N/2, \dots, N/2-1. \quad (3.4)$$

e sabemos que tem o erro  $\sum_{m=-\infty, m \neq 0}^{+\infty} \hat{f}_{k+Nm}$  se aproximamos  $\tilde{f}_k$  por  $\hat{f}_k$ .

## 3.2 Uma Nova Relação entre $\hat{f}_k$ e $\tilde{f}_k$

Nesta seção, deduziremos uma nova relação aproximada entre  $\hat{f}_k$  e  $\tilde{f}_k$ . Usando esta relação, seremos capazes de apresentar um novo filtro para reduzir o erro quando calculamos  $\{f_j, j = 0, \dots, N-1\}$  pela FFT. Por outro lado, podemos também obter um algoritmo aproximado para calcular eficientemente os coeficientes da expansão de Fourier. Antes de tudo, apresentamos o principal resultado da seção.

**Teorema 3.1** Suponhamos que uma função  $f(x)$  satisfaz as Suposições 1-3 (Capítulo 1),  $\{x_j\}_0^N$  e  $\{f_j\}_0^{N-1}$  são definidos em (1.7) e (1.8) respectivamente, e  $z_m \in \{x_j\}_1^{N-1}, m = 1, \dots, l$ . Então,

$$|\hat{f}_k - a_k \tilde{f}_k| \leq C(f) \Delta x, \text{ para qualquer inteiro } k \quad (3.5)$$

onde

$$a_k = \begin{cases} 1, & k = 0, \\ \frac{N(e^{-i2k\pi/N} - 1)}{-i2k\pi} = \frac{\sin(k\pi/N)}{k\pi/N} e^{-ik\pi/N}, & k \neq 0, \end{cases} \quad (3.6)$$

e  $C(f)$  é uma constante não-negativa que depende de  $f$  somente.

**Demonstração:** Primeiramente, fazemos a prova para o caso  $k \neq 0$ . Pelas partições dos nós  $\{x_j, j = 0, 1, \dots, N\}$  definidos em (1.7) e os valores discretos  $\{f_j, j = 0, 1, \dots, N-1\}$  definidos em (1.8), temos (notemos que  $x_j = \frac{j}{N}$ )

$$\int_{x_j}^{x_{j+1}} e^{-i2k\pi x} dx = \frac{e^{-i2k\pi x}}{-i2k\pi} \Big|_{x_j}^{x_{j+1}} = e^{-i2k\pi \frac{j}{N}} \frac{e^{-i2k\pi/N} - 1}{-i2k\pi},$$

e obtemos

$$\int_{x_j}^{x_{j+1}} e^{-i2k\pi x} dx = \frac{a_k}{N} e^{-\frac{i2k\pi j}{N}}. \quad (3.7)$$

Observamos que a igualdade acima também é válida se  $k = 0$ . Assim, aproximando  $f(x)$  pela constane por partes, obtemos



$$\hat{f}_k = \sum_{j=0}^{N-1} \int_{x_j}^{x_{j+1}} f(x) e^{-i2k\pi x} dx = \sum_{j=0}^{N-1} f_j \int_{x_j}^{x_{j+1}} e^{-i2k\pi x} dx + r_k(\Delta x),$$

onde  $r_k(\Delta x)$  é o termo do erro que será especificado mais adiante. Então,

$$\hat{f}_k = \sum_{j=0}^{N-1} f_j e^{-i2k\pi \frac{j}{N}} \frac{e^{-i2k\pi/N} - 1}{-i2k\pi} + r_k(\Delta x).$$

Usando que

$$\tilde{f}_k = \frac{1}{N} \sum_{j=0}^{N-1} f_j e^{-i2k\pi \frac{j}{N}}, \quad -N/2 \leq k \leq N/2 - 1,$$

temos

$$\hat{f}_k = \tilde{f}_k \frac{N(e^{-i2k\pi/N} - 1)}{-i2k\pi} + r_k(\Delta x) = a_k \tilde{f}_k + r_k(\Delta x). \quad (3.8)$$

Agora voltamos para estimar o termo  $r_k(\Delta x)$ . Supondo que a derivada de  $f(x)$  existe no subintervalo  $(x_j, x_{j+1})$ , para qualquer  $x \in (x_j, x_{j+1})$  temos que

$$f(x) = f_j + f'(\xi)(x - \eta_j), \quad \text{para algum } \xi \in (x_j, x_{j+1})$$

Usando a Suposição 3, obtemos

$$\left| \int_{x_j}^{x_{j+1}} [f(x) - f_j] e^{-i2k\pi x} dx \right| \leq \int_{x_j}^{x_{j+1}} \sup_{\xi} |f'(\xi)| |x - \eta_j| dx \leq \frac{1}{2} C(\Delta x)^2,$$

onde  $C$  é definido na Suposição 1. Escolhendo  $\eta_j$  como  $x_j + \Delta x/2$ , o limitante acima será  $\frac{1}{4} C(\Delta x)^2$ . Da desigualdade acima e de (3.1), (3.2), (3.6), (3.7) e (3.8), obtemos

$$\begin{aligned} |r_k(\Delta x)| &= |\hat{f}_k - a_k \tilde{f}_k| \leq \sum_{j=0}^{N-1} \int_{x_j}^{x_{j+1}} |f(x) - f_j| dx \\ &\leq N \frac{1}{2} C(\Delta x)^2 \equiv C(f) \Delta x, \end{aligned} \quad (3.9)$$

No caso  $k = 0$ , então, o valor da integral em (3.7) é  $1/N$ . Seguindo a demonstração no caso  $k \neq 0$ , podemos facilmente obter a mesma estimativa com  $a_0 = 1$ . Assim, terminamos a prova deste teorema.

As seguintes observações são necessárias:

1. É claro que, se  $f(x)$  é aproximada usando funções constantes por partes, então, coeficiente exato da sua expansão de Fourier, que é  $a_k \tilde{f}_k$ , será uma aproximação de  $\hat{f}_k$ . Tal idéia é semelhante à aquela no método de Splines [23]. Mas, nossa expressão mostra a relação entre  $\tilde{f}_k$  e  $\hat{f}_k$  de maneira que podemos utilizar a FFT nas operações numéricas. Além disto, a fórmula é exata para uma família de funções muito comuns.

2. Seja  $\sigma(k) = \frac{1}{a_k}$ , logo, por (3.6), temos que

$$\sigma(k) = \begin{cases} 1, & k = 0, \\ \frac{k\pi/N}{\sin(k\pi/N)} e^{ik\pi/N}, & k \neq 0, \end{cases} \quad (3.10)$$

e podemos verificar que

$$1 \leq |\sigma(k)| \leq \frac{\pi}{2}$$

a partir da desigualdade bem conhecida

$$\frac{2}{\pi} \leq \frac{\sin \theta}{\theta} \leq 1 \quad \text{para } 0 < \theta \leq \frac{\pi}{2}.$$

3. Quando calculamos os valores aproximados de  $\{f_j, j = 0, 1, \dots, N-1\}$  pela FFT inversa, podemos utilizar os valores aproximados de  $\tilde{f}_k$ :

$$\tilde{f}_k \approx \sigma(k) \hat{f}_k, \quad k = -N/2, \dots, N/2 - 1. \quad (3.11)$$

Notemos que  $\sigma(k)$  é um filtro no espaço de Fourier, compararemos este filtro com aqueles na seção 2.2:

1). O filtro de Fejér

$$\sigma_1\left(\frac{2k}{N}\right) = 1 - 2|k|/N.$$

2). O filtro de Lanczos

$$\sigma_2\left(\frac{2k}{N}\right) = \begin{cases} 1, & k = 0, \\ \frac{\sin(2k\pi/N)}{2k\pi/N}, & k \neq 0. \end{cases}$$

3). O filtro de Raised-Cosine

$$\sigma_3\left(\frac{2k}{N}\right) = \frac{1 + \cos(2k\pi/N)}{2}.$$

Todos os filtros acima são reais. Uma razão essencial para introduzir estes filtros é reduzir as oscilações numa vizinhança da descontinuidade, e além disto, a meticulosidade é a segunda consideração. Na nossa simulação numérica, achamos que os filtros sem o fator 2,  $\sigma_2(\frac{k}{N})$  e  $\sigma_3(\frac{k}{N})$ , devem ser melhores. Também comparamos com o filtro seguinte.

4).

$$\sigma_6\left(\frac{2k}{N}\right) = 1 - \frac{(2p-1)!}{(p-1)!^2} \int_0^{\frac{2k}{N}} [t(1-t)]^{p-1} dt.$$

A magnitude do filtro  $\sigma(k)$  (3.10) é a inversa do filtro de Lanczos, e ainda, tem uma mudança de fase  $e^{ik\pi/N}$  em  $\sigma(k)$ . Podemos reivindicar o fato de que nosso filtro é bem melhor no sentido de que estabelecemos uma relação mais exata entre  $\tilde{f}_k$  e  $\hat{f}_k$  via a fórmula (3.5). Além disso, conseguimos obter uma melhor aproximação para uma função  $f(x)$  que tem pontos de descontinuidade ou  $f(1) \neq f(0)$ .

4. Nos dois casos particulares, obtemos resultados surpreendentes (a reconstrução exata de  $\{f(\eta_j) = f_j\}_0^{N-1}$ ), que mostramos nos seguintes corolários.

**Corolário 1.** Suponhamos que  $f(x)$  é uma função escada com o suporte em  $[0, 1]$ . Sejam  $\{x_j\}_0^N$  e  $\{f_j\}_0^{N-1}$  definidos por (1.7) e (1.8) respectivamente, tal que todos os pontos interiores no subintervalo  $(x_j, x_{j+1})$  não são os pontos de descontinuidade, para  $j = 0, 1, \dots, N-1$ . Se  $f_j = f(\eta_j) = f(x_j + \delta)$  com  $0 < \delta < \Delta x$ , então

$$a_k \tilde{f}_k = \hat{f}_k, \quad k = -N/2, \dots, N/2 - 1. \quad (3.12)$$

Aplicando a FFT inversa de  $\{\sigma(k)\hat{f}_k\}_{-N/2}^{N/2-1}$  podemos reconstruir os valores  $\{f_j\}_0^{N-1}$  exatamente.

**Demonstração:** Sob as condições do corolário, temos  $C = 0$  na estimativa de  $r_k(\Delta x)$  na fórmula (3.9), logo,  $r_k(\Delta x) = 0$ . Então

$$\tilde{f}_k = \sigma(k)\hat{f}_k \quad k = -N/2, \dots, N/2 - 1,$$

completando a prova deste corolário.

**Corolário 2.** Suponhamos que  $f(x)$  é uma função com o suporte em  $[0, 1]$ , tal que  $f(x) = \alpha x + \beta$  para  $x \in [0, 1]$ . Sejam  $\{x_j\}_0^N$  e  $\{f_j\}_0^{N-1}$  definidos por (1.7) e (1.8) respectivamente. Se  $f_j = f(\eta_j) = f(x_j + \Delta x/2)$ , então

$$\hat{f}_k = a_k \tilde{f}_k, \quad k = -N/2, \dots, N/2 - 1, \quad (3.13)$$

e aplicando a IFFT de  $\{\sigma(k)\hat{f}_k\}_{-N/2}^{N/2-1}$ , podemos reconstruir os valores  $\{f_j\}_0^{N-1}$  exatamente.

**Demonstração:** Sob as condições do corolário, integrando por partes, temos

$$\int_{x_j}^{x_{j+1}} [f(x) - f_j] e^{-i2k\pi x} dx = B_k \alpha e^{-i2k\pi j/N}, \quad B_k = \begin{cases} 0, & k = 0, \\ i \left[ \frac{\cos(k\pi/N)}{2kN\pi} - \frac{\sin(k\pi/N)}{4k^2 N \pi^2} \right], & k \neq 0. \end{cases}$$

para  $k = -N/2, \dots, N/2 - 1$ . Notemos que  $B_k$  é independente de  $j$ . Substituindo a identidade acima dentro da fórmula de  $r_k(\Delta x)$  na prova do Teorema 3.1, obtemos, para  $k = -N/2, \dots, N/2 - 1$ ,

$$\begin{aligned} r_k(\Delta x) &= \hat{f}_k - a_k \tilde{f}_k = \sum_{j=0}^{N-1} \int_{x_j}^{x_{j+1}} [f(x) - f_j] e^{-i2k\pi x} dx \\ &= B_k \alpha \sum_{j=0}^{N-1} e^{-i2k\pi j/N} = 0, \end{aligned}$$

o que prova este corolário.

Observamos que os resultados acima também são válidos no caso em que o suporte de  $f(x)$  é  $[a, b]$ .

Assumindo as condições do Teorema 3.1, suponhamos que os pontos da descontinuidade de  $f(x)$  coincidem com alguns nós  $x_j$ . Mas, para uns problemas mais práticos, precisamos

considerar o caso de que os pontos de descontinuidade estão situados em qualquer lugar. Por conveniência, vamos supor que  $f(x)$  possui apenas um ponto de descontinuidade no intervalo  $(0,1)$ , denotado por  $z$ , e  $x_p < z < x_{p+1}$ ,  $0 < p < N - 1$ . Agora, utilizaremos as novas partições não-uniformes dos nós tal que a derivada de  $f(x)$  existe em cada subintervalo. Sejam

$$0 = x_0 < x_1 < \dots < x_p < z < x_{p+1} < \dots < x_{N-1} < x_N = 1, \quad (3.14)$$

$$\Delta x_j = x_{j+1} - x_j = \Delta x = \frac{1}{N}, \quad j \neq p, p+1, \quad \Delta x_p = z - x_p, \quad \Delta x_{p+1} = x_{p+1} - z, \quad (3.15)$$

e

$$\begin{aligned} f_j &= f(\eta_j), \quad j = 0, \dots, N-1 \quad \text{com} \quad \eta_j = x_j + \delta_j, \quad j \neq p, p+1 \\ f_p &= f(x_p + \delta_p), \quad f_{p+1} = f(z + \delta_{p+1}) \end{aligned} \quad (3.16)$$

onde  $0 \leq \delta_j = \delta \leq \Delta x$ ,  $j = 0, \dots, N-1$ ,  $j \neq p, p+1$ ;  $0 \leq \delta_p \leq \Delta x_p$  e  $0 \leq \delta_{p+1} \leq \Delta x_{p+1}$  são constantes.

Obtemos uma relação aproximada entre  $\hat{f}_k$  e  $\tilde{f}_k$  como segue.

**Teorema 3.2** Suponhamos que uma função  $f(x)$  satisfaz as Suposições 1-3, e  $f(x)$  possui apenas um ponto de descontinuidade interior  $z$ . As partições dos nós no intervalo  $[0,1]$  e  $\{f_j\}_0^{N-1}$  são as definidas em (3.14)-(3.16) respectivamente. Então

$$|\hat{f}_k - a_k \tilde{f}_k - (f_p - f_{p+1}) \hat{g}_k| \leq C(f) \Delta x, \quad (3.17)$$

onde

$$\hat{g}_k = \begin{cases} z - x_{p+1}, & k = 0, \\ \frac{1}{-i2k\pi} (e^{-i2k\pi z} - e^{-i2k\pi x_{p+1}}), & k \neq 0, \end{cases} \quad (3.18)$$

e  $C(f)$  é uma constante não-negativa que somente depende de  $f$ .  $\{a_k\}$  é a mesma do Teorema 3.1.

**Demonstração:** A prova é semelhante à aquela do Teorema 3.1. Primeiramente, consideraremos o caso  $k \neq 0$ . Usando a regra dos retângulos para a aproximação da integral, obtemos

$$\begin{aligned}\hat{f}_k &= \sum_{j=0, j \neq p, p+1}^{N-1} \int_{x_j}^{x_{j+1}} f(x) e^{-i2k\pi x} dx + \int_{x_p}^z f(x) e^{-i2k\pi x} dx + \int_z^{x_{p+2}} f(x) e^{-i2k\pi x} dx \\ &= \sum_{j=0, j \neq p, p+1}^{N-1} f_j \int_{x_j}^{x_{j+1}} e^{-i2k\pi x} dx + f_p \int_{x_p}^z e^{-i2k\pi x} dx + f_{p+1} \int_z^{x_{p+2}} e^{-i2k\pi x} dx + e_k(\Delta x),\end{aligned}$$

onde  $e_k(\Delta x)$  é o termo do erro. A fim de utilizar os resultados do Teorema 3.1, rescrevemos a igualdade acima como

$$\hat{f}_k = \sum_{j=0}^{N-1} f_j \int_{x_j}^{x_{j+1}} e^{-i2k\pi x} dx - f_p \int_z^{x_{p+1}} e^{-i2k\pi x} dx + f_{p+1} \int_z^{x_{p+2}} e^{-i2k\pi x} dx + e_k(\Delta x).$$

De acordo com o Teorema 3.1, temos

$$\hat{f}_k = a_k \tilde{f}_k + (f_p - f_{p+1}) \hat{g}_k + e_k(\Delta x). \quad (3.19)$$

Analogamente, podemos facilmente obter a estimativa de  $e_k(\Delta x)$

$$\left| \int_{x_p}^z [f(x) - f_p] e^{-i2k\pi x} dx \right| \leq \int_{x_p}^z \sup_{\xi} |f'(\xi)| |x - \eta_p| dx \leq \frac{1}{2} C (z - x_p)^2 \leq \frac{1}{2} C (\Delta x)^2$$

e

$$\left| \int_z^{x_{p+2}} [f(x) - f_{p+1}] e^{-i2k\pi x} dx \right| \leq \int_z^{x_{p+2}} \sup_{\xi} |f'(\xi)| |x - \eta_{p+1}| dx \leq \frac{1}{2} C (x_{p+2} - z)^2 \leq 2C (\Delta x)^2.$$

Logo, obtemos

$$\begin{aligned}|e_k(\Delta x)| &= |\hat{f}_k - a_k \tilde{f}_k - (f_p - f_{p+1}) \hat{g}_k| \\ &\leq \sum_{j=0, j \neq p, p+1}^{N-1} \int_{x_j}^{x_{j+1}} |f(x) - f_j| dx + \int_{x_p}^z |f(x) - f_p| dx + \int_z^{x_{p+2}} |f(x) - f_{p+1}| dx \quad (3.20) \\ &\leq (N-1) \frac{1}{2} C (\Delta x)^2 + 2C (\Delta x)^2 \equiv C(f) \Delta x.\end{aligned}$$

No caso  $k = 0$ , temos

$$\int_{x_{p+1}}^z dx = z - x_{p+1} = \hat{g}_0. \quad (3.21)$$

Seguindo a prova no caso  $k \neq 0$ , podemos obter a mesma estimativa (3.20) com  $a_0 = 1$ . E completamos a prova do teorema.

Já sabemos que o erro da aproximação será menor se aproximamos  $f(x)$  pela função linear por partes. Então, daremos uma relação mais exata baseada nesse tipo de aproximação. Aqui, utilizaremos os valores da função nos nós  $x_j, j = 0, \dots, N$ , isto é, seja  $\delta = 0$  em (1.8):

$$f_j = f(x_j), \quad j = 0, \dots, N \quad (3.22)$$

e denotamos

$$f_j^- = \lim_{x \rightarrow x_j^-} f(x), \quad j = 1, \dots, N \quad (3.23)$$

e

$$f_j^+ = \lim_{x \rightarrow x_j^+} f(x), \quad j = 0, \dots, N-1. \quad (3.24)$$

**Teorema 3.3** Suponhamos que uma função  $f(x)$  satisfaz as Suposições 1,2 e 4, com  $\{x_j\}_0^N$ ,  $\{f_j^-\}_1^N$  e  $\{f_j^+\}_0^{N-1}$  definidos por (1.7), (3.23) e (3.24) respectivamente, e  $z_m \in \{x_j\}_0^N, m = 1, \dots, l$ . Então

$$|\hat{f}_k - a_k^{(1)} \tilde{f}_k^{(1)} - a_k^{(2)} \tilde{f}_k^{(2)}| \leq C(f)(\Delta x)^2, \quad (3.25)$$

onde  $\{\tilde{f}_k^{(1)}, k = -N/2, \dots, N/2-1\}$  e  $\{\tilde{f}_k^{(2)}, k = -N/2, \dots, N/2-1\}$  são as DFT de  $\{f_j^+, j = 0, \dots, N-1\}$  e  $\{f_j^-, j = 1, \dots, N\}$  respectivamente e

$$a_k^{(1)} = \begin{cases} 1/2, & k = 0, \\ \frac{N}{i2k\pi} + \frac{N^2}{(i2k\pi)^2} (e^{-i2k\pi/N} - 1), & k \neq 0, \end{cases} \quad (3.26)$$

e

$$a_k^{(2)} = \begin{cases} 1/2, & k = 0, \\ -\frac{N}{i2k\pi} e^{-i2k\pi/N} - \frac{N^2}{(i2k\pi)^2} (e^{-i2k\pi/N} - 1), & k \neq 0, \end{cases} \quad (3.27)$$

e  $C(f)$  é uma constante não-negativa que somente depende de  $f$ .

**Demonstração:** Em primeiro lugar, consideraremos a prova no caso  $k \neq 0$ . Usando as partições dos nós  $\{x_j, j = 0, 1, \dots, N\}$  definidas em (1.7) e os valores discretos  $\{f_j^-\}_1^N$  e  $\{f_j^+\}_0^{N-1}$  definidos em (3.23) e (3.24), obtemos

$$\int_{x_j}^{x_{j+1}} (x - x_{j+1}) e^{-i2k\pi x} dx = -\frac{1}{i2k\pi N} e^{-i2\pi x_j} - \frac{1}{(i2k\pi)^2} e^{-i2\pi x_j} (e^{-i2k\pi/N} - 1)$$

e

$$\int_{x_j}^{x_{j+1}} (x - x_j) e^{-i2k\pi x} dx = -\frac{1}{i2k\pi N} e^{-i2\pi x_{j+1}} - \frac{1}{(i2k\pi)^2} e^{-i2\pi x_j} (e^{-i2k\pi/N} - 1)$$

usando que

$$\int_{x_j}^{x_{j+1}} \frac{x - x_{j+1}}{x_j - x_{j+1}} e^{-i2k\pi x} dx = \frac{a_k^{(1)}}{N} e^{-i2k\pi \frac{j}{N}} \quad (3.28)$$

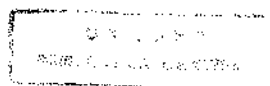
e

$$\int_{x_j}^{x_{j+1}} \frac{x - x_j}{x_{j+1} - x_j} e^{-i2k\pi x} dx = \frac{a_k^{(2)}}{N} e^{-i2k\pi \frac{j}{N}}. \quad (3.29)$$

Notemos que as igualdades (3.28) e (3.29), também são válidas se  $k = 0$ . Em seguida, utilizamos a regra dos trapézios para a aproximação da integral em (3.1) e obtemos

$$\begin{aligned} \hat{f}_k &= \sum_{j=0}^{N-1} \int_{x_j}^{x_{j+1}} f(x) e^{-i2k\pi x} dx \\ &= \sum_{j=0}^{N-1} \int_{x_j}^{x_{j+1}} \left[ \frac{f_j^+(x-x_{j+1})}{x_j-x_{j+1}} + \frac{f_{j+1}^-(x-x_j)}{x_{j+1}-x_j} \right] e^{-i2k\pi x} dx + r_k(\Delta x), \end{aligned}$$

onde  $r_k(\Delta x)$  é o termo do erro.





Usando o fato de que

$$\tilde{f}_k^{(1)} = \frac{1}{N} \sum_{j=0}^{N-1} f_j^+ e^{-i2k\pi \frac{j}{N}}, \quad -N/2 \leq k \leq N/2 - 1,$$

e

$$\tilde{f}_k^{(2)} = \frac{1}{N} \sum_{j=0}^{N-1} f_{j+1}^- e^{-i2k\pi \frac{j}{N}}, \quad -N/2 \leq k \leq N/2 - 1,$$

temos

$$\hat{f}_k = a_k^{(1)} \tilde{f}_k^{(1)} + a_k^{(2)} \tilde{f}_k^{(2)} + r_k(\Delta x). \quad (3.30)$$

Agora, estimaremos o termo  $r_k(\Delta x)$ . Suponhamos que a derivada segunda de  $f(x)$  existe no subintervalo  $(x_j, x_{j+1})$ , logo, temos que

$$f(x) = \frac{f_j^+(x - x_{j+1})}{x_j - x_{j+1}} + \frac{f_{j+1}^-(x - x_j)}{x_{j+1} - x_j} + \frac{1}{2} f''(\xi)(x - x_j)(x - x_{j+1}), \text{ para algum } \xi \in (x_j, x_{j+1})$$

para qualquer  $x \in (x_j, x_{j+1})$ , e pela Suposição 4, obtemos

$$\begin{aligned} & \left| \int_{x_j}^{x_{j+1}} \left[ f(x) - \frac{f_j^+(x - x_{j+1})}{x_j - x_{j+1}} - \frac{f_{j+1}^-(x - x_j)}{x_{j+1} - x_j} \right] e^{-i2k\pi x} dx \right| \\ & \leq \int_{x_j}^{x_{j+1}} \frac{1}{2} \sup_{\xi} |f''(\xi)| (x - x_j)^2 dx \leq \frac{1}{6} C(\Delta x)^3, \end{aligned}$$

onde  $C$  é definida na Suposição 1. A partir da desigualdade acima e usando (3.1), (3.2) e (3.26)-(3.30), obtemos

$$\begin{aligned} |r_k(\Delta x)| &= \left| \hat{f}_k - a_k^{(1)} \tilde{f}_k^{(1)} - a_k^{(2)} \tilde{f}_k^{(2)} \right| \\ &\leq \sum_{j=0}^{N-1} \int_{x_j}^{x_{j+1}} \left| f(x) - \frac{f_j^+(x - x_{j+1})}{x_j - x_{j+1}} - \frac{f_{j+1}^-(x - x_j)}{x_{j+1} - x_j} \right| dx \\ &\leq N \frac{1}{6} C(\Delta x)^3 \equiv C(f)(\Delta x)^2 \end{aligned} \quad (3.31)$$

No caso  $k = 0$ , os valores das integrais (3.28) e (3.29) são iguais a  $\frac{1}{2N}$ . Seguindo a prova no caso  $k \neq 0$ , podemos facilmente obter a mesma estimativa (3.25) com  $a_0^{(1)} = 1/2$  e  $a_0^{(2)} = 1/2$ . Então, completamos nossa prova.

Como no Teorema 3.1, suponhamos que os pontos da descontinuidade de  $f(x)$  coincidem com alguns nós  $x_j$  no Teorema 3.3. Se os pontos de descontinuidade situam em qualquer lugar em  $[0,1]$ , usando a mesma idéia do Teorema 3.2, utilizamos as novas partições não-uniformes dos nós de maneira que a segunda derivada de  $f(x)$  exista em cada subintervalo. Então, podemos generalizar o Teorema 3.3 à função com os pontos da descontinuidade situados em qualquer lugar.

Observamos que se  $f'''(x)$  existe em cada  $(x_{2j}, x_{2j+2}), j = 0, 1, \dots, N/2$ , podemos aproximar  $f(x)$  pela função quadrática e conseguimos obter uma melhor aproximação de  $\hat{f}_k$ .

Para calcular os coeficientes de Fourier  $\hat{f}_k$  de  $f(x)$  usando valores discretos de  $f(x)$  nos pontos escolhidos, já obtemos algumas expressões aproximadas de  $\hat{f}_k$  pelos Teoremas 3.1-3.3 acima. Considerando as diferentes características da função, podemos calcular eficientemente os coeficientes de Fourier de  $f(x)$  em  $[0,1]$  pela FFT.

Do Teorema 3.3, sabemos que precisamos  $\{f_j^+, j = 0, \dots, N-1\}$  e  $\{f_j^-, j = 1, \dots, N\}$  para calcular os coeficientes de Fourier  $\{\hat{f}_k\}$ . É claro que não conseguimos reconstruir  $\{f_j^+, j = 0, \dots, N-1\}$  e  $\{f_j^-, j = 1, \dots, N\}$  pela relação (3.25), se somente são dados  $\{\hat{f}_k, -N/2 \leq k \leq N/2 - 1\}$ . Então, precisamos mais coeficientes de Fourier  $\hat{f}_k$ . Nas aplicações, uma função possui apenas poucos pontos de descontinuidade. Temos  $f_k^+ = f_k^-$ , se  $f(x)$  é contínua em  $x_j$ . Por necessidade da reconstrução de funções, precisamos rescrever a relação (3.25). Por conveniência, suponhamos que uma função possui apenas um ponto de descontinuidade de  $f(x)$ , denotado por  $x_p, 1 \leq p \leq N-1$ . Então,

$$\hat{f}_k = c_k^{(1)} \bar{f}_k^{(1)} + c_k^{(2)} (f_0 - f_N) + c_k^{(3)} (f_p^+ - f_p^-) + O(\Delta x^2), \quad (3.32)$$

onde

$$c_k^{(1)} = a_k^{(1)} + e^{i\frac{2k\pi}{N}} a_k^{(2)}, \quad (3.33)$$

$$c_k^{(2)} = -e^{i\frac{2k\pi}{N}} \frac{a_k^{(2)}}{N}, \quad (3.34)$$

$$c_k^{(3)} = -e^{i\frac{2k\pi(1-p)}{N}} \frac{a_k^{(2)}}{N}. \quad (3.35)$$

Temos  $(N+2)$  valores desconhecidos  $\{f_0, \dots, f_{p-1}, f_p^-, f_p^+, f_{p+2}, \dots, f_N\}$  da função em (3.32) (incluindo especialmente os limites esquerdo e direito no ponto de descontinuidade). Se esperarmos recuperar  $N$  valores desconhecidos da função em (3.32), podemos deduzir a seguinte fórmula. Sem perda de generalidade, continuamos supondo que há apenas um ponto de descontinuidade de  $f(x)$  em  $[0,1]$ , denotado por  $z_p = x_p + \delta_p$ , onde,  $\delta_p$  é uma constante,  $0 \leq \delta_p < \frac{1}{N}$ ,  $p \in \{0, 1, \dots, N-1\}$  (Notemos que  $z_p$  pode não ser um nó).

**Teorema 3.4** Suponhamos que uma função  $f(x)$  satisfaz as Suposições 1,2 e 4.  $\{x_j\}_0^N$  e  $\{f_j\}_0^N$  definidos em (1.7) e (3.22) respectivamente. Há apenas um ponto de descontinuidade de  $f(x)$  em  $[0,1]$ , denotado por  $z_p = x_p + \delta_p$ , onde,  $\delta_p$  é uma constante,  $0 \leq \delta_p < \frac{1}{N}$ ,  $p \in \{0, 1, \dots, N-1\}$ . Se  $\delta_p=0$ , tomamos  $f_p$  como  $f_p^-$ . Então:

$$\hat{f}_k = c_k^{(1)} \tilde{f}_k + c_k^{(2)} f_1 + \overline{c_k^{(2)}} f_{N-1} + c_k^{(4)} f_p + c_k^{(5)} f_{p+1} + C(f) O(\Delta x)^2, \quad (3.36)$$

onde

$$c_k^{(4)} = \begin{cases} z_p - x_p - \frac{1}{2N}, & k = 0, \\ \frac{i}{2k\pi} (e^{-i2\pi k z_p} - e^{-i2\pi k x_p}) - \frac{1}{N} a_k^{(1)} e^{-i2\pi k \frac{z_p}{N}}, & k \neq 0. \end{cases} \quad (3.37)$$

e

$$c_k^{(5)} = \begin{cases} x_{p+1} - z_p - \frac{1}{2N}, & k = 0, \\ \frac{i}{2k\pi} (e^{-i2\pi k x_{p+1}} - e^{-i2\pi k z_p}) - \frac{1}{N} a_k^{(2)} e^{-i2\pi k \frac{z_p}{N}}, & k \neq 0. \end{cases} \quad (3.38)$$

$\overline{c_k^{(2)}}$  é o complexo conjugado de  $c_k^{(2)}$  e  $C(f)$  é uma constante não-negativa que depende de  $f$ .

Aqui, somente expressamos a fórmula aproximada de integral em (3.1) como segue:

$$\begin{aligned}\hat{f}_k &= \int_0^1 f(x)e^{-i2k\pi x} dx = \sum_{j=0, j \neq p}^{N-2} \int_{x_j}^{x_{j+1}} f(x)e^{-i2k\pi x} dx + \int_{x_p}^{x_{p+1}} f(x)e^{-i2k\pi x} dx + \int_{x_{N-1}}^{x_N} f(x)e^{-i2k\pi x} dx \\ &= \sum_{j=0, j \neq p}^{N-2} \int_{x_j}^{x_{j+1}} \left[ \frac{f_j(x-x_{j+1})}{x_j-x_{j+1}} + \frac{f_{j+1}(x-x_j)}{x_{j+1}-x_j} \right] e^{-i2k\pi x} dx + \\ &\quad \int_{x_p}^{x_p} f_p e^{-i2k\pi x} dx + \int_{x_p}^{x_{p+1}} f_{p+1} e^{-i2k\pi x} dx + \int_{x_{N-1}}^{x_N} f_{N-1} e^{-i2k\pi x} dx.\end{aligned}$$

Logo, podemos facilmente obter (3.36). Omitiremos a prova detalhada. Analisando a fórmula (3.36), achamos que os coeficientes  $c_k^{(2)}$ ,  $c_k^{(3)}$ ,  $c_k^{(4)}$  e  $c_k^{(5)}$  são pequenos, comparados com  $c_k^{(1)}$ . Isto coincide com os resultados numéricos mais adiante. Podemos pensar o termo  $c_k^{(1)} \tilde{f}_k$  como uma parte principal, e o resto como uma parte corrigida. Uma relação semelhante aparecia no Teorema 3.2, onde  $\hat{g}_k$  também é pequeno. Isto é proveitoso para nossa reconstrução no Capítulo 5, desde que implica que nosso método de reconstrução iterativo será convergente.

Neste capítulo, principalmente estabelecemos algumas relações entre a transformada de Fourier discreta  $\{\tilde{f}_k\}$  e os coeficientes de Fourier  $\{\hat{f}_k\}$  de uma função  $f(x)$ . Usando estes resultados, podemos substituir  $\{\hat{f}_k\}$  por expressões de  $\{\tilde{f}_k\}$  se for necessário, ou vice-versa.

# Capítulo 4

## Cálculo dos Coeficientes de Fourier Complexos

No Capítulo 3, estabelecemos algumas relações entre  $\{\hat{f}_k, -N/2 \leq k \leq N/2 - 1\}$  e  $\{\tilde{f}_k, -N/2 \leq k \leq N/2 - 1\}$ . Usando estas relações, apresentamos neste capítulo um algoritmo eficiente para calcular os coeficientes de Fourier complexos. Com o novo algoritmo podemos usar a transformada rápida de Fourier e chegamos a uma melhor exatidão que com os outros métodos encontrados na literatura. Mostramos alguns exemplos numéricos para ilustrar o comportamento do novo algoritmo.

### 4.1 Algoritmos Eficientes para Calcular os Coeficientes de Fourier Complexos

Já tendo as relações aproximadas (3.5), (3.17) e (3.25) obtidas no capítulo anterior, podemos facilmente calcular os coeficientes da expansão de Fourier de  $f(x)$  usando os valores da função no conjunto finito de pontos escolhidos. Apresentaremos agora três algoritmos, para aproximar  $f(x)$  satisfazendo as condições dos Teoremas 3.1, 3.2 e 3.3 respectivamente. Quando  $f(x)$  satisfaz as condições do Teorema 3.1, temos

**Algoritmo 4.1** Dados  $\{f_j, j = 0, \dots, N - 1\}$ ,

Passo 1. Calcular

$$\{\tilde{f}_k, k = -N/2, \dots, N/2 - 1\}$$

usando a FFT.

Passo 2. Calcular

$$R\hat{f}_k := a_k \tilde{f}_k \text{ para } k = -N/2, \dots, N/2 - 1,$$

onde  $a_k$  satisfaz (3.6) e  $R\hat{f}_k$  é a aproximação de  $\hat{f}_k$ .

Observamos que são necessárias apenas  $O(N \log_2 N)$  operações para aplicar a FFT no passo 1 e  $N$  multiplicações no passo 2. Então, a quantidade de cálculos neste algoritmo não é grande. Além disso, (3.6) é ainda válido para todo inteiro  $k$ , e podemos calcular todos os coeficientes de Fourier  $R\hat{f}_k$  usando  $N$  valores da função. Aqui,  $\tilde{f}_k$  será uma função periódica com período  $N$ . Resultados semelhantes, também, são válidos para os Algoritmos 4.2 e 4.3. Por outro lado, quanto maior  $N$ , menor é o erro da aproximação. Quando  $f(x)$  satisfaz as condições do Teorema 3.2, (3.17) é válido e podemos utilizar o algoritmo como segue.

**Algoritmo 4.2** Dadas as partições dos nós no intervalo  $[0,1]$  e  $\{f_j\}_0^{N-1}$  definidos em (3.14) - (3.16) respectivamente,

Passo 1. Calcular

$$\{\tilde{f}_k, k = -N/2, \dots, N/2 - 1\}$$

usando a FFT

Passo 2. calcular

$$R\hat{f}_k := a_k \tilde{f}_k + (f_p - f_{p+1})\hat{g}_k,$$

onde  $a_k$  e  $\hat{g}_k$  satisfazem (3.6) e (3.18) respectivamente.

Em geral, temos que escolher uma partição uniforme no intervalo  $[0,1]$  como (1.7). Mas, o Algoritmo 4.2 mostra que podemos escolher arbitrariamente a partição do intervalo  $[0,1]$ . Isto é uma vantagem para o caso de uma função descontínua.

De mesmo jeito, de (3.25) do Teorema 3.3 deduzimos o Algoritmo 4.3.

**Algoritmo 4.3** Dados  $\{f_j^+, j = 0, \dots, N-1\}$ ,  $\{f_j^-, j = 1, \dots, N\}$ ,

Passo 1. Calcular

$$\{\tilde{f}_k^{(1)}, k = -N/2, \dots, N/2 - 1\}$$

e

$$\{\tilde{f}_k^{(2)}, k = -N/2, \dots, N/2 - 1\}$$

usando a FFT

Passo 2. Calcular

$$R\hat{f}_k := a_k^{(1)} \tilde{f}_k^{(1)} + a_k^{(2)} \tilde{f}_k^{(2)} \quad \text{para } k = -N/2, \dots, N/2 - 1,$$

onde  $a_k^{(1)}$  e  $a_k^{(2)}$  satisfazem (3.26) e (3.27),  $R\hat{f}_k$  é a aproximação de  $\hat{f}_k$ .

Algumas observações importantes são as seguintes:

1. Nossos algoritmos são tão rápidos quanto o método da FFT, mas, obtemos uma melhor aproximação. O método da FFT precisa  $N$  pontos equidistantes igualmente espaçados, mas, o Algoritmo 4.2 mostra que isto não é necessário, desde que os pontos de descontinuidade pertençam à partição. Por outro lado, podemos calcular apenas  $N$  coeficientes de Fourier pelas  $N$  amostras usando a FFT, mas, conseguimos calcular todos os coeficientes de Fourier  $\hat{f}_k$  pelo nosso método, para qualquer inteiro  $k$ . Os exemplos mostram que o erro será grande quando  $|k|$  torna-se grande ( $N/2 \leq |k|$ ).

2. Se  $f(x) \in C^q[0,1]$ ,  $1 \leq q$ , o método dos Splines é um bom algoritmo. Mas, nosso algoritmo será melhor para uma função suave por partes. Por outro lado, são necessárias  $O(N^2)$  para  $N$  coeficientes de Fourier usando o método dos Splines quando tomamos  $N$  nós em  $[0,1]$ . Assim, a quantidade de cálculos será grande quando este método é estendido para dimensões maiores. No Algoritmo 4.3, também, utilizamos a interpolação linear da função, mas, a diferença com o método dos Splines é que não calculamos diretamente os coeficientes de Fourier da base linear. Além disso, utilizamos os limites a esquerda e a direita nos pontos de descontinuidade da função.

3. Usando a mesma idéia, podemos estender os algoritmos a dimensões maiores.

## 4.2 Exemplos Numéricos

Nesta seção, mostraremos os resultados de alguns experimentos numéricos para ilustrar o comportamento dos algoritmos deste capítulo. Todos os cálculos, nesta tese, foram realizados usando o pacote MATLAB numa estação SunSparc no IMECC, Universidade Estadual de Campinas, SP, Brasil. A precisão da máquina é  $u \sim 10^{-16}$ . O suporte da função é o intervalo  $[0,1]$  neste e nos próximos capítulos, isto é,  $f(x) = 0$ ,  $x \notin [0,1]$ .

Aplicamos os algoritmos às três funções com  $N = 128$ . Realizamos os cálculos para o exemplo 4.1 usando o Algoritmo 4.1, para o exemplo 4.2 usando os Algoritmos 4.1 e 4.2, e para o exemplo 4.3 usando os Algoritmos 4.1 e 4.3. Comparamos também, o erro entre  $R\hat{f}_k$  e  $\hat{f}_k$ , e o erro entre  $\tilde{f}_k$  e  $\hat{f}_k$ ,  $k = -N/2, \dots, N/2 - 1$ , como segue:

$$\hat{e}_k = \log_{10}|R\hat{f}_k - \hat{f}_k| \quad (4.1)$$

e

$$\tilde{e}_k = \log_{10}|\tilde{f}_k - \hat{f}_k| \quad (4.2)$$



o que é mostrado nas Figuras 4.1-4.3. Para todas as figuras, mostramos os erros como função de  $k$ .

Exemplo 4.1 Função  $f_1(x)$ :

$$f_1(x) = x, \quad x \in [0, 1] \quad (4.3)$$

Notemos que  $f_1(x)$  é linear em  $[0,1]$ .

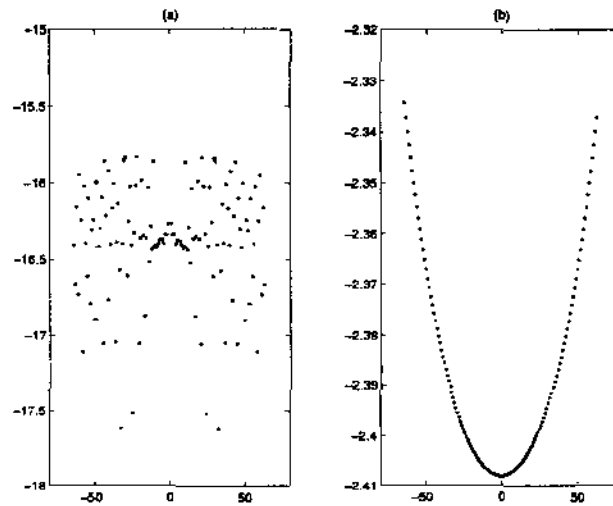


Figura 4.1: Erros:  $\hat{e}_k = \log_{10}|R\hat{f}_{1k} - \hat{f}_{1k}|$  em (a) e  $\tilde{e}_k = \log_{10}|\tilde{f}_{1k} - \hat{f}_{1k}|$  em (b)

Da Fig. 4.1 (a), observamos que o erro do nosso método é da ordem  $O(10^{-16})$  causado apenas pelo erro da máquina. Isto é devido ao fato de que  $f_1(x)$  é linear em  $[0,1]$ , verificando o Corolário 2 no Capítulo 3. Comparando a Fig. 4.1 (b) com (a), observamos que  $\tilde{e}_k$  é maior que  $\hat{e}_k$ .

Exemplo 4.2. Função  $f_2(x)$ :

$$f_2(x) = \begin{cases} x, & x \in [0, 0.5 + 1/256], \\ 0, & x \notin [0, 0.5 + 1/256]. \end{cases} \quad (4.4)$$

Em seguida, aproximamos os coeficientes da expansão de Fourier da  $f_2(x)$ , usando os Algoritmos 4.1 e 4.2.

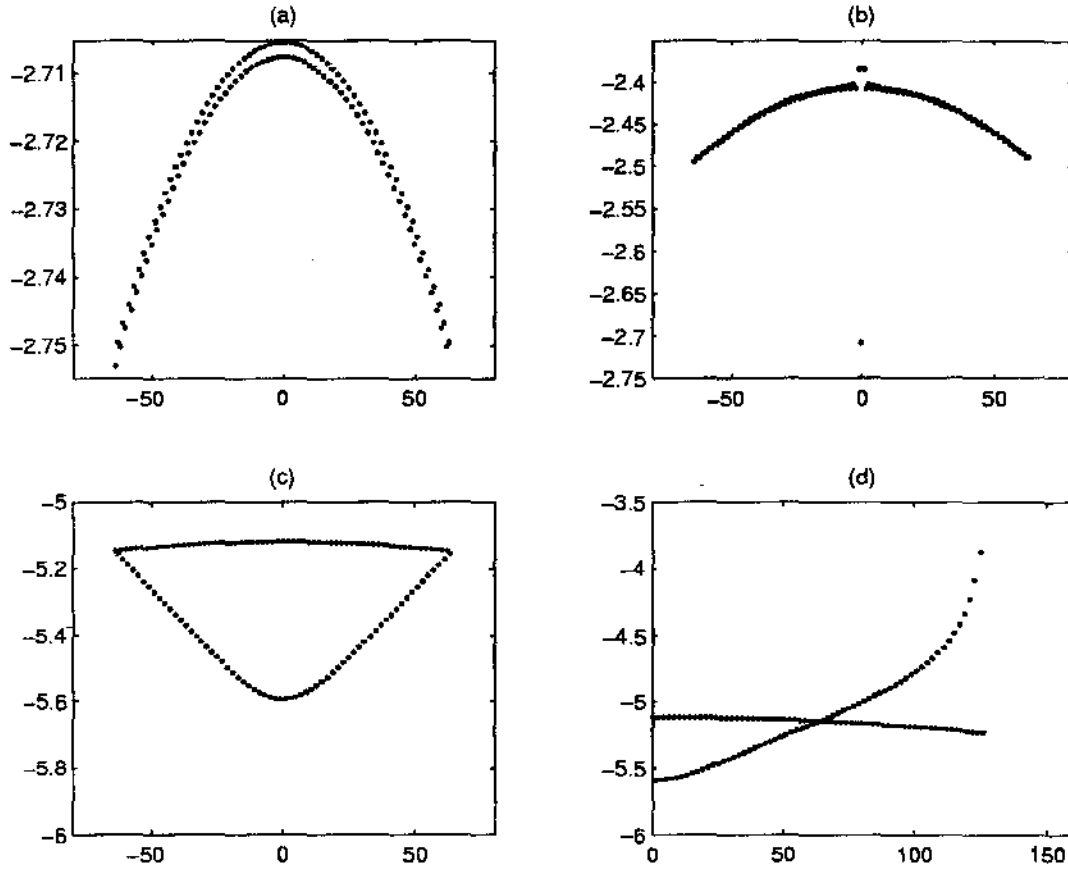


Figura 4.2: Erro  $\hat{e}_k = \log_{10}|R\hat{f}_{2k} - \hat{f}_{2k}|$  ( $R\hat{f}_{2k}$  calculado pelo Algoritmo 4.1) em (a) . Erro  $\tilde{e}_k = \log_{10}|\tilde{f}_{2k} - \hat{f}_{2k}|$  em (b). Erros  $\hat{e}_k = \log_{10}|R\hat{f}_{2k} - \hat{f}_{2k}|$  ( $R\hat{f}_{2k}$  calculado pelo Algoritmo 4.2): (c) para  $-N/2 \leq k \leq N/2 - 1$  e (d) para  $0 \leq k \leq N$ .

Se calculamos  $R\hat{f}_k$  pelo Algoritmo 4.1, a Fig. 4.2 (a) e (b) mostra que o erro  $\hat{e}_k$  é menor que  $\tilde{e}_k$ . A Fig. 4.2 (a) e (c) mostra que  $R\hat{f}_{2k}$  calculado pelo Algoritmo 4.2 é mais exato que aquele pelo Algoritmo 4.1. A Fig. 4.2 (d) mostra que podemos calcular todos os coeficientes de Fourier para qualquer  $k$ , mas, o erro é grande quanto  $k$  torna-se grande.

Exemplo 4.3. Função  $f_3(x)$ :

$$f_3(x) = \begin{cases} x^2, & x \in [0, 0.5), \\ \cos(x), & x \notin [0, 0.5). \end{cases} \quad (4.5)$$

Aplicamos os Algoritmos 4.1 e 4.3 para aproximar os coeficientes da expansão de Fourier de  $f_3(x)$ .

Se calculamos os coeficientes da expansão de Fourier  $R\hat{f}_{3k}$  pelo Algoritmo 4.1, a Fig. 4.3 (a) e (b) mostra que o erro  $\hat{e}_k$  é menor que  $\tilde{e}_k$ . A Fig. 4.3 (a) e (c) mostra que  $R\hat{f}_{3k}$  pelo Algoritmo 4.3 é mais exato que aquele pelo Algoritmo 4.1, o que coincide com os resultados do Teorema 3.3. A Fig 4.3 (d) mostra que conseguimos obter uma melhor aproximação dos coeficientes de Fourier  $\hat{f}_k$  para  $k = N/2, \dots, N$ . Achamos que o erro será grande se calculamos  $\hat{f}_k$ , para  $k = N/2, \dots, N$ , pelo Algoritmo 4.1. Isso é semelhante à Fig 4.2 (d).

Os exemplos acima ilustram a eficiência de nosso método para calcular os coeficientes da expansão de Fourier. Conseguimos uma precisão muito maior que outros métodos na literatura, e mais, o esforço computacional é pequeno, pelo fato de estar usando a FFT eficientemente.

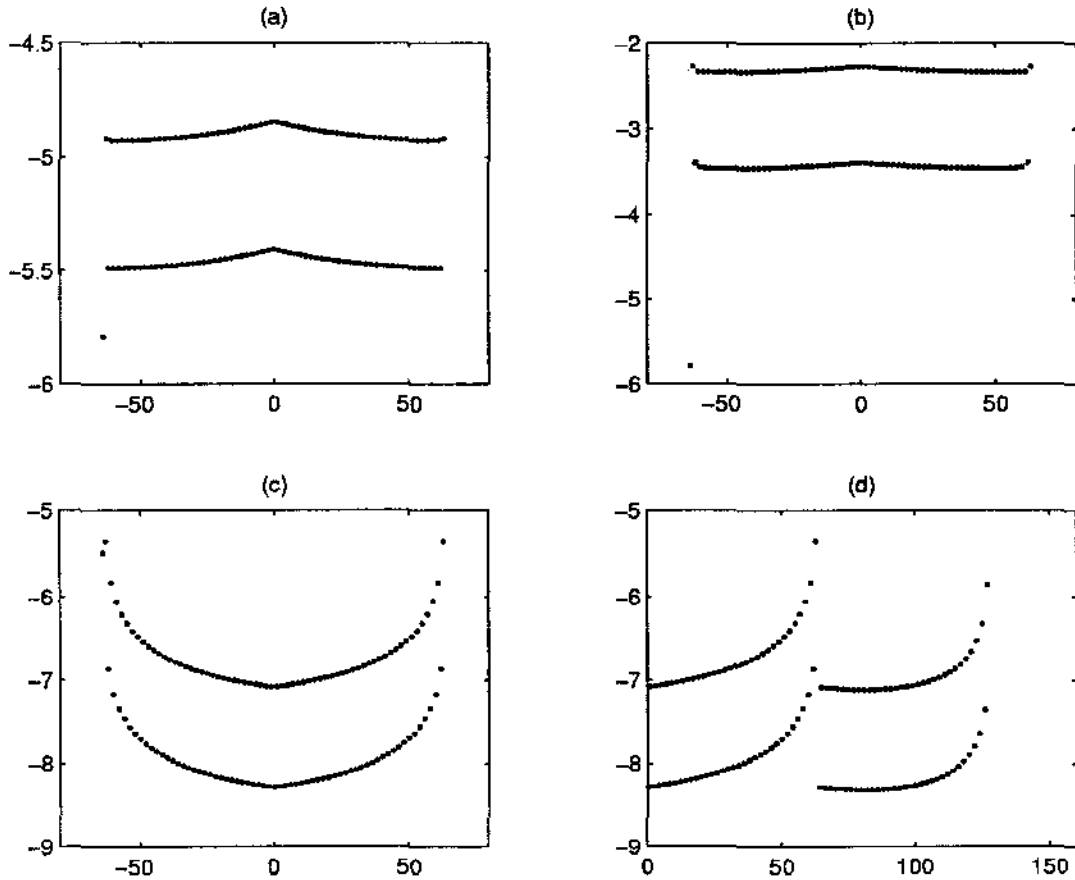


Figura 4.3: Erro  $\hat{e}_k = \log_{10}|R\hat{f}_{3k} - \hat{f}_{3k}|$  ( $R\hat{f}_{3k}$  calculado pelo Algoritmo 4.1) em (a). Erro  $\tilde{e}_k = \log_{10}|\tilde{f}_{3k} - \hat{f}_{3k}|$  em (b). Erros  $\hat{e}_k = \log_{10}|R\hat{f}_{3k} - \hat{f}_{3k}|$  ( $R\hat{f}_{3k}$  calculado pelo Algoritmo 4.3): (c) para  $-N/2 \leq k \leq N/2 - 1$  e (d) para  $0 \leq k \leq N$ .

## Capítulo 5

# A Reconstrução de $\{f_j, j = 0, 1, \dots, N - 1\}$ a Partir dos Coeficientes de Fourier $\{\hat{f}_k, k = -N/2, \dots, N/2 - 1\}$

No Capítulo 3, estabelecemos algumas relações entre  $\{\hat{f}_k, -N/2 \leq k \leq N/2 - 1\}$  e  $\{\bar{f}_k, -N/2 \leq k \leq N/2 - 1\}$ . Usando estas relações, apresentaremos agora alguns algoritmos para a reconstrução de funções. Como antes, nossos métodos permitem adotar a transformada rápida de Fourier conseguindo baixo custo computacional e mais precisão que outros algoritmos similares encontrados na literatura (particularmente nos pontos próximos dos pontos de descontinuidade). Também estimaremos o erro da função reconstruída.

### 5.1 Algoritmos para a Reconstrução de Funções

Já estabelecemos algumas relações aproximadas entre a transformada de Fourier discreta e os coeficientes de uma função  $f(x)$  no Capítulo 3, e obtivemos alguns algoritmos eficientes para calcular os coeficientes de Fourier no capítulo anterior. Mas, nosso principal objetivo é reconstruir o conjunto discreto  $\{f_j, j = 0, 1, \dots, N - 1\}$  de uma função  $f(x)$  a partir dos coeficientes de Fourier  $\{\hat{f}_k, k = -N/2, \dots, N/2 - 1\}$  (ou a transformada de Fourier de  $f(x)$  com a frequência  $k$ ). Apresentaremos diferentes algoritmos para os casos nos quais  $f(x)$  satisfaz as condições dos Teoremas 3.1, 3.2 e 3.3. Se  $f(x)$  satisfaz as condições do Teorema

3.1, então, temos

**Algoritmo 5.1** Dados  $\{\hat{f}_k, k = -N/2, \dots, N/2 - 1\}$ ,

Passo 1. Inicializar

$$\tilde{f}_k := \sigma(k)\hat{f}_k \text{ para } k = -N/2, \dots, N/2 - 1,$$

onde  $\sigma(k)$  satisfaz (3.10).

Passo 2. Calcular

$$Rf_j = \sum_{k=-N/2}^{N/2-1} \tilde{f}_k e^{i2k\pi \frac{j}{N}} \text{ para } j = 0, \dots, N - 1$$

usando a IFFT, onde  $\{Rf_j, j = 0, \dots, N - 1\}$  são os valores reconstruídos da função em  $\{\eta_j, j = 0, \dots, N - 1\}$ .

Precisamos apenas  $N$  multiplicações no passo 1 e  $O(N \log_2 N)$  para as operações da IFFT no passo 2, e então, a quantidade de cálculos não é grande neste algoritmo. Se  $f(x)$  satisfaz as condições do Teorema 3.2, então, aplicamos a fórmula (3.17), e temos

$$\sigma(k)\hat{f}_k \approx \tilde{f}_k + (f_p - f_{p+1})\sigma(k)\hat{g}_k, \quad k = -N/2, \dots, N/2 - 1$$

Denotando as transformadas de Fourier inversas de  $\{\sigma(k)\hat{f}_k\}_{-N/2}^{N/2-1}$  e  $\{\sigma(k)\hat{g}_k\}_{-N/2}^{N/2-1}$  como  $\{h_j\}_0^{N-1}$  e  $\{g_j\}_0^{N-1}$  respectivamente, obtemos

$$h_j = f_j + (f_p - f_{p+1})g_j, \quad j = 0, \dots, N - 1. \quad (5.1)$$

Para obter  $\{f_j\}_0^{N-1}$ , basta resolver o sistema de equações (5.1). Como  $\{\hat{g}_k\}$  é pequeno,  $\{g_j\}$  também o será. Claramente, o sistema de equações (5.1) tem solução. O algoritmo é descrito como segue:

**Algoritmo 5.2** Dados  $\{\hat{f}_k, k = -N/2, \dots, N/2 - 1\}$  e  $\{\hat{g}_k, k = -N/2, \dots, N/2 - 1\}$ ,

Passo 1. Inicializar

$$\tilde{h}_k := \sigma(k)\hat{f}_k \text{ para } k = -N/2, \dots, N/2 - 1,$$

$$\tilde{g}_k := \sigma(k)\hat{g}_k \text{ para } k = -N/2, \dots, N/2 - 1,$$

onde  $\sigma(k)$  satisfaz (3.10).

Passo 2. Calcular

$$h_j = \sum_{k=-N/2}^{N/2-1} \tilde{h}_k e^{i2k\pi \frac{j}{N}} \text{ para } j = 0, \dots, N - 1,$$

$$g_j = \sum_{k=-N/2}^{N/2-1} \tilde{g}_k e^{i2k\pi \frac{j}{N}} \text{ para } j = 0, \dots, N - 1,$$

pela IFFT.

Passo 3. Achar a solução do sistema de equações (5.1) como segue

Passo 3.1

$$\begin{cases} (1 - g_p)Rf_p - g_p Rf_{p+1} & = h_p \\ g_{p+1}Rf_p + (1 - g_{p+1})Rf_{p+1} & = h_{p+1} \end{cases} \quad (5.2)$$

Passo 3.2

$$Rf_j = h_j - (Rf_p - Rf_{p+1})g_j, \quad j = 0, \dots, N - 1, j \neq p, p + 1. \quad (5.3)$$

Precisamos apenas  $2N$  multiplicações no passo 1,  $O(N \log_2 N)$  operações para as duas IFFTs no passo 2, e algumas poucas operações para achar a solução do sistema de equações (5.2) e (5.3) no passo 3 (no passo 3.1, podemos achar diretamente as soluções de  $Rf_p$  e  $Rf_{p+1}$ , e em seguida, inserimos elas em (5.3)). Então, a quantidade do cálculos neste algoritmo é da mesma ordem do Algoritmo 5.1. Se há muitos pontos de descontinuidade de  $f(x)$  que possuem mesma característica como  $z$ , podemos desenvolver um algoritmo semelhante. Neste caso, precisamos mais operações por causa da IFFT no passo 2 e calculamos a solução de (5.2) e (5.3) por um método iterativo. Pela estrutura da matriz, o método iterativo converge muito rapidamente.

Usando a mesma idéia, podemos generalizar os Algoritmos 5.1 e 5.2 para dimensões maiores. Mostraremos alguns exemplos em duas-dimensões.

Para obter uma melhor reconstrução, precisamos uma relação mais exata entre  $\hat{f}_k$  e  $\tilde{f}_k$  como (3.32) e (3.36). Se  $f(x)$  satisfaz as condições do Teorema 3.3, e supondo que há apenas um ponto de descontinuidade de  $f(x)$ ,  $x_p, 1 \leq p \leq N-1$ , então, aplicamos a fórmula (3.32), e temos

$$\tilde{f}_k^{(1)} \approx \frac{\hat{f}_k}{c_k^{(1)}} - \frac{c_k^{(2)}}{c_k^{(1)}}(f_0 - f_N) - \frac{c_k^{(3)}}{c_k^{(1)}}(f_p^+ - f_p^-). \quad (5.4)$$

Denotando as transformadas de Fourier inversas de  $\frac{\hat{f}_k}{c_k^{(1)}}$ ,  $\frac{c_k^{(2)}}{c_k^{(1)}}$  e  $\frac{c_k^{(3)}}{c_k^{(1)}}$  por  $\{h_j^{(0)}\}_0^{N-1}$ ,  $\{h_j^{(1)}\}_0^{N-1}$  e  $\{h_j^{(2)}\}_0^{N-1}$ , respectivamente, obtemos

$$\begin{cases} f_j = h_j^{(0)} - h_j^{(1)}(f_0 - f_N) - h_j^{(2)}(f_p^+ - f_p^-), & j = 0, \dots, N-1, j \neq p, \\ f_p^+ = h_p^{(0)} - h_p^{(1)}(f_0 - f_N) - h_p^{(2)}(f_p^+ - f_p^-). \end{cases} \quad (5.5)$$

Agora, temos  $(N+2)$  variáveis desconhecidas  $\{f_j, j = 0, \dots, N, j \neq p, f_p^-, f_p^+\}$ . Mas, apenas  $N$  equações em (5.5), e ainda precisamos outras duas equações. Sejam  $k = N/2, N/2 + 1$  em (5.4), temos

$$\hat{f}_{N/2} = c_{N/2}^{(1)} \tilde{f}_{N/2}^{(1)} + c_{N/2}^{(2)}(f_0 - f_N) + c_{N/2}^{(3)}(f_p^+ - f_p^-) \quad (5.6)$$

e

$$\hat{f}_{N/2+1} = c_{N/2+1}^{(1)} \tilde{f}_{N/2+1}^{(1)} + c_{N/2+1}^{(2)}(f_0 - f_N) + c_{N/2+1}^{(3)}(f_p^+ - f_p^-). \quad (5.7)$$

Como

$$\hat{f}_{N/2} = \overline{\hat{f}_{-N/2}},$$

$$\tilde{f}_{N/2}^{(1)} = \tilde{f}_{-N/2}^{(1)}$$

e



$$\tilde{f}_{N/2+1}^{(1)} = \tilde{f}_{-N/2+1}^{(1)},$$

temos

$$\hat{f}_{-N/2} = c_{N/2}^{(1)} \tilde{f}_{-N/2}^{(1)} + c_{N/2}^{(2)} (f_0 - f_N) + c_{N/2}^{(3)} (f_p^+ - f_p^-) \quad (5.8)$$

e

$$\hat{f}_{N/2+1} = c_{N/2+1}^{(1)} \tilde{f}_{-N/2+1}^{(1)} + c_{N/2+1}^{(2)} (f_0 - f_N) + c_{N/2+1}^{(3)} (f_p^+ - f_p^-). \quad (5.9)$$

Notando que as duas equações acima são complexas, pegaremos então a parte real da primeira e a parte imaginária da segunda. Finalmente, podemos achar as  $(N + 2)$  variáveis desconhecidas a partir das  $(N + 2)$  equações.

**Algoritmo 5.3** Dados  $\{\hat{f}_k, k = -N/2, \dots, N/2 + 1\}$ ,  $\{c_k^{(1)}, k = -N/2, \dots, N/2 - 1\}$ ,  $\{c_k^{(2)}, k = -N/2, \dots, N/2 - 1\}$  e  $\{c_k^{(3)}, k = -N/2, \dots, N/2 - 1\}$ ,

Passo 1. Inicializar

$$\tilde{h}_k^{(1)} = \frac{\hat{f}_k}{c_k^{(1)}}, \quad k = -N/2, \dots, N/2 - 1,$$

$$\tilde{h}_k^{(2)} = \frac{c_k^{(2)}}{c_k^{(1)}}, \quad k = -N/2, \dots, N/2 - 1$$

e

$$\tilde{h}_k^{(3)} = \frac{c_k^{(3)}}{c_k^{(1)}}, \quad k = -N/2, \dots, N/2 - 1.$$

Passo 2. Calcular

$$h_j^{(1)} = \sum_{k=-N/2}^{N/2-1} \tilde{h}_k^{(1)} e^{i2k\pi \frac{j}{N}}, \quad j = 0, \dots, N - 1,$$

$$h_j^{(2)} = \sum_{k=-N/2}^{N/2-1} \tilde{h}_k^{(2)} e^{i2k\pi \frac{j}{N}}, \quad j = 0, \dots, N-1$$

e

$$h_j^{(3)} = \sum_{k=-N/2}^{N/2-1} \tilde{h}_k^{(3)} e^{i2k\pi \frac{j}{N}}, \quad j = 0, \dots, N-1,$$

pela IFFT.

Passo 3. Achar as soluções  $\{f_j, j = 0, \dots, N, j \neq p, f_p^-, f_p^+\}$  do sistema de equações (5.5), (5.8) e (5.9) por um método iterativo.

Como na resolução de (5.1), podemos afirmar que o método iterativo terá convergência rápida e obterá uma solução em poucos passos.

Quando utilizamos o resultado do Teorema 3.4, podemos obter um sistema de equações semelhantes ao sistema de equações (5.5), (5.8) e (5.9) e também utilizar um método iterativo para achar a solução .

Para todos os nossos algoritmos, podemos utilizar a FFT, e portanto, o esforço computacional será pequeno.

## 5.2 Estimativa do Erro da Função Reconstruída

Discutiremos a seguir o erro da reconstrução. Em primeiro lugar, descrevemos o gráfico da função do núcleo  $K_N(\xi)$

$$K_N(\xi) = \sum_{k=-N/2}^{N/2} \sigma(k) e^{i2k\pi\xi}, \quad 0 \leq \xi \leq 1, \quad (5.10)$$

que é uma função periódica com período 1. Ela não possui a mesma característica de (2.18) no Capítulo 2, isto é, para  $\delta > 0$ , e todo  $\epsilon > 0$  existe um inteiro  $N(\delta, \epsilon) > 0$  tal que  $|K_N(\xi)| < \epsilon$ , se  $N > N(\delta, \epsilon)$  e  $\delta \leq \xi < 1 - \delta$ . Mas, veremos que os valores da função  $K_N(\xi)$  serão grandes somente em poucos subintervalos, e pequenos (não suficientemente pequenos) nos outros subintervalos. A Fig. 5.1 mostra esta característica de  $K_N(\xi)$ .

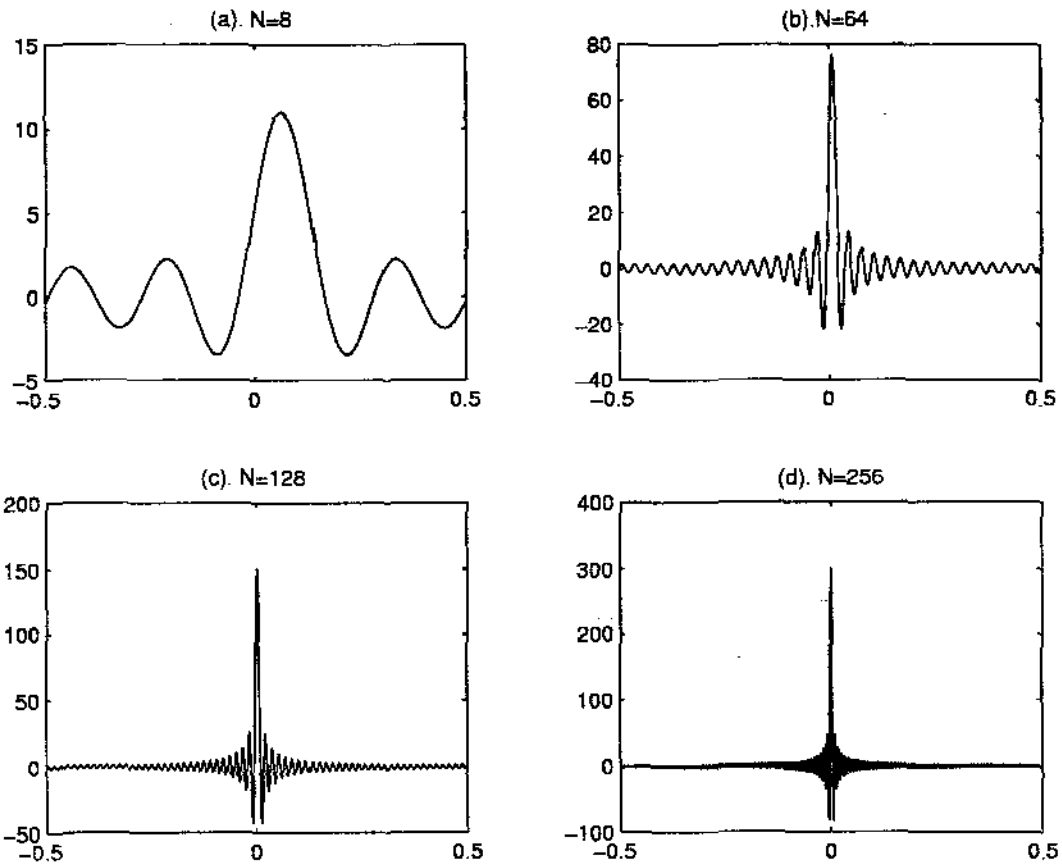


Figura 5.1: Função do núcleo  $K_N(\xi)$  para diferentes valores de  $N$  ( $N = 8, 64, 128$  e  $256$ ).

Da Fig. 5.1, podemos provar que existem algumas constantes  $M$ ,  $c_1$  e  $c_2$  tais que

$$\begin{aligned} |K_N(x)| &\leq c_1 N, & \text{para } x \in [x_l, x_{l+1}], l = 0, \dots, M-1 \text{ e } l = N-M, \dots, N-1, \\ |K_N(x)| &\leq c_2, & \text{para } x \in [x_l, x_{l+1}], l = M, \dots, N-M-1. \end{aligned} \quad (5.11)$$

Baseados nesta propriedade, podemos chegar à seguinte estimativa do erro da função reconstruída.

**Teorema 5.1** Suponha que uma função  $f(x)$  satisfaz as condições do Teorema 3.1. Então

$$|f_j - Rf_j| \leq C(f)\Delta x, \quad j = 0, \dots, N-1 \quad (5.12)$$

onde  $C(f)$  é uma constante não-negativa que depende somente de  $f$ .

**Demonstração:** De (3.8) e (3.9), obtemos

$$\tilde{f}_k = \sigma(k)\hat{f}_k + \sigma(k) \sum_{l=0}^{N-1} \int_{x_l}^{x_{l+1}} f'(\xi_l)(x - \eta_l)e^{-i2k\pi x} dx, \quad (5.13)$$

para  $-N/2 \leq k \leq N/2 - 1$ . Inserindo (5.12) em (3.3), temos que

$$\begin{aligned} f_j &= \sum_{k=-N/2}^{N/2-1} \tilde{f}_k e^{i2k\pi \frac{j}{N}} \\ &= \sum_{k=-N/2}^{N/2-1} \sigma(k)\hat{f}_k e^{i2k\pi \frac{j}{N}} + \sum_{k=-N/2}^{N/2-1} \sigma(k) \left[ \sum_{l=0}^{N-1} \int_{x_l}^{x_{l+1}} f'(\xi_l)(x - \eta_l)e^{-i2\pi kx} dx \right] e^{i2k\pi \frac{j}{N}}. \end{aligned}$$

Pelo Algoritmo 5.1, o primeiro termo é  $Rf_j$ , e denotamos o segundo termo (do erro) como  $e_j$ , então

$$f_j = Rf_j + e_j.$$

Agora, estimaremos o termo  $e_j$ , onde

$$\begin{aligned} e_j &= \sum_{k=-N/2}^{N/2-1} \sigma(k) \left[ \sum_{l=0}^{N-1} \int_{x_l}^{x_{l+1}} f'(\xi_l)(x - \eta_l)e^{-i2k\pi x} dx \right] e^{i2k\pi \frac{j}{N}} \\ &= \sum_{l=0}^{N-1} \int_{x_l}^{x_{l+1}} f'(\xi_l)(x - \eta_l) \sum_{k=-N/2}^{N/2-1} \sigma(k)e^{-i2k\pi(x - \frac{j}{N})} dx \\ &= e_j^{(1)} + e_j^{(2)}. \end{aligned}$$

Denotamos

$$e_j^{(1)} = \sum_{l=0}^{N-1} \int_{x_l}^{x_{l+1}} f'(\xi_l)(x - \eta_l) K_N^j(x) dx,$$

onde

$$K_N^j(x) = \sum_{k=-N/2}^{N/2} \sigma(k) e^{-i2\pi k(x - \frac{j}{N})} = K_N(-x + \frac{j}{N}) = K_N(x - \frac{j}{N}) \quad (5.14)$$

e

$$e_j^{(2)} = - \sum_{l=0}^{N-1} \int_{x_l}^{x_{l+1}} f'(\xi_l)(x - \eta_l) \sigma(N/2) e^{-iN\pi(x - \frac{j}{N})} dx.$$

Como em (3.9), podemos obter a estimativa de  $e_j^{(2)}$ :

$$|e_j^{(2)}| \leq \frac{C}{2} \Delta x, \text{ para } j.$$

Seguidamente, estimaremos  $e_j^{(1)}$ . Já sabemos que  $K_N(x)$  é uma função periódica com período 1. Então, todas as funções  $\{K_N^j(x), j = 0 \dots, N - 1\}$  possuem a mesma característica como  $K_N^0(x) = K_N(x)$ . Sem perda da generalidade, somente estimaremos  $e_0^{(1)}$ ; as outras estimativas são semelhantes,

$$e_0^{(1)} = \sum_{l=0}^{N-1} \int_{x_l}^{x_{l+1}} f'(\xi_l)(x - \eta_l) K_N(x) dx.$$

É claro que

$$\int_{x_l}^{x_{l+1}} |x - \eta_l| dx \leq 1/2(\Delta x)^2,$$

para  $l = 0, \dots, N - 1$ . Pela Suposição 3 no Capítulo 1, temos  $|f'(\xi_l)| \leq C$ , e por (5.11), temos que

$$\begin{aligned}
|e_0^{(1)}| &\leq \sup_{\xi \in [0,1]} f'(\xi) \sum_{l=0}^{N-1} \sup_{x \in [x_l, x_{l+1}]} |K_N(x)| \int_{x_l}^{x_{l+1}} |x - \eta_l| dx \\
&= \sup_{\xi \in [0,1]} f'(\xi) \left( \sum_{l=0}^{M-1} + \sum_{l=M}^{N-M-1} + \sum_{l=N-M}^{N-1} \right) \sup_{x \in [x_l, x_{l+1}]} |K_N(x)| \int_{x_l}^{x_{l+1}} |x - \eta_l| dx \\
&\leq C[2Mc_1 N(\Delta x)^2 + (N - 2M)c_2(\Delta x)^2].
\end{aligned}$$

Como  $\Delta x = 1/N$ , temos que

$$|e_0| \leq |e_0^{(1)}| + |e_0^{(2)}| \leq C(f)\Delta x,$$

onde  $C(f) = C[2Mc_1 + c_2] + C/2$  é independente de  $N$ .

Completamos assim nossa prova. Observamos que  $0 \leq 2Mc_1 + c_2 \leq 6$  se  $N = 128$ .

O Teorema 5.1 mostra o erro entre  $f_j$  e  $Rf_j$ ,  $j = 0, \dots, N-1$ . Usando os outros métodos existentes de filtragem, eles chegam a uma melhor reconstrução com filtros de ordem alta somente nos pontos que estão longe da descontinuidade. Nós podemos alcançar uma melhor aproximação dos valores da função em todos  $N$  pontos, incluindo os muitos próximos da descontinuidade. Analogamente, podemos conseguir estimativas semelhantes para outros Algoritmos.

### 5.3 Exemplos Numéricos

Apresentaremos alguns exemplos para ilustrar a reconstrução dos valores da função. De acordo com as diferentes características da função, efetuaremos o cálculo pelos Algoritmos 5.1, 5.2 e 5.3 respectivamente. Tomamos  $N$  como 64, 128 e 256 respectivamente. Usando os diferentes filtros  $\sigma(k)$  e  $\sigma_m(\frac{2k}{N})$ ,  $m = 1, 2, 3, 6$ , compararemos o erro em cada ponto pelo método da reconstrução. Escolhemos  $p = 10$  em  $\sigma_6(\frac{2k}{N})$ . Além disso, calcularemos o erro médio quadrático definido por:

$$e = \sqrt{\frac{1}{N} \sum_{j=0}^{N-1} |f_j - Rf_j|^2}. \quad (5.15)$$

Em todos os exemplos, calcularemos  $\{f_j = f(x_j + \Delta x/2), j = 0, \dots, N-1\}$  quando utilizamos os Algoritmos 5.1 e 5.2, e  $\{f_j = f(x_j), j = 0, \dots, N-1\}$  quando utilizamos o Algoritmo 5.3.

Exemplo 5.1 Função  $f_1(x)$ :

$$f_1(x) = x, \quad x \in [0, 1]. \quad (5.16)$$

Notemos que no intervalo  $(0, 1)$ ,  $f_1(x)$  é analítica e  $f_1(0) \neq f_1(1)$ .

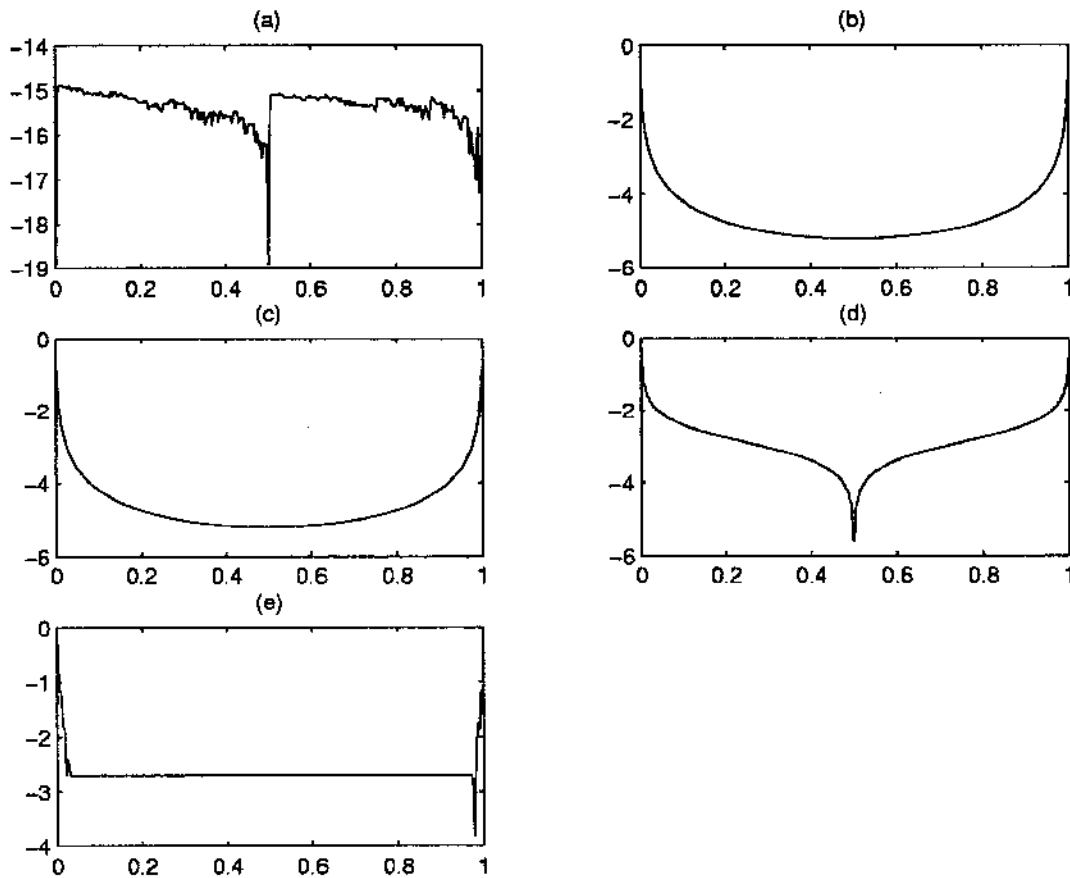


Figura 5.2: Erro  $\log_{10}|Rf_1(x) - f_1(x)|$  em cada ponto com  $N = 256$  e diferentes filtros:  $\{\sigma(k)\}$  em (a) e  $\{\sigma_m(\frac{2k}{N}), m = 1, 2, 3, 6\}$  em (b)-(e).

As Fig. 5.2 (a)-(e) mostram os gráficos do erro em cada ponto usando os filtros  $\{\sigma(k)\}$  e  $\{\sigma_m(\frac{2k}{N}), m = 1, 2, 3, 6\}$  respectivamente com  $N = 256$ .

Da Fig. 5.2, observamos que o erro da função reconstruída é quase  $O(10^{-16})$  causado apenas pelo erro da máquina. Por causa de que  $f_1(x)$  é linear em  $[0,1]$ , isto é assim no exemplo 4.1.

A Tabela 5.1 mostra o erro médio quadrático para os métodos de reconstrução usando os diferentes filtros com  $N = 64, 128$  e  $256$ .

Tabela 5.1 Erro médio quadrático para  $f_1(x)$

$N$	$\sigma(k)$	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_6$
64	$1.7408 \times 10^{-16}$	$2.0253 \times 10^{-2}$	$2.3987 \times 10^{-2}$	$4.8369 \times 10^{-2}$	$6.4480 \times 10^{-2}$
128	$2.0529 \times 10^{-16}$	$1.4317 \times 10^{-2}$	$1.6957 \times 10^{-2}$	$3.4700 \times 10^{-2}$	$4.5760 \times 10^{-2}$
256	$6.0281 \times 10^{-16}$	$1.0123 \times 10^{-2}$	$1.1990 \times 10^{-2}$	$2.4717 \times 10^{-2}$	$3.2416 \times 10^{-2}$

Exemplo 5.2. Função  $f_4(x)$ :

$$f_4(x) = \begin{cases} 1, & x \in I \equiv [0, 1/4) \cup [1/2, 5/8) \cup [3/4, 7/8), \\ 0, & \text{caso contrário.} \end{cases} \quad (5.17)$$

Notemos que  $f_4(x)$  é uma função com 6 pontos de descontinuidade.

Na Fig. 5.3, (a)-(e) são os gráficos do erro em cada ponto usando os diferentes filtros  $\{\sigma(k)\}$  e  $\{\sigma_m(\frac{2k}{N}), m = 1, 2, 3, 6\}$  respectivamente com  $N = 256$ .

Da Fig. 5.3, observamos que nosso erro da função reconstruída é quase  $O(10^{-16})$  causado apenas pelo erro da máquina, o que verifica o Corolário 1 do Capítulo 3.

A Tabela 5.2 mostra o erro médio quadrático para os métodos de reconstrução usando os diferentes filtros com  $N = 64, 128$  e  $256$ .

Tabela 5.2 Erro médio quadrático para  $f_4(x)$

$N$	$\sigma(k)$	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_6$
64	$1.1510 \times 10^{-15}$	$4.9144 \times 10^{-2}$	$5.8302 \times 10^{-2}$	$1.1888 \times 10^{-1}$	$1.5922 \times 10^{-1}$
128	$1.7342 \times 10^{-15}$	$3.4986 \times 10^{-2}$	$4.1457 \times 10^{-2}$	$8.5176 \times 10^{-2}$	$1.1250 \times 10^{-1}$
256	$2.4355 \times 10^{-15}$	$2.4781 \times 10^{-2}$	$2.9355 \times 10^{-2}$	$6.0614 \times 10^{-2}$	$7.9548 \times 10^{-2}$



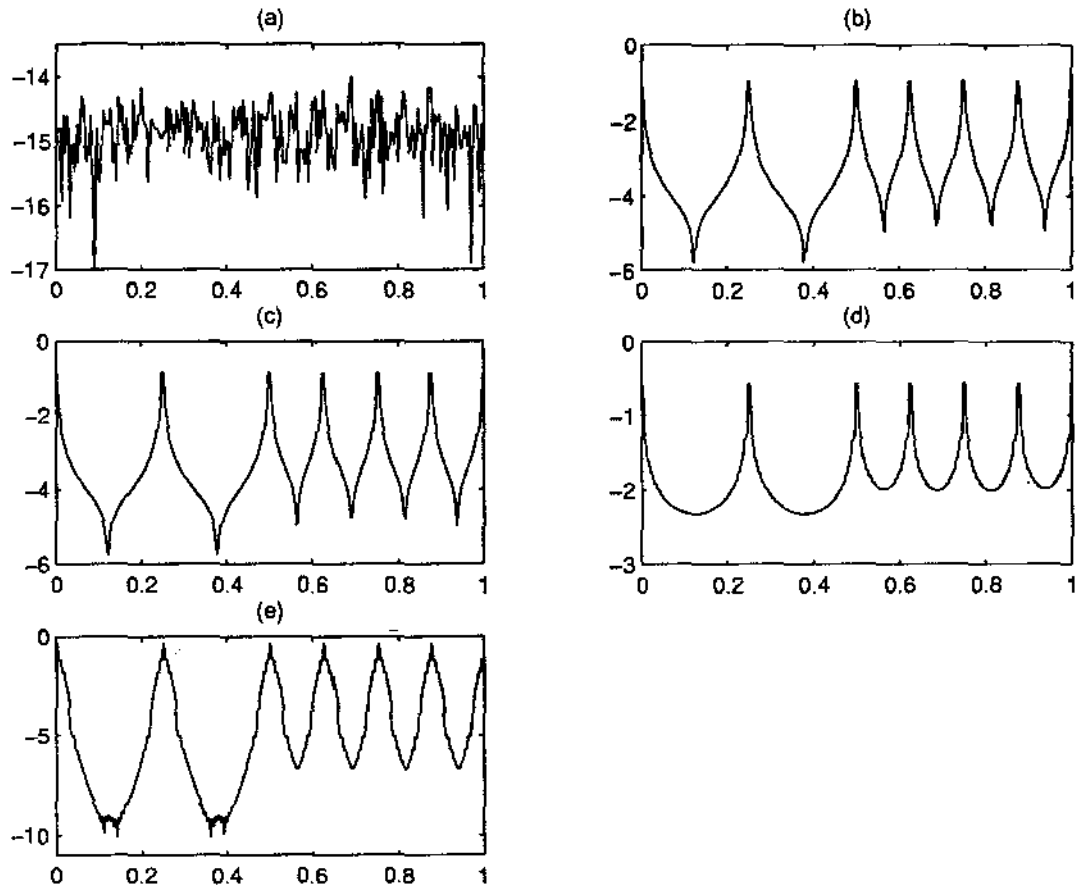


Figura 5.3: Erro  $\log_{10}|Rf_4(x) - f_4(x)|$  em cada ponto com  $N = 256$  e diferentes filtros:  $\{\sigma(k)\}$  em (a) e  $\{\sigma_m(\frac{2k}{N}), m = 1, 2, 3, 6\}$  em (b)-(e).

Exemplo 5.3. Função  $f_5(x)$ :

$$f_5(x) = x^2, x \in [0, 1]. \quad (5.18)$$

Notemos que no intervalo  $(0, 1)$ ,  $f_5(x)$  é analítica e  $f_5(0) \neq f_5(1)$ .

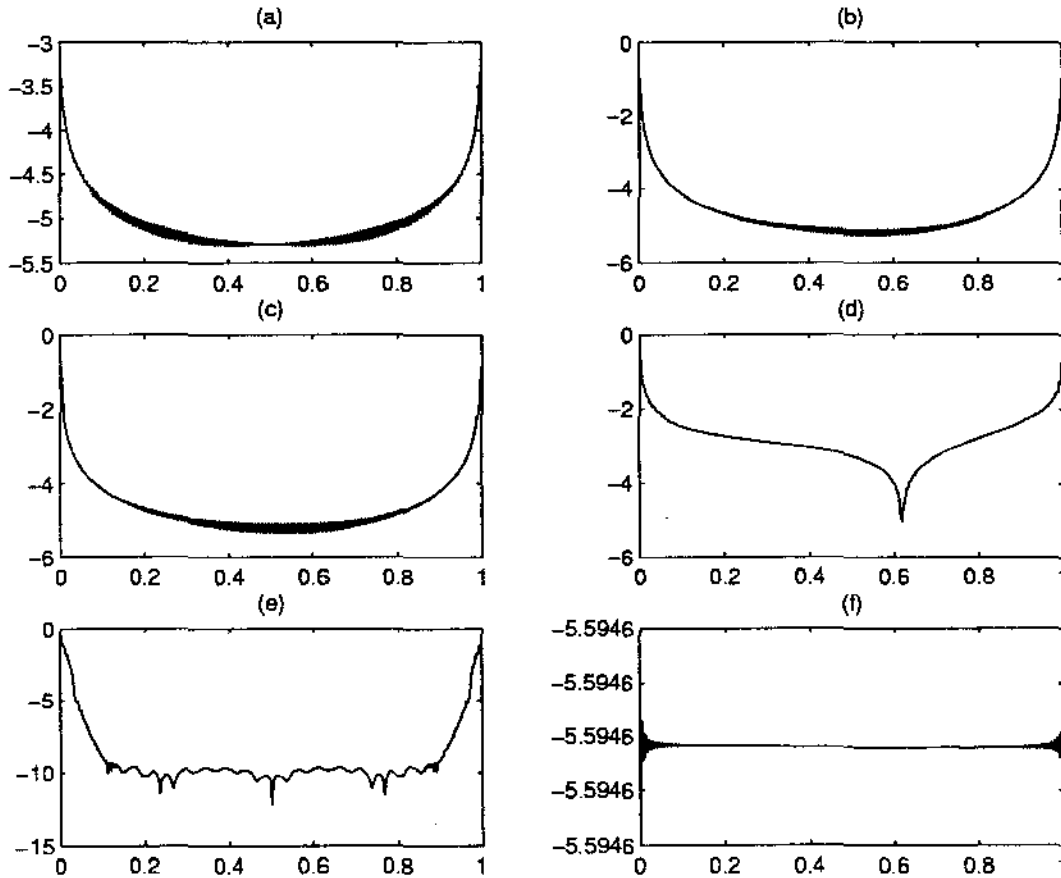


Figura 5.4: Erro  $\log_{10}|Rf_5(x) - f_5(x)|$  em cada ponto com  $N = 256$  e diferentes filtros:  $\{\sigma_k\}$  em (a) e  $\{\sigma_m(\frac{2k}{N}), m = 1, 2, 3, 6\}$  em (b)-(e). (f) mostra o erro pelo Algoritmo 5.3.

Na Fig. 5.4, (a)-(e) são os gráficos do erro em cada ponto usando os filtros  $\{\sigma(k)\}$  e  $\{\sigma_m(\frac{2k}{N}), m = 1, 2, 3, 6\}$  respectivamente com  $N = 256$ . (f) mostra o erro em cada ponto pelo Algoritmo 5.3.

Da Fig. 5.4, observamos que o nosso erro da função reconstruída usando o filtro  $\sigma(k)$  nos pontos próximos de  $x = 0$  e  $x = 1$  é bem menor que aquele erro usando quaisquer

outros filtros. Então, podemos reduzir fortemente as oscilações do fenômeno de Gibbs na função reconstruída. Notemos que já omitimos os pontos  $\eta_0$  e  $\eta_{N-1}$  quando calculamos o erro médio quadrático de  $f_1(x)$ ,  $f_4(x)$  e  $f_5(x)$  pelo filtro  $\sigma_6$ . Por outro lado, podemos reconstruir  $(N + 1)$  valores da função se utilizamos o Algoritmo 5.3. Se  $N = 256$ , os valores reconstruídos das fronteiras são  $Rf_5(x_0) = -2.5431 \times 10^{-6}$  e  $Rf_5(x_N) = 9.9999 \times 10^{-1}$ . Os valores verdadeiros são 0 e 1 respectivamente. Quando utilizamos o filtro  $\sigma_6$ , obtemos apenas  $Rf_5(x_0 + \Delta x/2) = 4.9833 \times 10^{-1}$  e  $Rf_5(x_{N-1} + \Delta x/2) = 4.9874 \times 10^{-1}$ . Seus valores verdadeiros são  $1.5259 \times 10^{-5}$  e  $9.9220 \times 10^{-1}$  respectivamente.

A Tabela 5.3 mostra o erro médio quadrático para os métodos de reconstrução usando os diferentes filtros com  $N = 64, 128$  e  $256$ .

Tabela 5.3 Erro médio quadrático para  $f_3(x)$

$N$	$\sigma(k)$	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_6$	Algoritmo 5.3
64	$4.14 \times 10^{-4}$	$2.02 \times 10^{-2}$	$2.40 \times 10^{-2}$	$4.84 \times 10^{-2}$	$1.76 \times 10^{-2}$	$4.06 \times 10^{-5}$
128	$1.46 \times 10^{-4}$	$1.43 \times 10^{-2}$	$1.69 \times 10^{-2}$	$3.47 \times 10^{-2}$	$1.24 \times 10^{-2}$	$1.01 \times 10^{-5}$
256	$5.19 \times 10^{-5}$	$1.01 \times 10^{-2}$	$1.20 \times 10^{-2}$	$2.47 \times 10^{-2}$	$8.83 \times 10^{-3}$	$2.54 \times 10^{-6}$

Também, fizemos a simulação para o filtro  $\sigma_6$  com diferentes valores de  $p$  ( $p=2,4,7,10$ ). Acharmos que o erro da função reconstruída é bem menor nos pontos relativamente afastados dos pontos de descontinuidade (ou dos pontos da fronteira), quando tomamos  $p$  grande. Mas, o erro da função reconstruída é maior nos pontos próximos às descontinuidades, quando a ordem  $p$  aumenta. A Tabela 5.4 mostra o erro da função reconstruída nos 6 pontos próximos dos pontos da fronteira. A Fig. 5.5 mostra os gráficos do erro da função reconstruída pelos filtros  $\sigma_6$  com diferentes valores de  $p$  ( $p = 2, 4, 7, 10$ ).

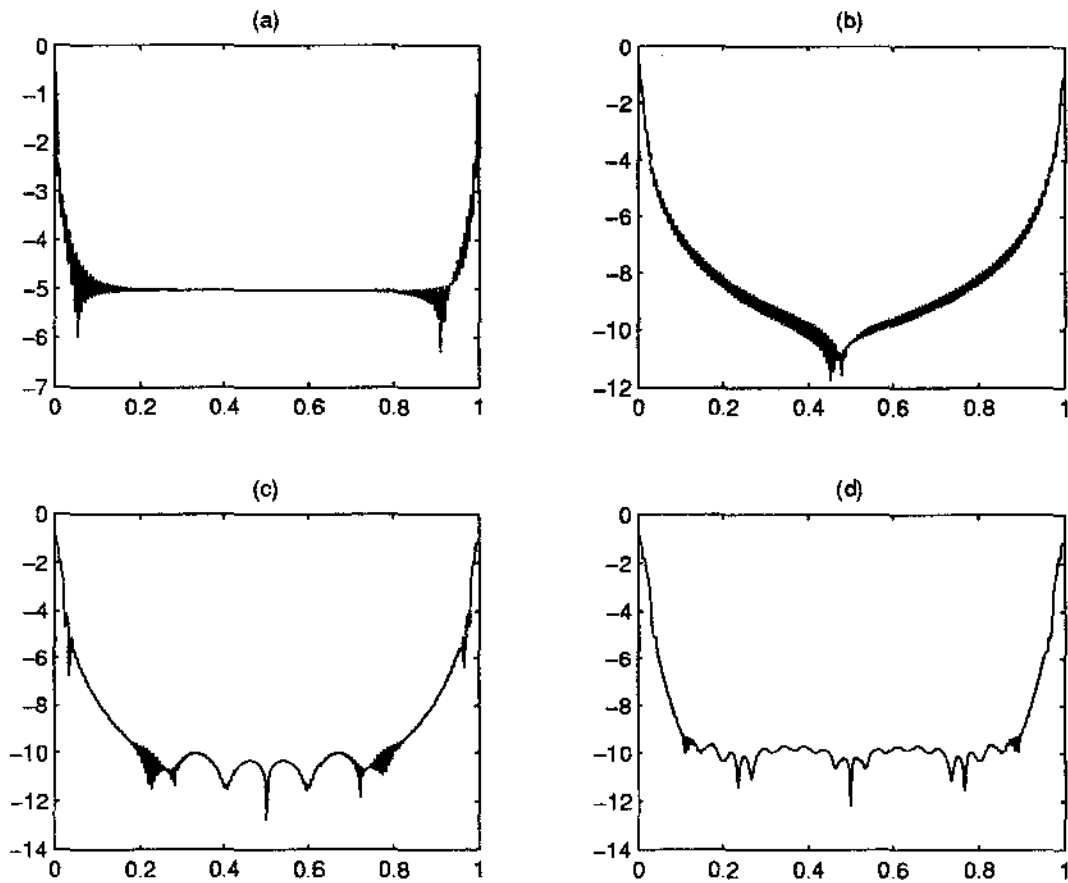


Figura 5.5: Erro  $\log_{10}|Rf_5(x) - f_5(x)|$  em cada ponto usando  $\sigma_6$  com  $N = 256$  e diferentes valores de  $p$ : (a)-(d) são para  $p=2,4,7$  e  $10$  respectivamente.

Tabela 5.4 Erros da função reconstruída nos 6 pontos

$p$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$
2	$9.37 \times 10^{-2}$	$2.09 \times 10^{-3}$	$3.71 \times 10^{-3}$	$3.00 \times 10^{-4}$	$8.16 \times 10^{-4}$	$8.71 \times 10^{-5}$
4	$8.10 \times 10^{-2}$	$3.82 \times 10^{-2}$	$1.07 \times 10^{-2}$	$1.18 \times 10^{-3}$	$9.04 \times 10^{-4}$	$1.54 \times 10^{-4}$
7	$7.44 \times 10^{-2}$	$5.77 \times 10^{-2}$	$1.47 \times 10^{-2}$	$8.14 \times 10^{-3}$	$2.20 \times 10^{-3}$	$2.31 \times 10^{-5}$
10	$7.15 \times 10^{-2}$	$6.64 \times 10^{-2}$	$1.52 \times 10^{-2}$	$1.55 \times 10^{-2}$	$4.00 \times 10^{-3}$	$2.09 \times 10^{-3}$

Exemplo 5.4. Função  $f_6(x)$ :

$$f_6(x) = \begin{cases} x^2, & x \in [0, 0.5), \\ \cos(x), & x \notin [0, 0.5). \end{cases} \quad (5.19)$$

Neste caso, no intervalo  $(0, 1)$ , há um ponto de descontinuidade de  $f_6(x)$ .

A Fig. 5.6, (a)-(e) mostra os gráficos do erro em cada ponto usando os filtros  $\{\sigma(k)\}$  e  $\{\sigma_m(\frac{2k}{N}), m = 1, 2, 3, 6\}$  respectivamente com  $N = 256$ . (f) mostra o erro em cada ponto para o Algoritmo 5.3.

Na Fig. 5.6, observamos que o erro da reconstrução usando o filtro  $\sigma(k)$  nos pontos próximos de  $x = 0$ ,  $x = 1$  e  $x = 1/2$  é bem menor que usando quaisquer dos outros filtros. Então, podemos reduzir fortemente as oscilações do fenômeno de Gibbs na função reconstruída. Notemos que omitimos os pontos  $\{\eta_j$  (ou  $x_j$ ),  $j = 0, N/2, N - 1\}$  quando calculamos o erro médio quadrático de  $f_5(x)$  pelo filtro  $\sigma_6$ . Além disso, podemos reconstruir  $(N + 2)$  valores da função quando utilizamos o Algoritmo 5.3. Os valores nas fronteiras e os limites a esquerda e a direita no ponto de descontinuidade da função são  $Rf_6(x_0) = -2.8830 \times 10^{-6}$ ,  $Rf_6(x_N) = 5.4030 \times 10^{-1}$ ,  $Rf_6(x_{N/2}^-) = 2.5000 \times 10^{-1}$  e  $Rf_6(x_{N/2}^+) = 8.7760 \times 10^{-1}$ , e seus valores verdadeiros são  $0$ ,  $\cos(1) \approx 5.4030 \times 10^{-1}$ ,  $0.25$  e  $\cos(1/2) \approx 8.7758 \times 10^{-1}$  respectivamente. Se utilizamos o Algoritmo 5.1 pelo filtro  $\sigma_6$ ,  $Rf_6(x_0 + \Delta x/2) = 2.7085 \times 10^{-1}$ ,  $Rf_6(x_{N/2} + \Delta x/2) = 5.6255 \times 10^{-1}$ , e os valores verdadeiros são  $4.9833 \times 10^{-1}$  e  $8.7570 \times 10^{-1}$  respectivamente. Não conseguimos obter os valores de  $f_6(x_N)$ ,  $f_6(x_{N/2}^-)$  e  $f_6(x_{N/2}^+)$ .

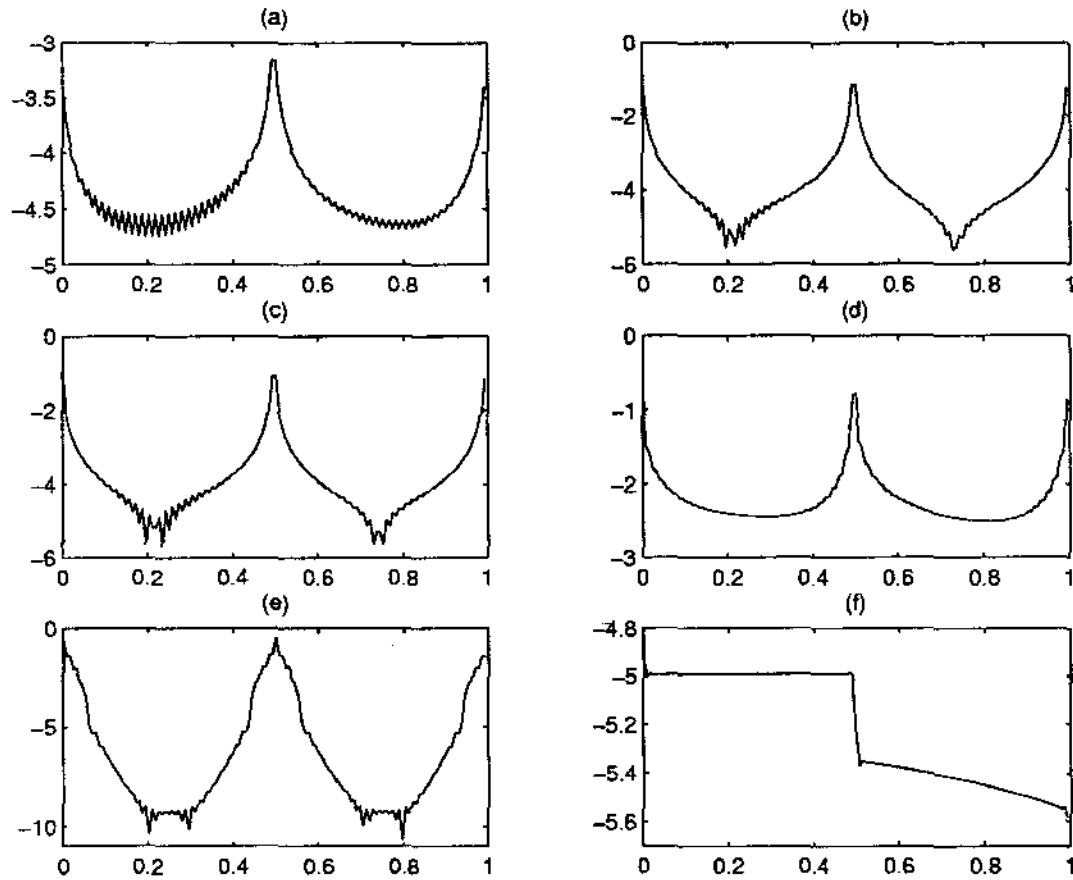


Figura 5.6: Erro  $\log_{10}|Rf_6(x) - f_6(x)|$  em cada ponto com  $N = 256$  e diferentes filtros:  $\{\sigma_k\}$  em (a) e  $\{\sigma_m(\frac{2k}{N}), m = 1, 2, 3, 6\}$  em (b)-(e). (f) mostra o erro pelo Algoritmo 5.3.

A Tabela 5.5 mostra o erro médio quadrático para os métodos de reconstrução usando os diferentes filtros com  $N = 64, 128$  e  $256$ .

Tabela 5.5 Erro médio quadrático para  $f_6(x)$

$N$	$\sigma(k)$	$\sigma_1$	$\sigma_2$	$\sigma_3$	$\sigma_6$	Algoritmo 5.3
64	$3.53 \times 10^{-4}$	$1.67 \times 10^{-2}$	$1.98 \times 10^{-2}$	$4.05 \times 10^{-2}$	$1.46 \times 10^{-2}$	$3.08 \times 10^{-5}$
128	$1.25 \times 10^{-4}$	$1.18 \times 10^{-2}$	$1.40 \times 10^{-2}$	$2.89 \times 10^{-2}$	$1.03 \times 10^{-2}$	$7.68 \times 10^{-6}$
256	$4.42 \times 10^{-5}$	$8.38 \times 10^{-3}$	$9.29 \times 10^{-3}$	$2.05 \times 10^{-2}$	$7.34 \times 10^{-3}$	$1.91 \times 10^{-6}$

Exemplo 5.5. Função  $f_7(x)$ :

$$f_7(x) = \begin{cases} 0, & x \in [0, 0.5 + 1/256], \\ 1, & x \notin [0, 0.5 + 1/256]. \end{cases} \quad (5.20)$$

Dividimos o intervalo  $[0,1]$  em 128 subintervalos típicos ( $N = 128$ ) como (3.14)-(3.15) (onde  $p = 64$ ), e apenas  $(N - 2)$  subintervalos são equidistantes; ou seja, este exemplo satisfaz as condições do Teorema 3.2.

A Fig. 5.7 mostra os gráficos da função reconstruída. Na Fig. 5.8, (a)-(d) são os gráficos do erro em cada ponto pelo Algoritmo 5.1 com quatro filtros  $\{\sigma(k)\}$  e  $\{\sigma_m, m = 1, 2, 3\}$  respectivamente e (e) mostra o erro pelo Algoritmo 5.2 com filtro  $\{\sigma(k)\}$ . Da Fig. 5.8, observamos que os erros da função reconstruída são quase iguais pelo Algoritmo 5.1 com quatro filtros  $\{\sigma(k)\}$  e  $\{\sigma_m, m = 1, 2, 3\}$ , e seus erros médios quadráticos são  $4.9869 \times 10^{-2}$ ,  $4.7283 \times 10^{-2}$ ,  $4.7745 \times 10^{-2}$  e  $5.8743 \times 10^{-2}$  respectivamente. Mas, o erro da função reconstruída é quase  $O(10^{-16})$  causado pelo erro da máquina usando o Algoritmo 5.2 com os quatro filtros, seu erro médio quadrático é  $1.1654 \times 10^{-15}$ .

Exemplo 5.6. Função  $f_8(x)$ :

$$f_8(x) = \begin{cases} x^2, & x \in [0, 0.5 + 1/256], \\ \cos(x), & x \notin [0, 0.5 + 1/256]. \end{cases} \quad (5.21)$$

Como no Exemplo 5.5, dividimos o intervalo  $[0,1]$  em 128 subintervalos típicos ( $N = 128$ ).

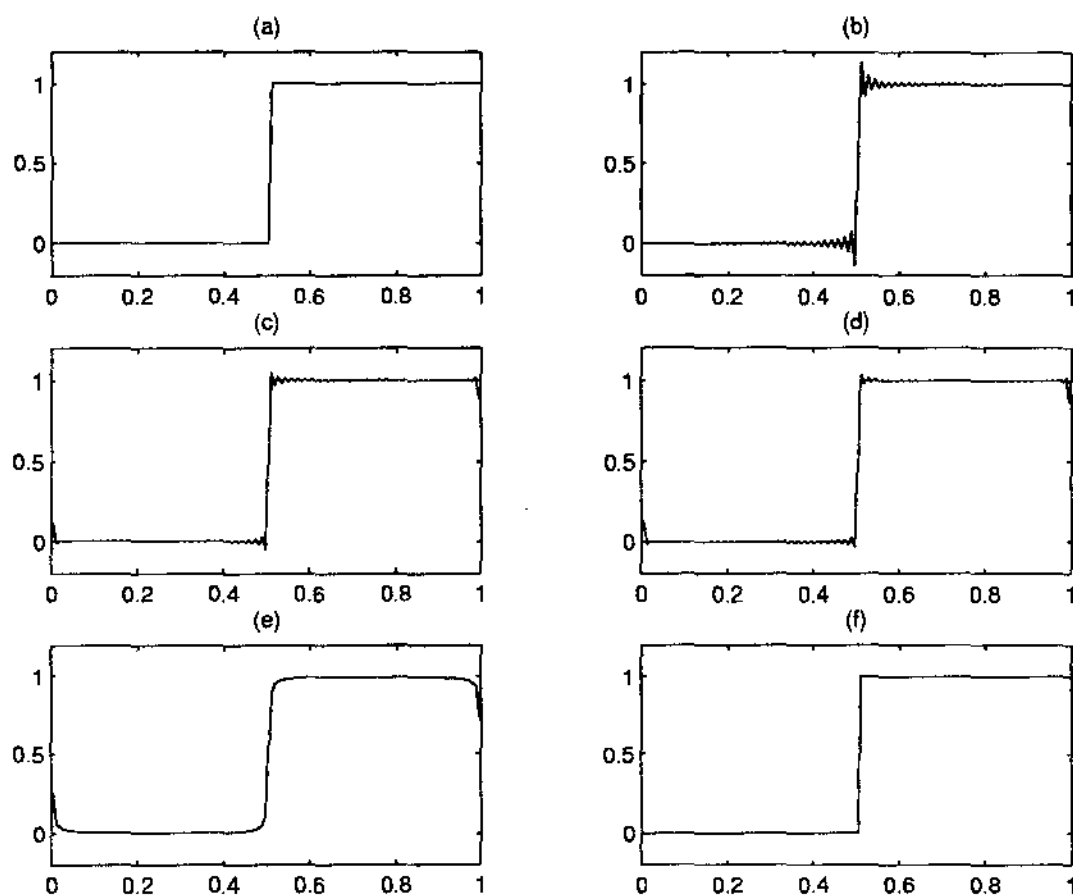


Figura 5.7: Gráfico original de  $f_7(x)$  em (a) e das reconstruções pelo Algoritmo 5.1 com diferentes filtros:  $\{\sigma(k)\}$  e  $\{\sigma_m, m = 1, 2, 3\}$  em (b)-(e) respectivamente. (f) mostra a reconstrução pelo Algoritmo 5.2 com  $\{\sigma(k)\}$ .



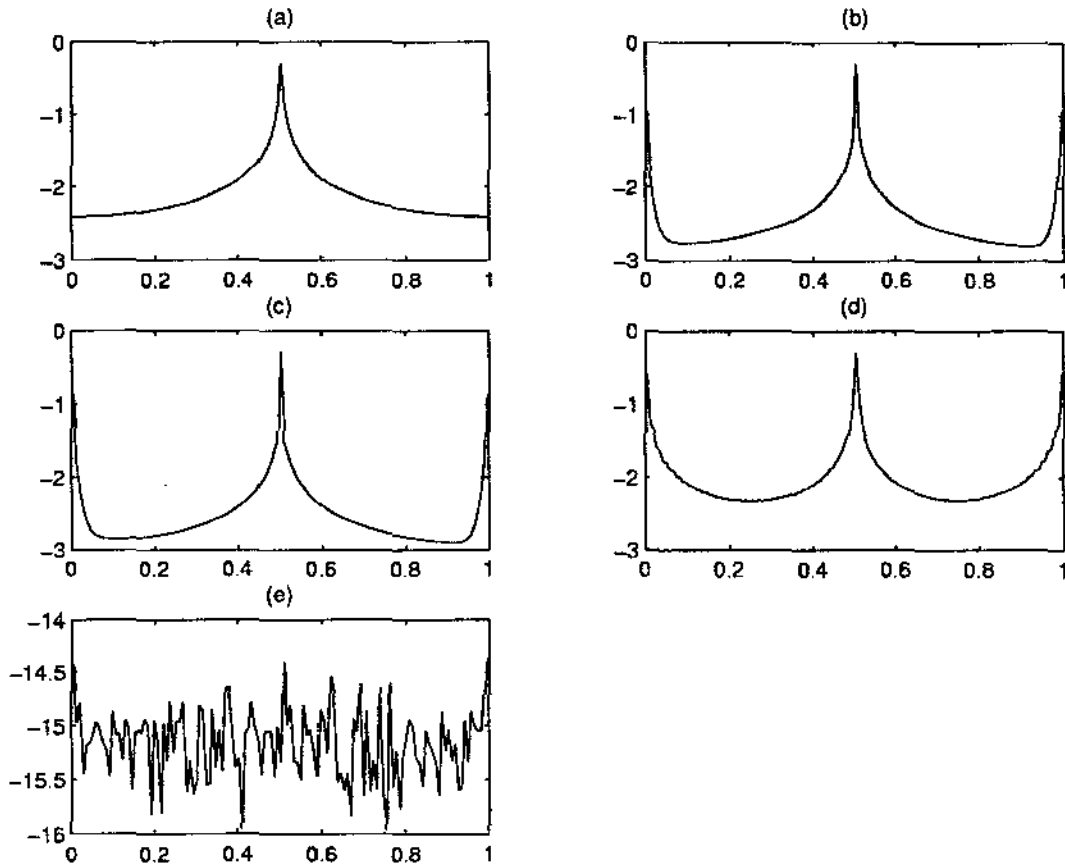


Figura 5.8: Erro  $\log_{10}|Rf_7(x) - f_7(x)|$  em cada ponto com  $N = 128$  pelo Algoritmo 5.1 com diferentes filtros:  $\{\sigma(k)\}$  e  $\{\sigma_m, m = 1, 2, 3\}$  em (a)-(d) respectivamente. (e) mostra o erro pelo Algoritmo 5.2 com  $\{\sigma(k)\}$ .

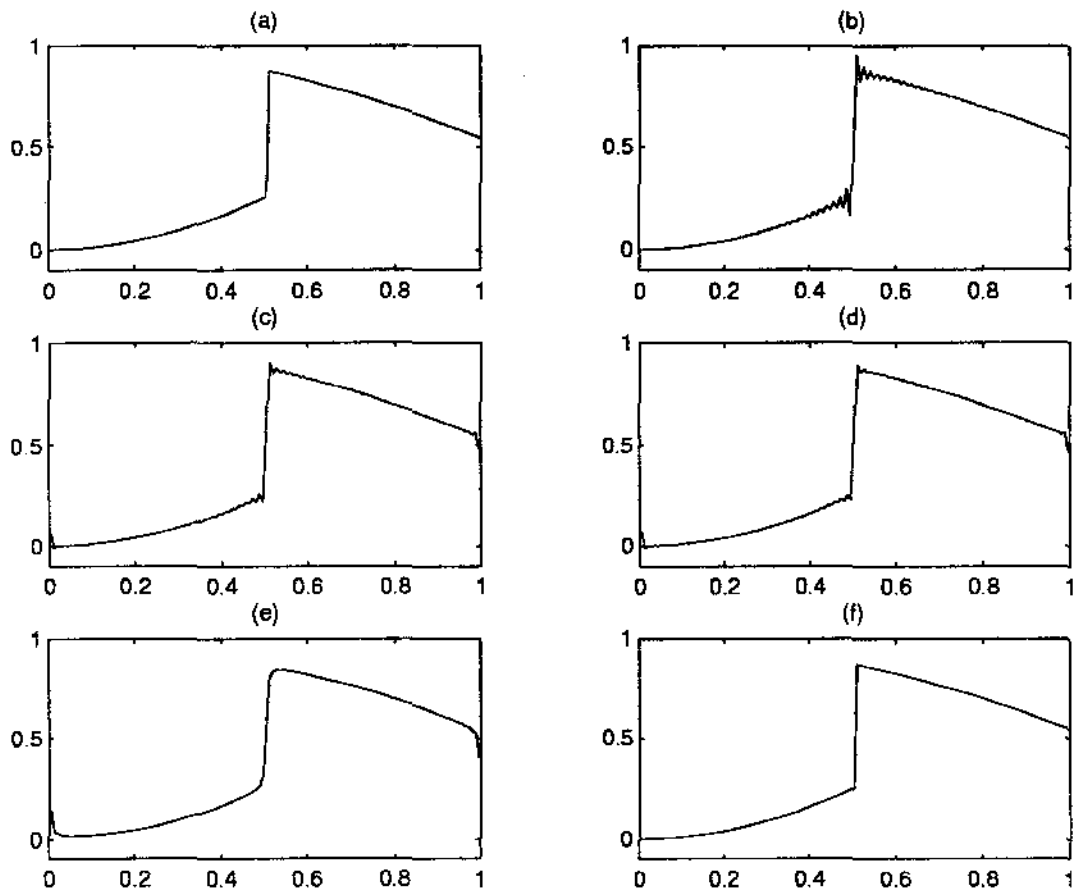


Figura 5.9: Gráfico original de  $f_8(x)$  em (a) e das reconstruções pelo Algoritmo 5.1 com diferentes filtros:  $\{\sigma(k)\}$  e  $\{\sigma_m, m = 1, 2, 3\}$  em (b)-(e) respectivamente. (f) mostra a reconstrução pelo Algoritmo 5.2 com  $\{\sigma(k)\}$ .

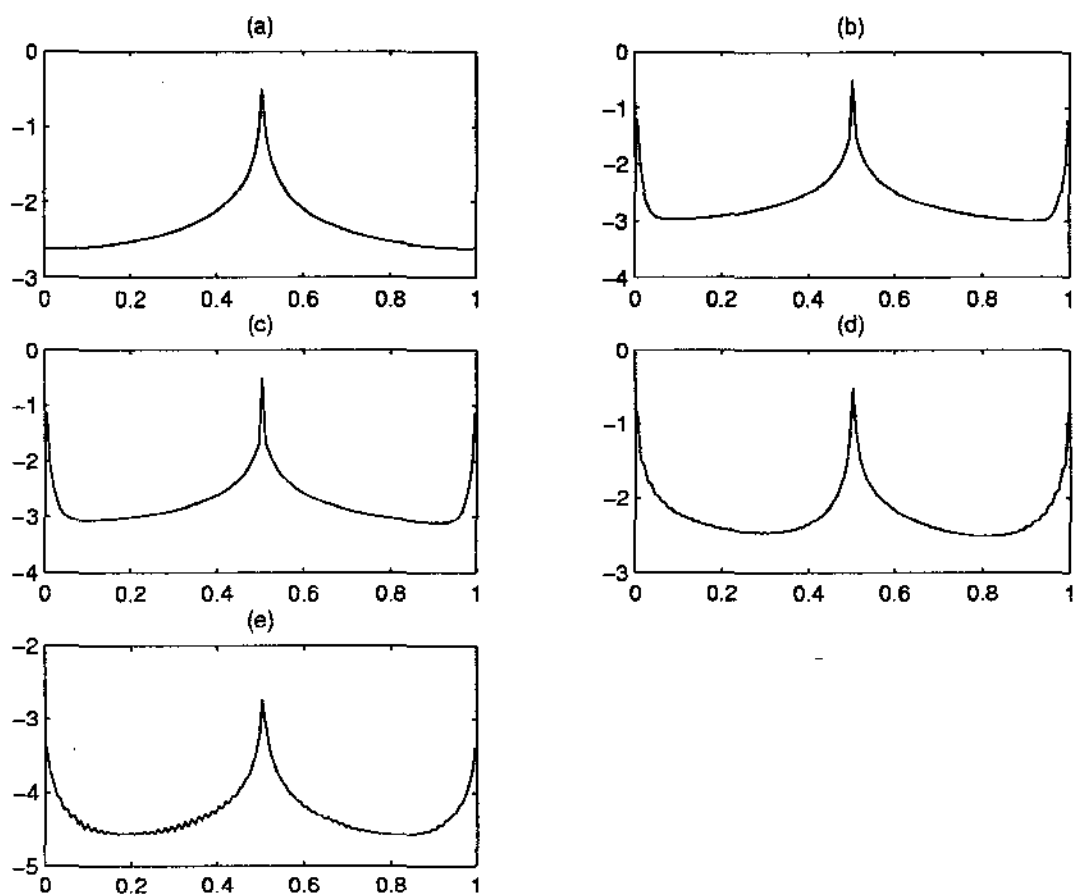


Figura 5.10: Erro  $\log_{10}|Rf_8(x) - f_8(x)|$  em cada ponto com  $N = 128$  pelo Algoritmo 5.1 com diferentes filtros:  $\{\sigma(k)\}$  e  $\{\sigma_m, m = 1, 2, 3\}$  em (a)-(d). (e) mostra o erro pelo Algoritmo 5.2 com  $\{\sigma(k)\}$ .

A Fig. 5.9 mostra o gráfico original da função e a reconstrução. Na Fig. 5.10, (a)-(d) são os gráficos do erro em cada ponto usando Algoritmo 5.1 com quatro filtros  $\{\sigma(k)\}$  e  $\{\sigma_m, m = 1, 2, 3\}$  respectivamente e (e) mostra o erro pelo Algoritmo 5.2 com o filtro  $\{\sigma(k)\}$ . Da Fig. 5.10 observamos que os erros da função reconstruída são quase iguais para o Algoritmo 5.1 com os quatro filtros  $\{\sigma(k)\}$  e  $\{\sigma_m, m = 1, 2, 3\}$ , e seus erros médios quadráticos são  $3.0958 \times 10^{-2}$ ,  $2.8930 \times 10^{-2}$ ,  $2.9071 \times 10^{-2}$  e  $3.4523 \times 10^{-2}$  respectivamente. Mas, usando o Algoritmo 5.2 com nosso filtro, o erro da função reconstruída no ponto próximo de  $x = 0.5 + 1/256$  é bem menor que aquele erro usando o Algoritmo 5.1 com quatro filtros, sendo seu erro quadrático é  $2.0141 \times 10^{-4}$ . Isto mostra que podemos reduzir fortemente as oscilações do fenômeno de Gibbs.

Os exemplos numéricos acima mostram que o filtro  $\sigma(k)$  introduzido no Capítulo 3 é o melhor para reconstruir uma função com o suporte compacto que possui pontos de descontinuidade e/ou é não periódica no intervalo do suporte. É óbvio que temos que utilizar os Algoritmos 5.1 e 5.2 respectivamente, de acordo com os pontos de descontinuidade típicos de  $f(x)$ . Se utilizamos o Algoritmo 5.3, obtemos um resultado melhor.

Em seguida, mostramos dois exemplos em duas dimensões. Assumimos que o suporte da função é  $[-0.5, 0.5] \times [-0.5, 0.5]$ , isto é,  $f(x, y) = 0$ , se  $(x, y) \notin [-0.5, 0.5] \times [-0.5, 0.5]$ . Podemos estender os Algoritmos 5.1 e 5.2 a duas-dimensões. Fizemos o cálculo pelo Algoritmo 5.1 com  $N = 64$  para o exemplo 5.7, e pelos Algoritmos 5.1 e 5.2 para o exemplo 5.8. Além disso, comparamos o erro da função reconstruída em cada ponto pelos métodos com os diferentes filtros  $\sigma(k_1, k_2), \sigma_m(k_1, k_2), m = 1, 2, 3$ . Tomamos  $\sigma(k_1, k_2)$  como  $\sigma(k_1)\sigma(k_2)$ , e  $\sigma_m(k_1, k_2) = \sigma_m(k_1)\sigma_m(k_2), m = 1, 2, 3$ .

Exemplo 5.7 Função  $f_9(x, y)$ :

$$f_9(x, y) = \begin{cases} -x + 1.5, & 0 \leq x \leq 0.25, -x \leq y \leq x, \\ -y + 1.5, & 0 \leq x \leq 0.25, -y \leq x \leq y, \\ x + 1.5, & -0.25 \leq x \leq 0, x \leq y \leq -x, \\ y + 1.5, & -0.25 \leq y \leq 0, y \leq x \leq -y. \end{cases} \quad (5.22)$$

Como em dimensão um, dividimos  $[-0.5, 0.5] \times [-0.5, 0.5]$  em  $N^2$  malhas. Notamos que há alguns pontos de descontinuidade da função, mas  $f(x, y)$  é contínua em cada malha. A Fig. 5.11 mostra o gráfico original de  $f_9(x, y)$ . A Fig. 5.12 mostra os gráficos do erro em cada ponto usando os filtros  $\sigma(k_1, k_2)$  e  $\{\sigma_m(k_1, k_2), m = 1, 2, 3\}$  respectivamente com  $N = 64$ . Da Fig. 5.12, observamos que nosso erro na função reconstruída pelo filtro  $\sigma(k_1, k_2)$  nos pontos próximos da descontinuidade é bem melhor que os erros dos outros filtros. Também, calculamos o erro quadrático. Se  $N=64$ , eles são  $3.04 \times 10^{-4}$ ,  $2.25 \times 10^{-2}$ ,  $2.23 \times 10^{-2}$  e  $2.42 \times 10^{-2}$  respectivamente. Se  $N = 128$ , eles são  $1.05 \times 10^{-4}$ ,  $1.59 \times 10^{-2}$ ,  $1.58 \times 10^{-2}$  e  $1.70 \times 10^{-2}$ . Então, podemos reduzir eficientemente as oscilações do fenômeno de Gibbs.

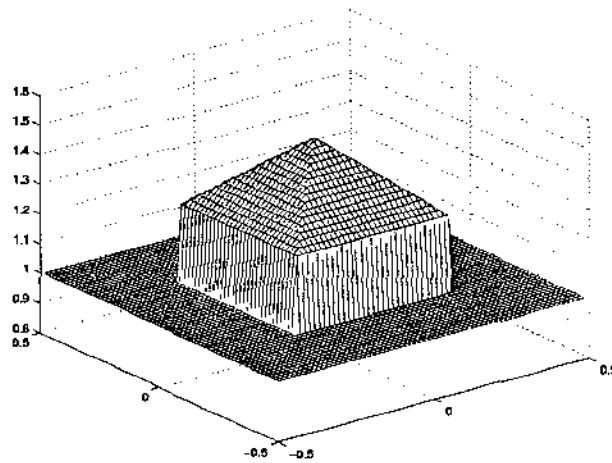


Figura 5.11: Gráfico original  $f_9(x, y)$

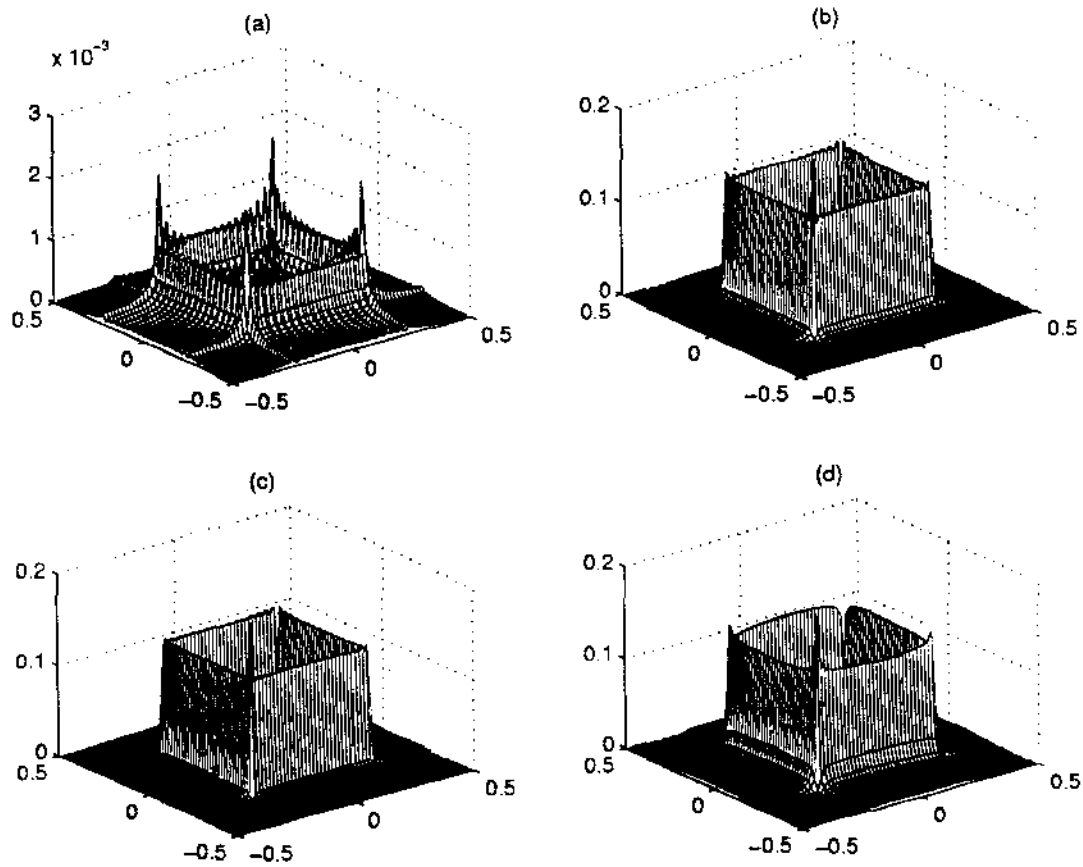


Figura 5.12: Erro  $|Rf_9(x, y) - f_9(x, y)|$  em cada ponto com  $N = 64$  pelo Algoritmo 5.1 com diferentes filtros:  $\{\sigma(k_1, k_2)\}$  e  $\{\sigma_m(k_1, k_2), m = 1, 2, 3\}$  em (a)-(d) respectivamente.

Exemplo 5.8 Função  $f_{10}(x, y)$ :

$$f_{10}(x, y) = \begin{cases} 0, & -0.25 \leq x + y \leq 0.25, -0.25 \leq x - y \leq 0.25, \\ 1, & \text{caso contrário} \end{cases} \quad (5.23)$$

Notamos que esta função é constante por partes. Dividiremos  $[-0.5, 0.5] \times [-0.5, 0.5]$  em  $N^2$  malhas, tal que seja constante em cada malha. Então, chegaremos a uma reconstrução exata.

A Fig. 5.13 mostra o gráfico de  $f_{10}(x, y)$ . A Fig. 5.14 mostra os gráficos do erro em cada ponto usando os filtros  $\sigma(k_1, k_2)$  e  $\{\sigma_m(k_1, k_2), m = 1, 2, 3\}$  pelo Algoritmo 5.1 respectivamente com  $N = 64$ . A Fig. 5.15 mostra o erro da função reconstruída em cada ponto pelo Algoritmo 5.2 com  $\sigma(k_1, k_2)$ . Da Fig. 5.14, observamos que os erros da função reconstruída são quase os mesmos pelo Algoritmo 5.1 com os outros quatro filtros  $\sigma(k_1, k_2)$  e  $\{\sigma_m, m = 1, 2, 3\}$ . Se  $N = 64$ , os erros médios quadráticos são  $7.46 \times 10^{-2}$ ,  $6.55 \times 10^{-2}$ ,  $6.53 \times 10^{-2}$  e  $8.04 \times 10^{-2}$  respectivamente. Se  $N = 128$ , eles são  $5.40 \times 10^{-2}$ ,  $4.54 \times 10^{-2}$ ,  $4.52 \times 10^{-2}$  e  $5.63 \times 10^{-2}$  respectivamente. A Fig. 5.15 mostra que o erro da função reconstruída é cerca de  $O(10^{-16})$  causado pelo erro da máquina usando o Algoritmo 5.2 com filtro  $\sigma(k_1, k_2)$ , e seu erro médio quadrático é  $9.96 \times 10^{-16}$ .

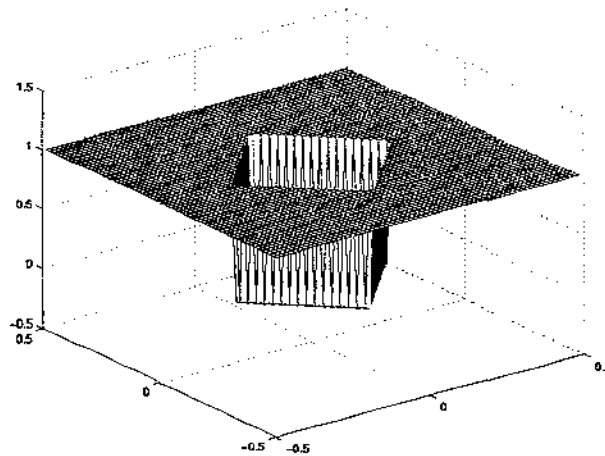


Figura 5.13: Gráfico original  $f_{10}(x, y)$

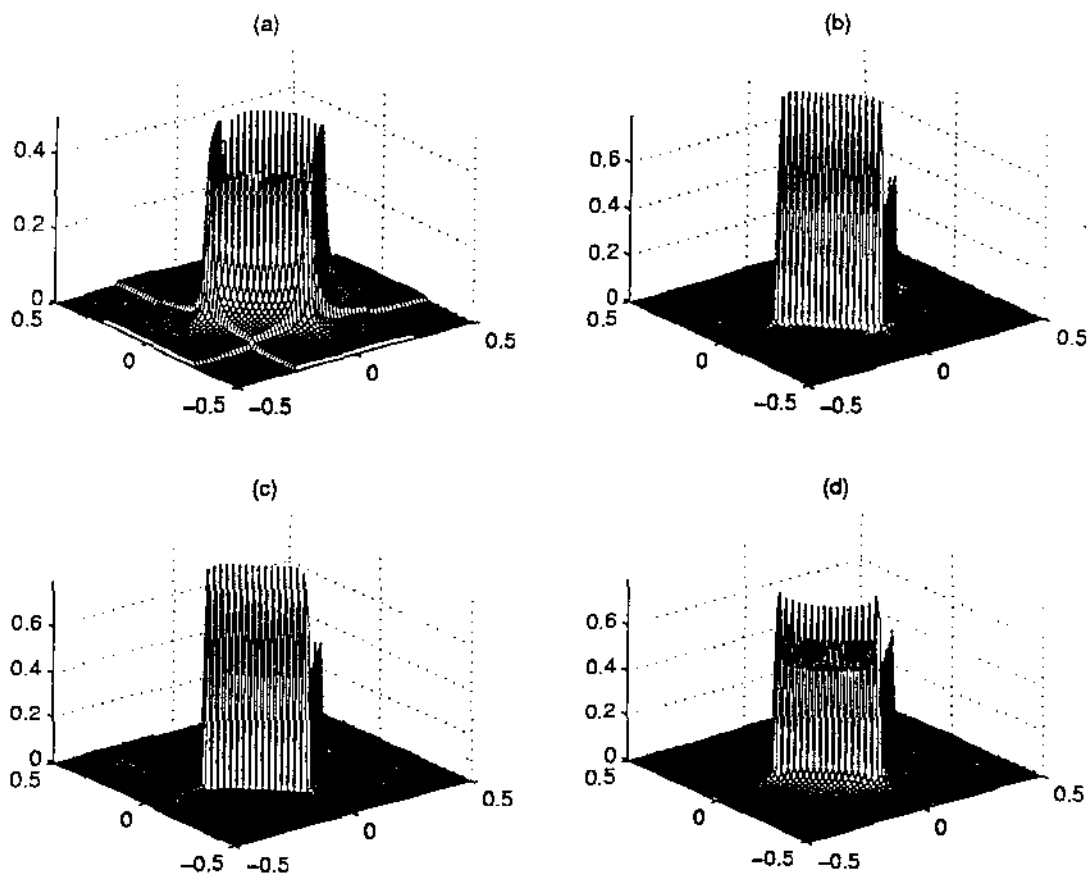


Figura 5.14: Erro  $|Rf_{10}(x,y) - f_{10}(x,y)|$  em cada ponto com  $N = 64$  pelo Algoritmo 5.1 com diferentes filtros:  $\{\sigma(k_1, k_2)\}$  e  $\{\sigma_m(k_1, k_2), m = 1, 2, 3\}$  em (a)-(d) respectivamente.



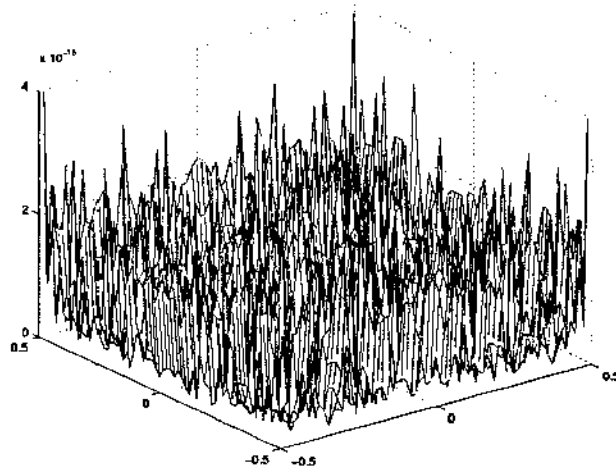


Figura 5.15: Erro  $|Rf_{10}(x, y) - f_{10}(x, y)|$  em cada ponto com  $N = 64$  pelo Algoritmo 5.2 com filtro  $\{\sigma(k_1, k_2)\}$ .

Neste capítulo, apresentamos alguns algoritmos eficientes para a reconstrução de funções a partir de uma amostra discreta da transformada de Fourier. Alguns exemplos numéricos ilustram como, com a nossa metodologia, podemos adotar a transformada rápida de Fourier e também alcançar melhores resultados que os encontrados na literatura.

## Capítulo 6

# Análise da Reconstrução de um Sinal com Ruído Baseado nos Momentos Ortogonais

Suponha que temos um conjunto de dados discretos de um sinal com ruído e que reconstruímos o sinal original a partir dos seus momentos ortogonais, baseados nos polinômios de Legendre. Neste capítulo, deduziremos uma nova estimativa do erro do sinal reconstruído, que revela as conexões entre o número de momentos, a suavidade do sinal e a taxa de discretização. Apresentamos simulações numéricas consistentes com nossa análise teórica.

A Seção 6.1 é uma revisão das propriedades básicas dos polinômios de Legendre e a Seção 6.2 apresenta os algoritmos de reconstrução. Na Seção 6.3 deduzimos a estimativa do erro do sinal reconstruído e na Seção 6.4 apresentamos exemplos numéricos para ilustrar o resultado.

### 6.1 Resultados Preliminares

Nesta seção, introduzimos os polinômios de Legendre e mostramos suas propriedades principais.

**Definição 6.1** Em [11] o polinômio de Legendre de grau  $n$ ,  $P_n(x)$ , é definido pela solução da seguinte equação diferencial:

$$\frac{d}{dx}[(1-x^2)\frac{dP_n(x)}{dx}] + n(n+1)P_n(x) = 0, \quad -1 \leq x \leq 1 \quad (6.1)$$

que satisfaz  $P_n(1) = 1$ .

Então,  $P_0(x) = 1, P_1(x) = x, P_2(x) = \frac{1}{2}(3x^2 - 1)$ , e assim sucessivamente. Os polinômios de Legendre satisfazem a relação da ortogonalidade:

$$\int_{-1}^1 P_i(x)P_j(x)dx = \frac{2}{2i+1}\delta_{ij} \quad \text{com} \quad \delta_{ij} = \begin{cases} 1, & i = j, \\ 0, & i \neq j. \end{cases} \quad (6.2)$$

As seguintes propriedades são válidas (ver [11]):

$$\begin{aligned} (n+1)P_{n+1}(x) &= (2n+1)xP_n(x) - nP_{n-1}(x), \\ P_n(\pm 1) &= (\pm 1)^n, P'_n(\pm 1) = (\pm 1)^{n-1}\frac{1}{2}n(n+1), \\ |P_n(x)| &\leq 1, |P'_n(x)| \leq \frac{1}{2}n(n+1), \forall x \in [-1, 1] \end{aligned} \quad (6.3)$$

e

$$P_n(\cos \theta) = \left(\frac{2}{n\pi \sin \theta}\right)^{\frac{1}{2}} \text{sen}\left[\left(n + \frac{1}{2}\right)\theta + \frac{\pi}{4}\right] + O(n^{-\frac{3}{2}}), \quad n \rightarrow \infty, 0 < \theta < \pi. \quad (6.4)$$

Definindo

$$p_j(x) = \sqrt{\frac{2j+1}{2}}P_j(x), \quad j = 0, 1, \dots \quad (6.5)$$

temos que  $\{p_j(x)\}_0^\infty$  é o sistema ortonormal de Legendre.

Neste capítulo, supomos que o sinal  $f(x)$  é definido no intervalo  $[-1, 1]$  e  $f(x) \in H^p[-1, 1]$ . Se  $f(x)$  é definido em  $[a, b] \neq [-1, 1]$ , podemos aplicar uma mudança de variáveis  $\xi = \frac{2x-a-b}{b-a}$  para mudar o domínio para  $[-1, 1]$ . Aqui, o espaço  $H^p[-1, 1]$  é definido como  $H^p[-1, 1] = \{f | f, f', \dots, f^{(p-1)} \in C[-1, 1], f^{(p)} \in L_2[-1, 1]\}$ . O produto interno de  $f(x)$  com  $g(x)$  é definido por

$$\langle f(x), g(x) \rangle = \int_{-1}^1 f(x)g(x)dx \quad (6.6)$$

e a norma de  $f(x)$  é dada por

$$\|f\|^2 = \int_{-1}^1 f(x)^2 dx. \quad (6.7)$$

É bem conhecido que  $\{p_j(x)\}_0^\infty$  constitui uma base ortonormal do espaço  $L_2[-1, 1]$ . Portanto, toda função  $f(x) \in H^p[-1, 1]$  com  $p \geq 0$ ,  $f(x)$  pode se expressar como combinação linear infinita de  $\{p_j(x)\}_0^\infty$ ,

$$f(x) = \sum_{j=0}^{\infty} a_j p_j(x), \quad (6.8)$$

(converge na norma 2) onde o momento  $a_j$  é dado por

$$a_j = \int_{-1}^1 f(x) p_j(x) dx. \quad (6.9)$$

O algoritmo de reconstrução empregando um número finito de  $a_j$ 's é

$$f_N(x) = \sum_{j=0}^N a_j p_j(x), \quad (6.10)$$

que é a melhor aproximação na norma  $\|f\|$  em  $\text{span}\{p_0(x), p_1(x), \dots, p_N(x)\}$ , isto é,

$$\|g - f_N\| = \min_{g \in \text{span}\{p_0(x), p_1(x), \dots, p_N(x)\}} \|g - f\|.$$

De (6.8), para  $f(\cdot) \in L_2[-1, 1]$ , o erro quadrático do sinal reconstruído  $\int_{-1}^1 (f_N(x) - f(x))^2 dx$  tende a zero quando  $N \rightarrow \infty$ . É claro que podemos fazer o erro arbitrariamente pequeno, usando os momentos de ordem maior. Mas, a situação será diferente quando o sinal é contaminado por ruído e/ou apenas poucas observações finitas estão disponíveis. Neste caso, mostramos que o erro diminui inicialmente (não necessariamente de modo monótono) e aumenta quando  $N \rightarrow \infty$ . Mas, mostraremos que o erro é sempre limitado para qualquer  $N$ , e a melhor escolha do número de momentos depende da suavidade do sinal, da taxa de

amostragem (discretização) e do tamanho do ruído. Isto difere essencialmente dos resultados relatados em [24].

## 6.2 Algoritmos de Reconstrução

Seja  $h(x)$  uma medida de um sinal  $f(x)$  degradado com ruído aleatório, ou seja, assumimos que:

$$h(x) = f(x) + n(x) \quad (6.11)$$

onde  $n(x)$  é um ruído satisfazendo as seguintes condições:

$$E(n(x)) = 0 \text{ e } \text{cov}(n(x), n(x')) = \begin{cases} \sigma^2 w(x), & x = x', \\ 0, & x \neq x'. \end{cases} \quad (6.12)$$

( $E$  denota a esperança e  $\text{cov}$  a covariância) com

$$0 \leq w(x) \leq W, \quad \forall x \in [-1, 1]. \quad (6.13)$$

Precisamos recuperar  $f(x)$  a partir de  $(m + 1)$  dados medidos  $\{h(x_i)\}$ ,  $-1 = x_0 < x_1 < \dots < x_m = 1$ . Denotamos

$$f_i = f(x_i), \quad h_i = h(x_i), \quad w_i = w(x_i), \quad \text{e } n_i = n(x_i), \quad i = 0, 1, \dots, m, \quad (6.14)$$

$$\Delta = \Delta_i = x_{i+1} - x_i, \quad i = 0, 1, \dots, m - 1. \quad (6.15)$$

Sem perda de generalidade, assumimos aqui que o intervalo  $[-1, 1]$  está dividido em  $m$  subintervalos iguais. A análise para subintervalos diferentes é semelhante. Sejam

$$\bar{f}^{(0)}(x) = f_i, \quad x \in [x_i, x_{i+1}), \quad (6.16)$$

$$\bar{f}^{(1)}(x) = \frac{x - x_{i+1}}{x_i - x_{i+1}} f_i + \frac{x - x_i}{x_{i+1} - x_i} f_{i+1}, \quad x \in [x_i, x_{i+1}], \quad (6.17)$$

para  $i = 0, 1, \dots, m - 1$ ; e

$$\bar{f}^{(2)}(x) = \frac{(x-x_{2i+1})(x-x_{2i+2})}{(x_{2i}-x_{2i+1})(x_{2i}-x_{2i+2})} f_{2i} + \frac{(x-x_{2i})(x-x_{2i+2})}{(x_{2i+1}-x_{2i})(x_{2i+1}-x_{2i+2})} f_{2i+1} + \frac{(x-x_{2i})(x-x_{2i+1})}{(x_{2i+2}-x_{2i})(x_{2i+2}-x_{2i+1})} f_{2i+2}, \quad x \in [x_{2i}, x_{2i+2}], \quad (6.18)$$

para  $i = 0, 1, \dots, m' - 1$ . Em (6.18), por simplicidade, assumimos que  $m = 2m'$ .  $\bar{f}^{(k)}(x)$ , para  $k = 0, 1, 2$ , são as aproximações constante, linear e quadrática por partes de  $f(x)$  respectivamente, interpolando nos pontos dados em [31]. As funções  $\bar{h}^{(k)}(x)$  e  $\bar{n}^{(k)}(x)$  são definidas analogamente. Em seguida, definimos

$$\begin{aligned} \bar{f}_N^{(k)}(x) &= \sum_{j=0}^N \bar{a}_j^{(k)} p_j(x) \text{ com } \bar{a}_j^{(k)} = \int_{-1}^1 \bar{f}^{(k)}(x) p_j(x) dx, \\ \bar{h}_N^{(k)}(x) &= \sum_{j=0}^N \bar{b}_j^{(k)} p_j(x) \text{ com } \bar{b}_j^{(k)} = \int_{-1}^1 \bar{h}^{(k)}(x) p_j(x) dx, \\ \bar{n}_N^{(k)}(x) &= \sum_{j=0}^N \bar{c}_j^{(k)} p_j(x) \text{ com } \bar{c}_j^{(k)} = \int_{-1}^1 \bar{n}^{(k)}(x) p_j(x) dx, \end{aligned} \quad (6.19)$$

onde  $j = 0, 1, \dots, N$  e  $k = 0, 1, 2$ .

Na próxima seção mediremos a diferença entre a aproximação  $\bar{h}_N^{(k)}(x)$  e  $f(x)$  na norma de  $L_2$ ,

$$T(N, k, \Delta) = \|\bar{h}_N^{(k)}(x) - f(x)\|. \quad (6.20)$$

### 6.3 Erro do Sinal Reconstruído

Para deduzir a estimativa do erro, precisamos os seguintes lemas, que são válidos para  $k = 0, 1, 2$ . Primeiramente, aplicando diretamente as definições de  $\bar{f}^{(k)}(x)$ ,  $\bar{h}^{(k)}(x)$  e  $\bar{n}^{(k)}(x)$ , temos

#### Lema 6.1

$$\bar{h}_N^{(k)}(x) = \bar{f}_N^{(k)}(x) + \bar{n}_N^{(k)}(x). \quad (6.21)$$

#### Lema 6.2

$$E(\|\bar{h}_N^{(k)} - f\|^2) = \|\bar{f}_N^{(k)} - f\|^2 + E(\|\bar{n}_N^{(k)}\|^2). \quad (6.22)$$

**Demonstração:** Observamos que para cada  $k = 0, 1, 2, \dots$ ,  $\{\bar{n}_N^{(k)}(x)\}$  e  $\{\bar{f}_N^{(k)}(x)\}$  são as combinações lineares (para fixo  $x$ ) de  $\{n_0, n_1, \dots, n_m\}$ , que são variáveis aleatórias com as propriedades (6.12), então, usando (6.12) e o Lema 6.1, temos

$$E(\bar{n}_N^{(k)}(x)) = 0 \text{ e } E\bar{h}_N^{(k)}(x) = \bar{f}_N^{(k)}(x).$$

Aplicando a esperança a (6.20) temos

$$\begin{aligned} E(\|\bar{h}_N^{(k)} - f\|^2) &= E(\|\bar{n}_N^{(k)} + \bar{f}_N^{(k)} - f\|^2) \\ &= E[\|\bar{n}_N^{(k)}\|^2 + \|\bar{f}_N^{(k)} - f\|^2 + 2\langle \bar{n}_N^{(k)}, \bar{f}_N^{(k)} - f \rangle] \\ &= E(\|\bar{n}_N^{(k)}\|^2) + \|\bar{f}_N^{(k)} - f\|^2; \end{aligned}$$

Na última igualdade, usamos o fato de que

$$E(\langle \bar{n}_N^{(k)}, \bar{f}_N^{(k)} - f \rangle) = 0.$$

para chegar no resultado desejado.

### Lema 6.3

$$\|\bar{f}_N^{(k)} - f\| \leq \|f_N - f\| + 2\|\bar{f}^{(k)} - f\|. \quad (6.23)$$

**Demonstração.** Aplicando a desigualdade triangular, temos que

$$\begin{aligned} \|\bar{f}_N^{(k)} - f\| &= \|\bar{f}_N^{(k)} - \bar{f}^{(k)} + \bar{f}^{(k)} - f\| \leq \|\bar{f}_N^{(k)} - \bar{f}^{(k)}\| + \|\bar{f}^{(k)} - f\| \\ &\leq \|f_N - \bar{f}^{(k)}\| + \|\bar{f}^{(k)} - f\| \leq \|f_N - f\| + 2\|\bar{f}^{(k)} - f\| \end{aligned}$$

onde já utilizamos que  $\|\bar{f}_N^{(k)} - \bar{f}^{(k)}\| \leq \|f_N - \bar{f}^{(k)}\|$ , pois  $\bar{f}_N^{(k)}$  é a melhor aproximação para  $\bar{f}^{(k)}$  em  $\text{span}\{p_0(x), \dots, p_N(x)\}$ .

Usando a ortonormalidade de  $\{p_j(x), j = 0, 1, \dots, N\}$  e a fórmula de Parseval, é fácil obter o lema seguinte.

**Lema 6.4**

$$\|\bar{n}_N^{(k)}\|^2 = \sum_{j=0}^N (\bar{c}_j^{(k)})^2 \leq \|\bar{n}^{(k)}\|^2. \quad (6.24)$$

Os dois lemas seguintes apresentam as estimativas para  $E(\|\bar{n}^{(k)}\|^2)$  e  $E(\|\bar{n}_N^{(k)}\|^2)$ .

**Lema 6.5**

$$E\|\bar{n}_i^{(k)}\|^2 \leq \beta^{(k)}\sigma^2W \text{ com } \beta^{(k)} = \begin{cases} 2, & k = 0, \\ \frac{4}{3}, & k = 1, \\ \frac{8}{5}, & k = 2. \end{cases} \quad (6.25)$$

**Demonstração:** Da definição de  $\bar{n}^{(k)}$ , fazendo alguns cálculos e usando (6.12)-(6.15), temos que

$$E\|\bar{n}^{(0)}\|^2 = E\left\{\sum_{i=0}^{m-1} \int_{x_i}^{x_{i+1}} n_i^2 dx\right\} = \sum_{i=0}^{m-1} \sigma^2 w_i \Delta \leq 2\sigma^2 W,$$

$$\begin{aligned} E\|\bar{n}^{(1)}\|^2 &= E\left\{\sum_{i=0}^{m-1} \int_{x_i}^{x_{i+1}} \left[\frac{x-x_{i+1}}{\Delta}n_i + \frac{x-x_i}{\Delta}n_{i+1}\right]^2 dx\right\} \\ &= \sum_{i=0}^{m-1} \frac{1}{3}\sigma^2(w_i + w_{i+1})\Delta \leq \frac{4}{3}\sigma^2 W, \end{aligned}$$

e

$$\begin{aligned} E\|\bar{n}^{(2)}\|^2 &= \sum_{i=0}^{m-1} \sigma^2 \left[ \frac{w_{2i}}{4\Delta^4} \int_{x_{2i}}^{x_{2i+2}} (x-x_{2i+1})^2 (x-x_{2i+2})^2 dx \right. \\ &\quad \left. + \frac{w_{2i+1}}{\Delta^4} \int_{x_{2i}}^{x_{2i+2}} (x-x_{2i})^2 (x-x_{2i+2})^2 dx + \frac{w_{2i+2}}{4\Delta^4} \int_{x_{2i}}^{x_{2i+2}} (x-x_{2i})^2 (x-x_{2i+1})^2 dx \right] \\ &= \sigma^2 \Delta \sum_{i=0}^{m-1} \left[ \frac{4}{15}w_{2i} + \frac{16}{15}w_{2i+1} + \frac{4}{15}w_{2i+2} \right] \leq \frac{8}{5}\sigma^2 W. \end{aligned}$$

O seguinte lema mostra a taxa do decréscimo para os coeficientes de  $n^{(k)}$ .

**Lema 6.6** Para  $j = 0, 1, \dots$ ,

$$E(\bar{c}_j^{(k)})^2 \leq \alpha^{(k)}\sigma^2W\Delta, \text{ onde } \alpha^{(k)} = \begin{cases} 1, & k = 0, \\ \frac{4}{3}, & k = 1, \\ \frac{32}{15}, & k = 2. \end{cases} \quad (6.26)$$

E mais, para  $j$  suficientemente grande, existe uma constante  $C$  tal que,

$$E(\bar{c}_j^{(k)})^2 \leq \frac{C}{j^2}. \quad (6.27)$$



**Demonstração:** Basta provar o resultado para o caso  $k = 0$ . As provas para  $k = 1, 2$  são semelhantes. Da definição de  $\bar{c}_j^{(k)}$ , temos  $\bar{c}_j^{(0)} = \sum_{i=0}^{m-1} n_i \int_{x_i}^{x_{i+1}} p_j(x) dx$ . Logo,

$$E(\bar{c}_j^{(0)})^2 = \sum_{i=0}^{m-1} \sigma^2 w_i \left( \int_{x_i}^{x_{i+1}} p_j(x) dx \right)^2. \quad (6.28)$$

Aplicando a desigualdade de Cauchy-Schwartz, obtemos

$$\left( \int_{x_i}^{x_{i+1}} p_j(x) dx \right)^2 \leq \left( \int_{x_i}^{x_{i+1}} 1^2 dx \right) \left( \int_{x_i}^{x_{i+1}} p_j(x)^2 dx \right) = \Delta \int_{x_i}^{x_{i+1}} p_j(x)^2 dx,$$

e de (6.28), temos que

$$E(\bar{c}_j^{(0)})^2 \leq \sigma^2 W \Delta \int_{-1}^1 p_j(x)^2 dx = \sigma^2 W \Delta.$$

Para  $j$  suficientemente grande, podemos utilizar a expressão assintótica de  $p_j(x)$  em (6.4)-(6.5) e, aplicando o teorema de Riemann para integrais com a mudança de variável apropriada, obtemos

$$\left| \int_{x_i}^{x_{i+1}} p_j(x) dx \right| = \sqrt{\frac{2j+1}{2}} \left| \int_{\arccos(x_{i+1})}^{\arccos(x_i)} \left( \frac{2 \operatorname{sen} \theta}{j\pi} \right)^{\frac{1}{2}} \operatorname{sen} \left[ \left( j + \frac{1}{2} \right) \theta + \frac{\pi}{4} \right] d\theta \right| + O\left(\frac{1}{j}\right) \leq \frac{C_1}{j},$$

onde  $C_1$  é uma constante positiva. A desigualdade acima é uma consequência do fato de que a integral é limitada por  $\frac{1}{j}$ . Pela estimativa acima e (6.28) obtemos, para  $j$  suficientemente grande,

$$E(\bar{c}_j^{(0)})^2 = \sum_{i=0}^{m-1} \sigma^2 w_i \left( \int_{x_i}^{x_{i+1}} p_j(x) dx \right)^2 \leq \frac{C}{j^2},$$

onde  $C$  é outra constante positiva; completando assim a demonstração.

Agora, estamos prontos para apresentar o principal resultado teórico deste Capítulo.

**Teorema 6.1** Seja  $f(x) \in H^p[-1, 1]$ ,  $0 \leq w(x) \leq W$ ,  $x \in [-1, 1]$  com  $p \geq k + 1$ . Então, para  $N$  pequeno,

$$E(\|\bar{h}_N^{(k)} - f\|^2) \leq \left[ \frac{C(f, p)}{N^p} + D(f, k)\Delta^{k+1} \right]^2 + \sigma^2 \alpha^{(k)} W \Delta(N+1), \quad (6.29)$$

e para  $N \rightarrow +\infty$ ,

$$E(\|\bar{h}_N^{(k)} - f\|^2) \leq \left[ \frac{C(f, p)}{N^p} + D(f, k)\Delta^{k+1} \right]^2 + \sigma^2 \beta^{(k)} W, \quad (6.30)$$

onde  $\beta^{(k)}$  e  $\alpha^{(k)}$  são as constantes respectivamente definidas nos Lemas 6.5 e 6.6,  $C(f, p) > 0$  é uma constante que somente depende de  $f$  e  $p$ , e  $D(f, k) > 0$  uma constante que só depende de  $f$  e  $k$ .

**Demonstração:** Aplicando diretamente os Lemas 6.2 e 6.3 temos

$$E(\|\bar{h}_N^{(k)} - f\|^2) = \|\bar{f}_N^{(k)} - f\|^2 + E(\|\bar{n}_N^{(k)}\|^2) \leq (\|f_N - f\| + 2\|\bar{f}^{(k)} - f\|)^2 + E(\|\bar{n}_N^{(k)}\|^2). \quad (6.31)$$

Agora, basta provar que  $\|f_N - f\| \leq \frac{C(f, p)}{N^p}$ . Sejam

$$\begin{aligned} F_0(x) &= f(x), \quad F_{2n+1}(x) = (1-x^2) \frac{d}{dx} F_{2n}(x), \\ F_{2n+2}(x) &= \frac{d}{dx} F_{2n+1}(x) \text{ para } 0 \leq n \leq 2\left[\frac{p}{2}\right], \end{aligned} \quad (6.32)$$

onde

$$\left[\frac{p}{2}\right] = \frac{p}{2} \text{ para } p \text{ par e } \left[\frac{p}{2}\right] = \frac{p-1}{2} \text{ para } p \text{ ímpar.}$$

De (6.32) deduzimos que  $F_l(\cdot) \in H^{p-l}[-1, 1]$  e  $F_{2n-1}(\pm 1) = 0$ . Usando (6.1), (6.5) e (6.9), para  $j > N$ , e integrando por partes de novo, obtemos

$$\begin{aligned} a_j &= \int_{-1}^1 f(x) p_j(x) dx = -\frac{1}{j(j+1)} \int_{-1}^1 F_0(x) \frac{d}{dx} [(1-x^2) \frac{d}{dx} p_j(x)] dx \\ &= -\frac{1}{j(j+1)} \left( F_0(x) [(1-x^2) \frac{d}{dx} p_j(x)] \Big|_{-1}^1 - \int_{-1}^1 F_1(x) \frac{d}{dx} p_j(x) dx \right) \\ &= -\frac{1}{j(j+1)} \left( F_1(x) p_j(x) \Big|_{-1}^1 + \int_{-1}^1 F_2(x) p_j(x) dx \right) \\ &= \frac{-1}{j(j+1)} \int_{-1}^1 F_2(x) p_j(x) dx. \end{aligned}$$

Continuando este procedimento, temos que, para  $l \leq \left[\frac{p}{2}\right]$ ,

$$a_j = \frac{-1}{(-j(j+1))^l} \int_{-1}^1 F_{2l-1}(x) \frac{d}{dx} p_j(x) dx = \frac{1}{(-j(j+1))^l} \int_{-1}^1 F_{2l}(x) p_j(x) dx. \quad (6.33)$$

Da fórmula acima, para  $j$  suficientemente grande e usando o teorema de Riemann, temos que  $a_j^2 \leq \frac{C_1(f,p)}{j^{2p+1}}$  para alguma constante  $C_1(f,p)$ . Agora, utilizamos a limitação anterior para estimar a limitação do erro aproximando  $f$  por  $f_N$ ,

$$\|f_N - f\| = \sqrt{\sum_{j=N+1}^{+\infty} a_j^2} \leq C_1(f,p)^{\frac{1}{2}} \sqrt{\sum_{j=N+1}^{+\infty} \frac{1}{j^{2p+1}}} \leq \frac{C(f,p)}{N^p},$$

onde já utilizamos o fato de que para  $p > 0$ ,

$$\int_{N+1}^{\infty} \frac{1}{x^{2p+1}} dx < \sum_{j=N+1}^{+\infty} \frac{1}{j^{2p+1}} < \int_N^{\infty} \frac{1}{x^{2p+1}} dx$$

Assim,  $\sqrt{\sum_{j=N+1}^{+\infty} \frac{1}{j^{2p+1}}} = \frac{C}{N^p}$  para alguma constante positiva  $C$ .

Em seguida, provaremos  $\|\bar{f}^{(k)} - f\| \leq D(f,k)\Delta^{k+1}$ . Pelo fato de que, em cada subintervalo,  $\bar{f}^{(k)}(x)$  é a  $k$ -ésima interpolação de Lagrange de  $f(x)$  em [31], temos que

$$\begin{aligned} f(x) - \bar{f}^{(0)}(x) &= (x - x_i)f'(y), \quad x, y \in [x_i, x_{i+1}], \\ f(x) - \bar{f}^{(1)}(x) &= (x - x_i)(x - x_{i+1})f''(y)/2, \quad x, y \in [x_i, x_{i+1}], \\ f(x) - \bar{f}^{(2)}(x) &= (x - x_{2i})(x - x_{2i+1})(x - x_{2i+2})f'''(y)/6, \quad x, y \in [x_{2i}, x_{2i+2}]. \end{aligned}$$

Das fórmulas acima, é fácil obter a desigualdade

$$\|\bar{f}^{(k)} - f\| \leq D(f,k)\Delta^{k+1}.$$

A limitação para o segundo termo em (6.30) segue dos Lemas 6.4, 6.5 e 6.6.

Agora, apresentamos algumas observações que podem ser deduzidas das estimativas do Teorema 6.1:

1. De (6.33) deduzimos que  $C(f,p) = O(\|F_p\|)$ . Se  $f(\cdot) \in C^\infty$ , podemos escolher  $p$  como qualquer número positivo, mas,  $C(f,p)$  aumenta muito rapidamente quando  $p$  aumenta. Então, se  $N$  é relativamente pequeno, não podemos escolher  $p$  tão grande, mas, se  $N$  é suficientemente grande, podemos escolher qualquer  $p > 0$ . Além disso, se  $N$  é suficientemente grande,  $\|f_N - f\| \rightarrow 0$  exponencialmente.

2. Do Teorema 6.1 deduzimos que o erro do sinal reconstruído depende dos parâmetros  $\Delta$ ,  $N$ ,  $\sigma$ ,  $p$  e  $k$ . Do Teorema 6.1, temos que, para  $N$  pequeno,

$$\|\bar{h}_N^{(k)} - f\| = O\left(\frac{C(f, p)}{N^p} + D(f, k)\Delta^{k+1} + \sigma\sqrt{\alpha^{(k)}W\Delta(N+1)}\right), \quad (6.34)$$

e para  $N \rightarrow +\infty$ ,

$$\|\bar{h}_N^{(k)} - f\| = O\left(\frac{C(f, p)}{N^p} + D(f, k)\Delta^{k+1} + \sigma\sqrt{\beta^{(k)}W}\right). \quad (6.35)$$

Para obter a melhor reconstrução, um compromisso é necessário entre os valores de todos os parâmetros.

3. Se o sinal  $f(x)$  é suave por partes e há finitos pontos de descontinuidade, então,  $p = 0$ . Isso significa que, com ou sem ruído aleatório, sempre ocorre o fenômeno de Gibbs. Neste caso, o erro  $\|f_N - f\|$  domina o erro do sinal reconstruído, e assim, os algoritmos na seção 6.2 fracassam na obtenção de um bom resultado. Este efeito é demonstrado nos experimentos numéricos descritos na próxima seção.

## 6.4 Exemplos Numéricos

Nesta seção, apresentamos alguns experimentos numéricos para ilustrar os resultados teóricos da seção anterior. Escolhemos  $m = 256$  e então  $\Delta = \frac{1}{128}$ .

Para o primeiro conjunto de experimentos, a função dada é

$$f(x) = \frac{1}{2-x} + \sin(2\pi x) \quad \text{para } x \in [-1, 1], \quad (6.36)$$

então  $f(\cdot) \in C^\infty[-1, 1]$ .

A Figura 6.1 mostra a relação entre  $N$  e  $\|f_N - f\|$  para a reconstrução sem ruído. Clara-

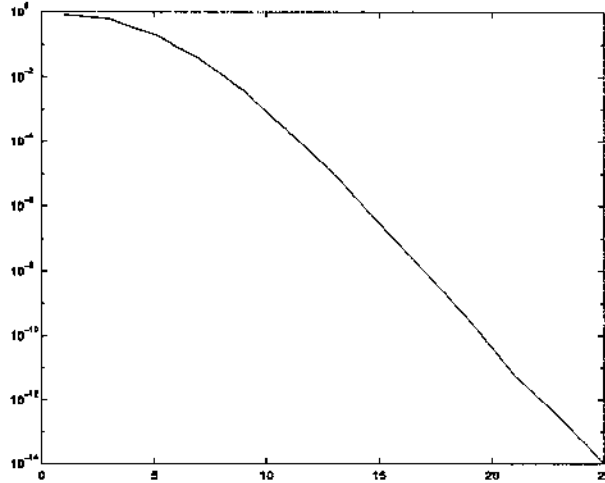


Figura 6.1: Erro  $\|f_N - f\|$  como função do número de momentos  $N$ .

mente, quando  $N$  aumenta,  $\|f_N - f\|$  decresce com uma taxa exponencial. Temos

$$\|f_N - f\| = \begin{cases} \sim 10^{-1}, & \text{para } 0 \leq N \leq 6, \\ \sim 10^{-2}, & \text{para } 7 \leq N \leq 8, \\ \sim 10^{-3}, & \text{para } 9 \leq N \leq 10, \\ \sim 10^{-4}, & \text{para } 11 \leq N \leq 12, \\ \sim 10^{-5}, & \text{para } 13 \leq N \leq 14, \\ \sim 10^{-7}, & \text{para } 15 \leq N \leq 16, \\ \sim 10^{-8}, & \text{para } 17 \leq N \leq 18, \\ \leq 10^{-10}, & \text{para } 19 \leq N, \\ \leq u, & \text{para } 26 \leq N. \end{cases} \quad (6.37)$$

A Figura 6.2 mostra o comportamento do erro  $\|\bar{f}_N^{(k)} - f\|$  em função de  $N$  ( $N = 0, \dots, 25$ ), para diferentes valores  $k$  ( $k = 0, 1, 2$ ). Considerando

$$\Delta^{k+1} = \begin{cases} O(10^{-2}), & \text{para } k = 0, \\ O(10^{-4}) \sim O(10^{-5}), & \text{para } k = 1, \\ O(10^{-6}) \sim O(10^{-7}), & \text{para } k = 2. \end{cases}$$

junto com (6.37) e a Figura 6.1, observamos que se  $N \geq 19$ , então  $\|f_N - f\| \ll \|\bar{f}^{(k)} - f\|$ , e deduzimos

$$\|\bar{f}^{(k)} - f\| = \begin{cases} O(10^{-2}), & \text{para } k = 0, \\ O(10^{-4}), & \text{para } k = 1, \\ O(10^{-7}), & \text{para } k = 2. \end{cases} \quad (6.38)$$

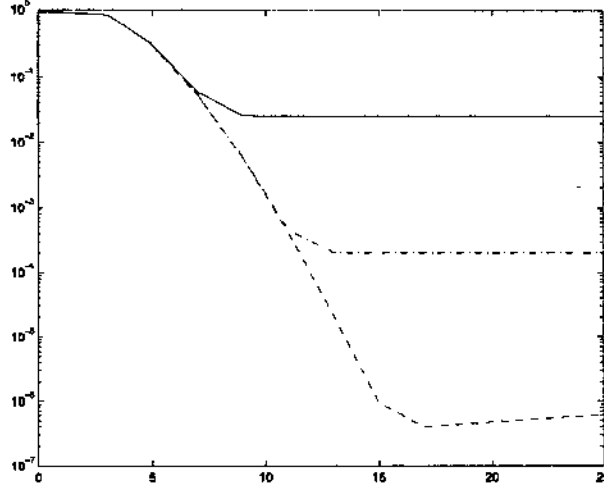


Figura 6.2: Erro  $\|\bar{f}_N^{(k)} - f\|$  em função de  $N$  para diferentes valores de  $k$ :  $k = 0$  (—),  $k = 1$  (-.-) e  $k = 2$  (- -).

Agora, tomando  $\{n_j\}$  como um conjunto das variáveis aleatórias com  $w_j = W = 1$  para  $j = 0, \dots, 256$  com várias  $\sigma$ , que são independentes e normalmente distribuídas, calculamos  $\|\bar{h}_N^{(k)} - f\|$  para alguns valores de  $N$  e  $k = 0, 1, 2$ .

As Figuras 6.3 (a)-(e) mostram a relação entre  $\sigma = 10^{-5}, 10^{-3}, 10^{-2}, 10^{-1}, 0.5$ ,  $N = 0, 1, \dots, 25$ , e  $\|\bar{h}_N^{(k)} - f\|$  para  $k = 0, 1, 2$ , respectivamente.

Se  $\sigma = 10^{-5}$ , o erro mínimo é  $\|\bar{h}_N^{(k)} - f\| = 3.20 \times 10^{-6}$ , em  $N = 15$  e  $k = 2$ . Quando  $15 \leq N \leq 25$ , para  $k = 0, 1, 2$ , o erro quase não muda.

Se  $\sigma = 10^{-3}$ , o erro mínimo é  $\|\bar{h}_N^{(k)} - f\| = 3.52 \times 10^{-4}$  em  $N = 13$ , para  $k = 1, 2$ . Quando  $13 \leq N \leq 20$ , para  $k = 1, 2$ ,  $\|\bar{h}_N^{(k)} - f\|$  também quase não muda.

Se  $\sigma = 10^{-2}$ , o erro mínimo é  $\|\bar{h}_N^{(k)} - f\| = 3.01 \times 10^{-3}$  em  $N = 11$ , para  $k = 1, 2$ . Quando  $11 \leq N \leq 15$ , para  $k = 1, 2$ ,  $\|\bar{h}_N^{(k)} - f\|$  é da mesma ordem.

Se  $\sigma = 10^{-1}$ , o menor erro é alcançado para  $N = 9$ , para  $k = 1, 2$ , com  $\|\bar{h}_N^{(k)} - f\| = 2.31 \times 10^{-2}$ . Quando  $9 \leq N \leq 12$ , para  $k = 1, 2$ ,  $\|\bar{h}_N^{(k)} - f\|$  é da mesma ordem.

Se  $\sigma = 0.5$ , vemos que precisamos escolher  $N = 7$ , para  $k = 0, 1, 2$  para atingir o menor erro, com  $\|\bar{h}_N^{(k)} - f\| = 7.87 \times 10^{-1}$ .

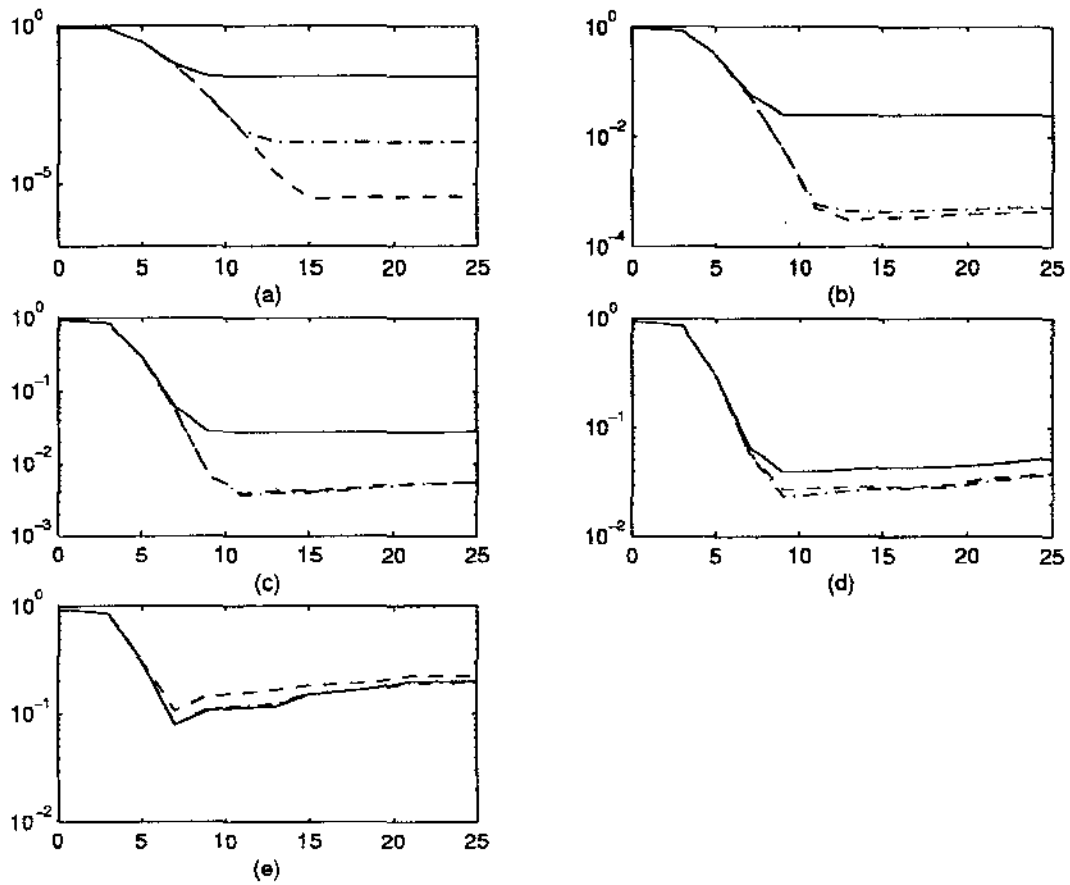


Figura 6.3: Erro  $\|\bar{h}_N^{(k)} - f\|$  em função de  $N$  para diferentes valores de  $k$ :  $k = 0$  (---),  $k = 1$  (-.-) e  $k = 2$  (- -) e diferentes valores  $\sigma$ : (a) - (e) para  $\sigma = 10^{-5}, 10^{-3}, 10^{-2}, 10^{-1}, 0.5$  respectivamente.

As Figuras 6.3 (a)-(e) mostram as reconstruções do sinal  $f(x)$  suficientemente suave; o menor erro  $\|\bar{h}_N^{(k)} - f\|$  depende de todos os parâmetros  $N$ ,  $k$  e  $\sigma$ . Para  $k = 0, 1, 2$ , quando  $\sigma$  decresce, o melhor  $N$  (com o menor erro  $\|\bar{h}_N^{(k)} - f\|$ ) aumenta. Como prevíamos,  $k = 2$  fornece sempre a melhor reconstrução, especialmente para  $\sigma$  pequeno. Portanto,  $k = 2$  deve ser a escolha para os cálculos.

O segundo sinal  $f(x)$  testado foi

$$f(x) = \begin{cases} 0, & \text{para } x \in [-1, 0), \\ 1, & \text{para } x \in [0, 1]. \end{cases} \quad (6.39)$$

Neste caso,  $f(x)$  não é contínua no ponto  $x = 0$ ; então, o fenômeno de Gibbs aparece como no capítulo anterior. A Figura 6.4 mostra os gráficos de  $f(x)$  e  $\bar{f}_N^{(k)}(x)$  com  $N = 8, 20, 30$ . A Figura 6.5 mostra  $f(x)$  e  $\bar{h}_N^{(k)}(x)$  com  $N = 8, 20, 30$ ,  $k = 2$ , e o ruído dado por  $\mathcal{N}(0, 0.1)$ .

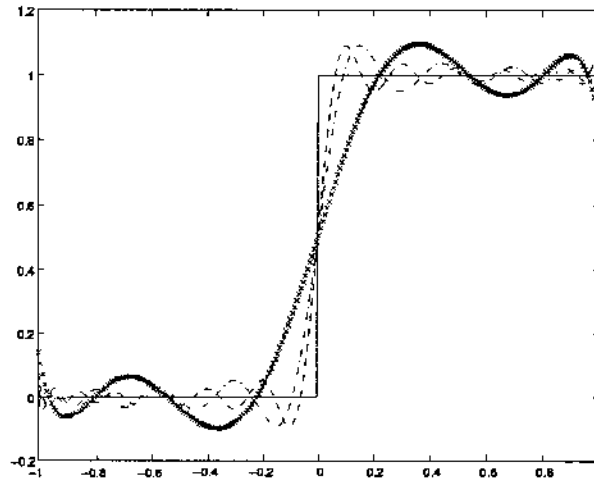


Figura 6.4: Sinal reconstruído  $\bar{f}_N^{(2)}$  como função de  $x$  para diferentes valores de  $N$ :  $N = 8$  ( $\times \times$ ),  $N = 20$  (-.-), e  $N = 30$  (- -). Sinal original é representado por “—”.

As Figuras 6.4 e 6.5 mostram o efeito do fenômeno de Gibbs. Como o erro  $\|f_N - f\|$  domina o erro da reconstrução, o algoritmo apresentado na Seção 6.2 fracassa na obtenção dos resultados exatos neste caso, o que indica a necessidade de algum tipo de filtro.

Neste capítulo, apresentamos três algoritmos diferentes ( $k = 0, 1, 2$ ) para recuperar um



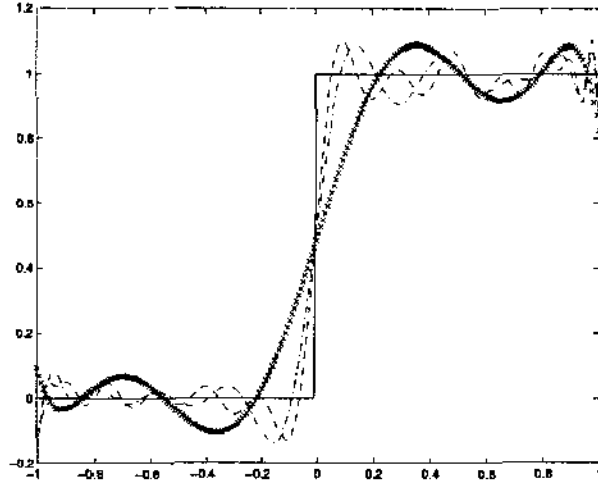


Figura 6.5: Reconstrução  $\bar{h}_N^{(2)}$  do sinal com ruído ( $\sigma = 0.1$ ) em função de  $x$  para diferentes valores de  $N$ :  $N = 8$  ( $\times \times$ ),  $N = 20$  (-.-), e  $N = 30$  (- - ). Sinal original é por “—”.

sinal com ruído a partir de momentos ortogonais de Legendre. Analisamos o erro quadrático do sinal reconstruído pelo método dos momentos, e provamos que este erro depende da suavidade do sinal, da taxa de amostragem, da ordem da interpolação, do nível de ruído, e do número de momentos usados. Nossa estimativa, dada pelo Teorema 6.1, apresenta uma nova relação entre todos esses parâmetros. Como indicamos na Introdução, uma consequência importante desta estimativa é que o erro do sinal reconstruído não tende a infinito quando o número de momentos aumenta. Alguns experimentos ilustram este resultado. A exatidão deste resultado para dimensões maiores é direta. De mesma forma, podemos utilizar outro tipo de polinômios ortogonais em [27].

# Capítulo 7

## Conclusões

Neste último capítulo da dissertação resumimos brevemente os resultados obtidos, a sua importância e originalidade e descrevemos algumas direções que podem ser a continuação natural de nosso trabalho e que são consequência das inúmeras perguntas surgidas ao longo desta pesquisa.

Para reconstruir os valores discretos de uma função com suporte compacto  $f(x)$  a partir de um número finito dos seus coeficientes de Fourier, no Capítulo 3, deduzimos algumas relações aproximadas entre a transformada de Fourier discreta e os coeficientes de Fourier da função. No Capítulo 5, usando estas relações, apresentamos alguns algoritmos para a reconstrução de funções. No Capítulo 4, como uma aplicação direta daquelas relações, apresentamos um algoritmo para calcular os coeficientes de Fourier de  $f(x)$ . Obtemos estimativas do erro aproximado dos coeficientes de Fourier (Teoremas 3.1-3.3) e a função reconstruída (Teorema 5.1). Também, discutimos a reconstrução do sinal com ruído a partir dos momentos ortogonais e obtivemos a uma estimativa do erro muito melhor (Teorema 6.1) que as já existentes, o que pode permitir uma melhor escolha do número de coeficientes e da discretização nas aplicações.

É um problema essencial superar o fenômeno de Gibbs que aparece na aproximação de  $f(x)$  da série de Fourier finita. Os métodos tradicionais são: o método de filtragem (tanto no domínio das frequências como no tempo) e o método de Fourier-Gegenbauer. Teórica-

mente, o fenômeno de Gibbs pode ser completamente evitado, usando o método de FG. Os resultados de D. Gottlieb *et al.* em [12] são os melhores sobre como remover o fenômeno de Gibbs. Mas a quantidade de cálculos envolvidos no método de FG pode ser muito grande para algumas funções. Em alguns casos pode ser proibitivo para os computadores atuais (especialmente para dimensões maiores). Os métodos de filtragem são um caminho mais simples e prático. No nosso trabalho, os novos filtros introduzidos tem propriedades em alguns aspectos melhores que os existentes e podem abrir uma nova perspectiva no desenvolvimento de filtros alternativos ainda melhores para eliminar o fenômeno de Gibbs com um custo computacional aceitável.

Os métodos de filtragem comuns baseam-se na seguinte idéia: aumentando a taxa de decaimento dos coeficientes de Fourier, sem perder a exatidão, reduz o fenômeno de Gibbs. Claramente, o decrescimento nas frequências altas provocará perda de informação no caso de uma função descontínua. Porém esses métodos com o filtros de ordem  $p$  grande funcionam bem longe da descontinuidade; o erro da reconstrução é muito grande numa vizinhança da descontinuidade (ver o Teorema 2.1). Por outro lado, usar filtros de ordem maior depende da suavidade da função reconstruída. E ainda mais, verificamos que o erro da função reconstruída aumenta nos pontos próximos às descontinuidades, quando a ordem  $p$  aumenta (ver a Fig. 5.5).

No nosso trabalho, obtemos os filtros partindo de um novo ponto de vista, diferente daquele dos métodos tradicionais. Deduzimos uma nova relação entre a transformada de Fourier discreta de valores discretos de  $f(x)$  e os seus coeficientes de Fourier (ver o Teorema 3.1 e suas extensões nos Teoremas 3.2 e 3.3). Pelo Teorema 3.1, obtemos um novo filtro. Usando estas relações, apresentamos os algoritmos para a reconstrução de funções. O Teorema 5.1 mostra o erro da função reconstruída. Comparado com o Teorema 5.1 e 2.1, verificamos que, usando o filtro da mesma ordem  $p$ ,  $p \leq 2$ , nosso filtro é melhor em dois aspectos: a exatidão da reconstrução e os bons resultados numa vizinhança das descontinuidades (ver Exemplo 5.2, 5.4 e etc). Podemos reduzir fortemente o fenômeno de Gibbs

na função reconstruída pelo nosso método.

É natural apresentar um algoritmo eficiente para calcular os coeficientes de Fourier pelos Teoremas 3.1-3.2. Comparados com outros métodos, nossos algoritmos são tão rápidos quanto o método da FFT, mas obtemos uma melhor aproximação. As amostras da função não precisam ser equidistantes desde que os pontos de descontinuidade pertençam à partição (ver o Algoritmo 4.2). E conseguimos calcular todos os coeficientes de Fourier  $\hat{f}_k$  pelo nosso método, para qualquer inteiro  $k$ . Se  $f(x) \in C^q[0,1], 1 \leq q$ , o método dos Splines é um bom algoritmo. Mas, nosso algoritmo será melhor para uma função suave por partes. Para um função descontínua, utilizamos os limites a esquerda e a direita nos pontos de descontinuidade. Por outro lado, precisamos  $O(N^2)$  operações para  $N$  coeficientes de Fourier usando o método dos Splines quando tomamos  $N$  nós em  $[0,1]$ . Igualmente, a quantidade de cálculos será grande quando generalizamos este método para dimensões maiores. Também, fizemos simulações para duas dimensões. Dada uma função de duas variáveis, conseguimos obter seus  $N^2$  ( $N = 256$ ) coeficientes de Fourier rapidamente (não mostrada neste trabalho). Este resultado não é viável com outros algoritmos (combinação de velocidade e precisão)

Sobre a reconstrução de um sinal com ruído a partir dos momentos, apresentamos três algoritmos e suas estimativas do erro do sinal reconstruído. No Teorema 6.1 deduzimos que o erro do sinal reconstruído revela correta e claramente uma relação entre a ordem dos momentos, a suavidade de sinal, a taxa da amostra e a característica do ruído. Além disso, mostra que o erro da reconstrução para um sinal sem ou com ruído não tende a infinito se a discretização do sinal é fixa. Se o sinal  $f(x)$  é suave por partes e há finitos pontos de descontinuidade, sem ou com ruído aleatório, sempre ocorre o fenômeno de Gibbs. Neste caso, o erro  $\|f_N - f\|$  domina o erro do sinal reconstruído, e assim, os algoritmos na seção 6.2 fracassam na obtenção de um bom resultado.

Apresentamos na tese uma amostra representativa das muitas simulações numéricas efetuadas que evidenciam as vantagens de nossos novos métodos.

Ao longo de nosso trabalho de pesquisa apareceram idéias e perguntas que podem ser uma

continuação dos resultados aqui apresentados. Seguem alguns exemplos que consideramos interessantes.

1. O Teorema 3.2 pode originar uma metodologia para detectar discontinuidades. A simulação (não apresentada na tese) mostra que a eficiência do filtro depende da distância das discontinuidades à fronteira do intervalo de discretização, então, detectar as pontos de discontinuidade se torna essencial para aumentar a resolução. É possível então desenvolver um algoritmo iterativo para detectar as discontinuidades de uma função.
2. Se  $f(x)$  é uma função analítica por partes, podemos estimar o erro da reconstrução de funções novamente. Melhoraria a reconstrução numa vizinhança de discontinuidade.
3. Extensão dos resultados neste trabalho para o caso bidimensional, o que apresenta algumas complicações próprias pelo fato de ter que lidar com curvas de discontinuidades, e não apenas com pontos.
4. Considerar a possibilidade de combinar a FFT com wavelets para obter melhores resultados.
5. Aplicação dos novos algoritmos para resolver problemas reais em reconstrução de imagens e em processamento de sinais.
6. Aplicação dos novos algoritmos para resolver equações diferenciais parciais por métodos espectrais.

# Bibliografia

- [1] M.T. Abuelma'atti, "Simple method for calculating Fourier coefficients of experimentally obtained waveforms", IEE Proc.-Sci. Meas. Technol., Vol.141, No.3, (1994) 177-178.
- [2] M. Bertero, C. Demol, and E.R. Pike, "Linear inverse problems with discrete data, I: General formulation and singular system analysis", Inverse Probl., Vol. (1985) 301-330.
- [3] R.P. Boas, *Entire Functions*, Academic Press, New York, 1954.
- [4] J.F. Boyce and W. J. Hossack, "Moment invariants for pattern recognition", Pattern Recognit. Lett., Vol. 1, (1983) 451-456.
- [5] E.O. Brigham, *The fast Fourier transform*, Prentice-Hall, Englewood Cliffs, New York, 1974.
- [6] W. Cai, David Gottlieb and C. W. Shu, "On one-sided Filters for spectral Fourier approximations of discontinuous functions", SIAM J. Numer. Anal. Vol. 29, No. 4, (1992) 905-916.
- [7] C. Canuto, M.Y. Hussaini, A. Quarteroni and T. A. Zang, *Spectral Methods in Fluid Dynamics*, Springer-Verlag, 1988.
- [8] W. K. Cheung and R. M. Lewitt, "Modified Fourier reconstruction method using shifted transform samples", Phys. Med. Biol., Vol. 36, No.2, (1991) 269-277.

- [9] P. Craven and G. Wahba, "Smoothing noisy data with spline functions", *Numer. Math.* Vol. 31, (1979) 377-403.
- [10] W. Gautschi, "Attenuation factors in practical Fourier analysis", *Numer. Math.*, Vol. 18, (1972) 373-400.
- [11] D. Gottlieb and S. Orszag, *Numerical Analysis of Spectral Methods: Theory and Applications*, SIAM, Philadelphia, PA, 1977.
- [12] D. Gottlieb, C.W. Shu, "On the Gibbs phenomenon and its resolution", *Siam Rev.*, Vol. 39, (1997) 644-668.
- [13] D. Gottlieb, C.W. Shu, A. Solomonoff and H. Vandeven, "On the Gibbs phenomenon I: recovering exponential accuracy from the Fourier partial sum of the a nonperiodic analytic function", *J. Comput. Appl. Math.*, Vol. 43, (1992) 81-98.
- [14] D. Gottlieb, E. Tadmor, "Recovering pointwise values of discontinuous data with spectral accuracy", In progress and Supercomputing in computational Fluid Dynamics, Proceedings of US-Israel Workshop, Birkhäuser-Bostn, Cambridge, MA, (1984) 357-375.
- [15] G. Hammerlin, K.-H. Hoffmann, *Numerical Mathematics*, Springer-Verlag, 1991.
- [16] G. Hamming, *Numerical Methods for Scientists and Engineers*, 2nd Edit., McGraw-Hill, New York, 1973.
- [17] G.T. Herman, *Image Reconstruction from Projections: the Fundamentals of Computerized Tomography*, Academic Press, New York, 1980.
- [18] A. K. Jain, *Fundamentals of Digital Image Processing*, Englewood Cliffs, NJ: Prentice-Hall, 1989.

- [19] A. C. Kak and M. Slaney, *Principles of Computerized Tomographic Imaging*, New York: IEEE Press, 1988.
- [20] D. A. Kopriva, "A practical assessment of spectral accuracy for Hyperbolic problems with discontinuities", *J. of Scientific Computing*, Vol. 2, (1987) 249-262.
- [21] E.D. Kolaczyk, "A wavelet shrinkage approach to tomographic image reconstruction", *J. Am. Statist. Assoc.*, vol. 91, 1996.
- [22] F. Natterer, *The Mathematics of Computerized Tomography*, New York: Wiley, 1985.
- [23] E. Neuman, "Calculation of complex Fourier coefficients using natural splines", *Computing*, 29, (1982) 327-336.
- [24] M. Pawlak, "On the reconstruction aspects of moment descriptors", *IEEE, Trans. Inform. Theory*, Vol. 38, No. 6, (1992) 1698-1708.
- [25] M. A. Slonim, I. Rapoport and P. P. Biringer, "Calculation of Fourier coefficients of experimentally obtained waveforms", *IEEE Trans. Ind. Electron. Contr. Instrum.*, Vol. IECI-28, No.4, (1981) 330-335.
- [26] G. Talenti, "Recovering a function from a finite number of moments", *Inverse Probl.*, Vol.3, (1987) 501-517.
- [27] C. H. Teh and R. T. Chin, "On image analysis by the methods of moments", *IEEE Trans. Pattern and Machine Intell.*, Vol. 10, (1988) 496-512.
- [28] L. Vozovoi, M. Israeli and A. Averbuch, "Analysis and application of Fourier-Gegenbauer method to stiff differential equations", *SIAM. J. Numer. Anal.*, Vol. 33, No. 5, (1996) 1844-1863.



- [29] L. Vozovoi, A. Weill and M. Israeli, "Spectrally Accurate solution of nonperiodic differential equations by the Fourier-Gegenbauer method", *SIAM J. Numer. Anal.* Vol. 34, No. 4, (1997) 1451-1471.
- [30] H. Vandeven, "Family of spectral filters for discontinuous problems", *J. Scientific Computing*, Vol. 6, (1991) 159-192.
- [31] David V. Widder, *Advanced Calculus*, Prentice-Hall, Inc. Englewood Cliffs, N.J. 1961
- [32] J.H. Yin, Alvaro R. De Pierro and M.S. Wei, "Reconstruction of a compactly supported function from the discrete sampling of its Fourier transform", Research Report No RP/98, Institute of Mathematics, Statistics and Scientific Computing, State University of Campinas, January 1998.