



UNIVERSIDADE FEDERAL DO RIO GRANDE DO NORTE
CENTRO DE CIÊNCIAS EXATAS E DA TERRA
DEPARTAMENTO DE FÍSICA TEÓRICA E EXPERIMENTAL
PROGRAMA DE PÓS-GRADUAÇÃO EM FÍSICA

**ANÁLISE ESTATÍSTICA DE PADRÕES SÍSMICOS:
DECOMPOSIÇÃO EM MULTIESCALA**

Francisco Edcarlos Alves Leite

Natal, Dezembro de 2007

Livros Grátis

<http://www.livrosgratis.com.br>

Milhares de livros grátis para download.

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Física do Departamento de Física Teórica e Experimental da Universidade Federal do Rio Grande do Norte como requisito parcial à obtenção do grau de Doutor em Física.

Autor: Francisco Edcarlos Alves Leite

Orientador: Prof. Dr. Liacir dos Santos Lucena

Dedico este trabalho à meus avós (in memoriam);

MATERNO:

Zacarias Alves de Queiroz

Maria Alves de Queiroz

PATERNO:

Francisca Leite de Souza

Antonio Silvino de Oliveira

Sumário

Resumo	vii
Abstract	ix
Agradecimentos	xi
Epígrafe	xii
Introdução	1
Capítulo 1	11
1 Representação de Funções: aspectos teóricos	12
1.1 Introdução	12
1.2 Transformada Contínua em Ondaletas	14
1.2.1 Propriedades da Transformada Contínua em Ondaletas	17
1.2.2 Transformada inversa em ondaletas	22
1.3 Transformada discreta em ondaletas	23

1.3.1	Análise em Multiresolução	24
1.3.2	Transformada discreta em ondaletas	27
1.3.3	Transformada rápida em ondaletas	29
Capítulo 2		34
2	Transformada de Karhunen-Loève	35
2.1	Introdução	35
2.2	Definição e propriedades principais	36
2.3	Transformada de Karhunen-Loève e Decomposição em Valores Sin- gulares	40
2.3.1	Decomposição em Valores Singulares e Aproximação de matrizes	41
Capítulo 3		42
3	Sondagem Sísmica	43
3.1	Introdução	43
3.2	Prospecção Sísmica	44
3.3	Sísmica de Reflexão	49
3.3.1	O Ruído de Rolamento Superficial	51
3.3.2	As reflexões	52
3.4	Técnicas de Filtragem	53
Capítulo 4		54

4	Supressão do <i>Ground roll</i>: um filtro ótimo	55
4.1	Introdução	55
4.1.1	Ondaletas: passagem do contínuo ao discreto	57
4.1.2	Transformada de Karhunen-Loève	59
4.2	O filtro ótimo em ondaletas	59
4.2.1	Resultados	63
4.3	Comentários finais	66
	Capítulo 5	67
5	Supressão do <i>Ground roll</i>: decomposição em multiescala	68
5.1	Introdução	68
5.1.1	Remoção do <i>ground roll</i>	72
5.2	Remoção do <i>ground roll</i> : decomposição em multiescala	73
5.3	Comentários finais	79
	Conclusões	89
	Apêndices A, B e C	90
A	Análise de Fourier	91
A.1	Introdução	91
A.1.1	Notação	92
A.2	Transformada de Fourier	93
A.2.1	Transformada de Fourier-Gabor	95

A.3	Comentários	96
B	Propriedades dos Filtros h e g	99
B.1	O Filtro de aproximação h	99
B.2	O Filtro de Detalhe g	100
C	Funções Ondaletas	102
C.1	Introdução	102
C.2	Exemplo de ondaletas	103
C.2.1	Ondaletas derivadas da Gaussiana	103
C.2.2	Ondaletas de Morlet	104
C.2.3	Ondaletas de Shannon	104
C.2.4	Ondaletas de Daubechies	105
C.2.5	Coeficientes $h[k]$ das ondaletas de Daubechies	105
	Lista de Figuras	115
	Referências Bibliográficas	116

Resumo

O processamento de registros sísmicos é uma tarefa muito importante dentro da Geofísica e que representa um desafio permanente na exploração de petróleo. Embora esses sinais forneçam uma imagem adequada da estrutura geológica do subsolo, eles são contaminados por ruídos e, o *ground roll* é o componente principal. Este fato exige um esforço grande para o desenvolvimento de metodologias para filtragem. Dentro desse contexto, este trabalho tem como objetivo apresentar um método de remoção do ruído *ground roll* fazendo uso de ferramentas da Física Estatística. No método, a Análise em Ondaletas é combinada com a Transformada de Karhunen-Loève para a remoção em uma região bem localizada.

O processo de filtragem começa com a *Decomposição em Multiescala*. Essa técnica permite uma representação em tempo-escala fazendo uso das ondaletas discretas implementadas a filtros de reconstrução perfeita. O padrão sísmico original fica representado em multipadrões: um por escala. Assim, pode-se atenuar o *ground roll* como uma operação cirúrgica em cada escala, somente na região onde sua presença é forte, permitindo preservar o máximo de informações relevantes. A atenuação é realizada pela definição de um *fator de atenuação* A_f . Sua escolha é feita pelo com-

portamento dos modos de energia da Transformada de Karhunen-Loève. O ponto correspondendo a um mínimo de energia do primeiro modo é identificado como um fator de atenuação ótimo.

Abstract

Seismic signal processing is an important task in geophysics sounding an represents a permanent challenge in petroleum exploration. Although those signals could give a picture of a geological structures of the underground, they are very contaminated by noises and the ground roll is the main component. This fact demands a big effort in developing new filtering methodologies. In this context, the main purpose of the work is present a method to remove the noise using of tools of the Statistical Physics. In the method, the analysis in wavelets is combined with Karhunen-Loève Transform for to removal in a given located region.

The process filtering begins with the *Multiscale Decomposition*. That technique allows a representation in time-scale using discrete wavelet implemented to filters of perfect reconstruction. The original seismic pattern is represented in multiple-pattern: one for scale. So, the ground roll is can attenuate as a surgical operation in each scale, only in the region where your presence is strong, allowing to preserve maximum of important information. The attenuation is reached by the definition of a *attenuation factor* A_f . It's choice is made for the behavior of the energy modes of the Transforms in Karhunen-Loève. The point corresponding to a minimum of

energy of the first mode is identified as a optimal attenuation factor.

Agradecimentos

- A **Deus**, pela vida, pela vontade, pela esperança, pois tornaram possível a conclusão deste trabalho.
- Aos meus pais, **Ester Alves** e **Sebastião Leite** pelo apoio na escolha profissional que fiz e pela pureza do amor.
- A minha irmã **Marta Alves Leite** e ao meu irmão **Erizomar Alves Leite** pelo carinho e incentivos, mesmo quando ausentes.
- Ao **Professor Liacir dos Santos Lucena**, por sua orientação, pela dedicação e pela oportunidade para que pudesse realizar este trabalho.
- Ao **Professor Gilberto Corso** e ao **Dr. Raúl Montagne**, pela orientação, esforços e esclarecimentos, os quais auxiliaram valiosamente em todo o desenvolvimento deste trabalho.
- A todos os **Professores do DFTE**, que contribuíram para a minha formação acadêmica e profissional.
- Aos **funcionários do DFTE**, cujo apoio foi de grande importância, pelos serviços prestados e sobretudo a amizade.
- Ao **CNPq** pelo apoio financeiro.

Epígrafe

Deus no teto do Céu não botou telhas pra dar sombra ao pináculo das montanhas, não tem teia que vença a das aranhas, nem tem lã que supere a das ovelhas. No cortiço a indústria das abelhas o zangão não faz nada nessa empresa, a rainha com pompas de princesa sem tempero no mel bota sabor... pra quem sente o poder do criador não tem nada maior que a natureza.

Nonato Neto.

Pra quem tem consciência do poder do maior arquiteto universal, imagine se o mar fosse sem sal tudo insosso teríamos que comer. Se as nuvens parassem de chover faltaria alimento em nossa mesa, se a lanterna do sol não fosse acesa quem faria outra fonte de calor... pra quem sente o poder do criador não tem nada maior que a natureza.

Nonato Costa.

Os Nonatos, Mote em 10.

Introdução

Em processamento de sinais, faz-se uso de ferramentas da Física Matemática para extrair características relevantes visando a melhoria da interpretação dos sinais. Melhorias podem ser obtidas pelo uso da Transformada de Fourier (TF). Essa ferramenta é adequada para se analisar sinais estacionários e que faz uso das funções harmônicas (senos e co-senos) como *Funções de Análise* para a base (funções que compõem a base). Quando se aplica a transformada de Fourier sobre um sinal $f(t) \in L^2(\mathbb{R})$, é obtida uma representação “nova” definida sobre o $L^2(\mathbb{R})$ (veja equação A.6). A representação “nova” $\hat{f}(\omega)$ é completamente caracterizada por seu conteúdo na frequência, mas nenhuma informação sobre a localização no tempo, pois as funções de análise que formam a base possuem suporte infinito no domínio temporal. Portanto, para analisar sinais não-estacionários, isto é, sinais que apresentam variações no tempo de suas propriedades estatísticas (como acontece com sinais sísmicos na prospecção de petróleo), as metodologias que fazem uso da transformada de Fourier, são razoavelmente precárias. A limitação encontra-se no fato de que a transformada de Fourier não permite analisar, em separado, trechos diferentes dos sinais [Grossmann et al. (1984)]. Caso um trecho seja extremamente ruidoso ou

contenha pontos anômalos, o processamento de todo o sinal ficará comprometido. Para suprir essa preacriedade, Gabor modificou a transformada de Fourier [Gabor (1946)]. A alternativa de Gabor foi usar uma função auxiliar $g(t)$ (equação A.17) no integrando da transformada de Fourier, delimitando-a no tempo (*Short Time Fourier Transform* ou ainda *Transformada de Fourier Janelada*¹ - (TFJ)). Para cada valor de $t \in \mathbb{R}$, a exponencial (na transformada de Fourier) é multiplicada pela função $g(t - u)$ com certa localização no tempo. Da definição da equação (A.16), se a função² $f \in L^2(\mathbb{R})$, então $\hat{f}(\omega, t) \in L^2(\mathbb{R})$. Portanto, a transformada de Fourier “janelada” está definida no domínio tempo-frequência (ω, t) .

A “janela” usada na TJJ é de tamanho fixo. Isto torna difícil a análise das componentes de altas e de baixas frequências simultaneamente (Figura 1.2) pois, há um limite na precisão da localização deste domínio. Esta limitação está no princípio (que governa as transformadas tempo-frequência) conhecido como *Princípio da Incerteza de Heinsenberg*. O tamanho do intervalo da localização é dado pela equação (A.15). Note que uma localização “boa” na frequência (σ_ω) não se tem uma localização “boa” no tempo $\sigma_t \geq 1/4\pi\sigma_\omega$. A localização da função no domínio em tempo-frequência está representada geometricamente pela dimensão do retângulo $\sigma_t \times \sigma_\omega$. Do Princípio da Incerteza, a área deste retângulo é $\geq 1/4\pi$, como mostrada na Figura (A.1).

A limitação de análise da Transformada de Fourier faz surgir a Transformada em Ondas [Meyer (1993)]. Essa transformada faz uso das *Funções de Análise*³

¹Este termo *janelada* tem concordância com a delimitação no tempo para a função auxiliar $g(t)$

²No contexto deste trabalho, a definição de função e sinal é semelhante

³As funções que formam a base são chamadas de ondas.

com “janelas” de larguras variáveis (Figura 1.2). Assim, as ondaletas podem ser melhor ajustadas às características de cada trecho da função. Em particular, a transformada em ondaletas é de interesse na análise de funções não-estacionárias, tais como processamentos de dados geofísicos [Hoekstra (1996), Deighan and Watts (1997), Galvão (2001), Corso et al. (2003)] onde é possível uma representação adequada em níveis de resoluções diferentes (Capítulo 1). Esse fato foi convenientemente explorado por Mallat [Mallat (1989)] e Meyer [Meyer (1990)], que lhes permitiram sistematizar o conceito de *Análise em Multiresolução* (AM). Esse método de análise constitui um enquadramento ideal que facilita a construção e aplicação de algoritmos baseados nas ondaletas [Meyer (1990); Daubechies (1988)]. A aplicação destes algoritmos traz solução a problemas de filtragem e processamento de sinais, imagens. Em processamento de imagens, os algoritmos propostos por Mallat, permitem decompor uma imagem inicial em uma outra de mesmo padrão (de imagem), porém, com um nível de resolução inferior (o que corresponderá a uma imagem menos nítida). Simultaneamente, são obtidos os detalhes (coeficientes no espaço das ondaletas), os quais permitirão recompor a imagem inicial, partindo das informações contidas nos padrões da imagem decomposta.

Revisão Bibliográfica das Ondaletas

No início do século XX, Alfred Haar [Haar (1910)] definiu um sistema de funções que constituíam uma base ortonormal gerada por dilatações e translações de uma função protótipo⁴. Esse conjunto completo de funções ortogonais é con-

⁴Função de análise a qual gera uma base do $L^2(\mathbb{R})$ por suas translações e dilatações.

siderado hoje, na literatura da matemática, como o primeiro sistema de ondaletas construído no $L^2(\mathbb{R})$. Embora possuindo uma definição bastante simples (as funções são constituídas por ondas retangulares), o *Sistema de Haar*, como era chamado na época, possui as propriedades e características, que hoje em dia, definem as ondaletas. Mesmo assim, a formalização da teoria das ondaletas foi desenvolvida recentemente com base na generalização de conceitos já conhecidos, oriundos de diversos campos da ciência. Foi somente na década de oitenta que se definiu com rigor os conceitos que permitem compreender, de forma clara e, onde, as ondaletas passaram a ter uma identidade própria. Envolvidos neste trabalho pioneiro, estiveram vários investigadores. Destacam-se, entre outros, os trabalhos de Morlet e Grossmann [Grossmann et al. (1984)], Daubechies [Daubechies (1988)], Mallat [Mallat (1989)], Meyer [Meyer (1990)].

A importante contribuição de Morlet se deu no processamento sísmico. Foi trabalhando com sinais geofísicos, em uma companhia de petróleo, que Jean Morlet deparou-se com a análise de sinais não-estacionários, os quais apresentavam, ora componentes de frequências altas para um tempo curto ora componentes de frequências baixas para um tempo longo. Com a Transformada de Fourier Localizada Morlet foi capaz apenas de analisar as componentes de altas frequências usando “janelas” estreitas (para a função auxiliar $g(t)$) ora componentes de baixas frequências usando “janelas” largas, mas não ambas. Morlet então propôs uma modificação na transformada de Fourier para melhorar a análise dos sinais sísmicos. Ele usou “janelas” variáveis (para a função de análise) para as diferentes bandas de frequências analisadas. Essas “janelas” são geradas por dilatações ou compressões

de uma função Gaussiana protótipo. Devido à natureza dessas funções “janeladas”, Morlet as nomeou de ondaletas⁵. Convencido de suas idéias, Morlet recorreu ao físico teórico Alex Grossmann, o qual firmou uma base matemática para suas teorias. Anos depois, as ondaletas de Morlet, atraíram a atenção do matemático Yves Meyer. Fascinado pelo esquema elegante de análise de funções não-estacionárias, Meyer ajudou no enriquecimento da teoria nova, conciliando-a a outros campos da Matemática antes estudados separadamente.

Talvez a contribuição mais importante para a análise moderna em ondaletas, tenha sido o trabalho de Daubechies sobre *Bases Ortonormais de Ondaletas com Suporte Compacto* que contém a ondaleta de Haar como um caso particular. Daubechies definiu suas ondaletas sobre um intervalo limitado (suporte compacto) e ao mesmo tempo garantiu a ortogonalidade entre elas [Daubechies (1988)]. Com essas propriedades, pode-se descrever os detalhes de uma função convenientemente e com certa precisão. Outra vantagem, acerca da ortogonalidade, é que um erro na função de entrada não crescerá com a transformada, mantendo assim, a estabilidade numérica computacional. Uma outra contribuição muito importante foi a de Mallat sobre *Análise Multiresolução* (seção 1.3.1). Em seus trabalhos, Mallat introduziu um algoritmo (conhecido como algoritmo de decomposição em árvore e/ou piramidal) que permitiu a computação fácil da transformada em ondaletas implementada a filtros de reconstrução perfeita.

Atualmente, existem várias ondaletas construídas em $L^2(\mathbb{R})$ que podem ser usadas eficientemente, mas que todas parecem ter as mesmas características gerais

⁵O autor desta tese escolheu esse termo. Porém, do inglês, *wavelets* que significa ondas pequenas.

(veja Apêndice C). Os que as distinguem são, fundamentalmente, as suas maiores ou menores localizações no domínio espaço-frequência, as suas continuidades e as suas diferenciabilidades.

Ondaletas e a Sísmica do Petróleo

Os dados sísmicos⁶ obtidos de uma exploração sísmica feita em terra, em geral, contêm vários componentes ondulatórios indesejáveis. A mais proeminente é a presença de ondas superficiais de grande amplitude e baixa frequência. Essas ondas são denominadas de rolamento superficial (*ground roll*). Sua característica forte é que não penetram no interior da Terra e, no entanto, não conduzem informações relacionadas com as interfaces refletoras. Essa onda é propriamente caracterizada como ruído⁷ e precisa ser eliminada dos registros sísmicos. Uma análise sísmica, portanto, torna-se necessária [Yilmaz (2003)] e as ferramentas usuais neste processo são filtros “passa alta” e filtros que dependem da eliminação de frequências (ou número de onda) específicas baseadas na Transformada de Fourier. Isto não representa grande dificuldade quanto ao ruído *ground roll* em si, pois realmente apresenta baixas frequências e decaimento lento. Entretanto, essa supressão afeta o sinal sísmico como um todo, “borrando” as reflexões que podem conter as informações sobre a formação geológica do meio terrestre. Isso acontece porque as reflexões são componentes que decaem rapidamente durante as medições e, por este motivo, não são fáceis de serem separadas no domínio da frequência. Isto corresponde a afirmação comumente feita de

⁶As expressões registro sísmico, dados sísmicos ou sismograma, são utilizadas aqui para definir um conjunto de informações (sinais) oriundas do subsolo terrestre através de vibrações do meio. As representações destes sinais podem ser vistas nas figuras (3.3) e (5.1).

⁷Aqui o termo ruído para o *ground roll*, significa que este é um sinal indesejável.

que o sinal sísmico, como um todo, é não-estacionário. Então, os senos e cossenos que têm localização na frequência, mas não no tempo, não constituem uma base apropriada para tratar as componentes responsáveis pela não-estacionariedade, como é o caso das reflexões.

A não-estacionariedade do sinal sísmico já conduz à deficiência na técnica baseada na Transformada de Fourier, pois a mesma assume que o sinal seja estacionário. Um procedimento alternativo consiste em decompor o sinal sísmico em uma base formada por ondaletas, cujos vetores da base tem alcance finito. A transformada em ondaletas é uma transformada localizada que leva em conta as propriedades tempo-frequência do sinal. Portanto, a maior vantagem de uma análise com ondaletas sobre Fourier é a sua capacidade de uma análise local distinta [Deighan and Watts (1997)]. Esta propriedade, juntamente com uma implementação fácil e eficiente de um algoritmo computacional, são vantajosas para se extrair informações de um sinal, as quais não podem ser obtidas via a transformada de Fourier.

Objetivos

O objetivo deste trabalho é apresentar o método de análise que desenvolvemos para o processamento dos sismogramas e que permite uma caracterização detalhada destes sinais sísmicos. Em certo grau, no contexto tratado aqui, o foco está dentro da Geofísica, especificamente, na exploração de reservatórios de petróleo que está relacionada à Física de Sistemas Complexos. O método apresentado trata-se de uma técnica de processamento que possibilita a extração de características mensuráveis que serão analisadas pela mudança de escala visando em se obter uma

representação (visualização) mais eficiente. Contudo, a atenção maior está voltada para a eliminação do ruído *ground roll* e a base desta técnica está na Análise em Tempo e Frequência utilizando a ferramenta da Física Matemática que é a Transformada em Ondaletas Discretas. Uma análise quantitativa e qualitativa dentro do escopo da Física Estatística é utilizada buscando uma identificação padrão para os diversos eventos coerente presentes nos sismogramas. A fundamentação teórica para o método não é nova, porém a novidade introduzida está, amplamente, direcionada à técnica da *Decomposição em Multiescala*.

Certamente, a decomposição em multiescala por meio da Transformada em Ondaletas consiste em realizar uma decomposição (projeção) do sinal em uma base de ondaletas discretas. Esta distribuição em coeficientes ondaletas fornece informações em componentes diferentes por escalas. A idéia desta representação é a separação do registro sísmico em várias partes (cada parte indexada por uma escala) seguida de uma análise em separado de cada uma das partes. Dessa forma, coeficientes que representam sinais ruidosos são desprezados (zerados), restando apenas coeficientes que fornecem informações mais relevantes sobre o sinal de interesse físico geológico. O último passo é a aplicação da transformada inversa em ondaletas para recuperar o sinal, já com o ruído atenuado.

Uma ferramenta auxiliar à transformada em ondaletas utilizada para o processamento dos sismogramas é a transformada de Karhunen-Loève (KL). Essa transformada é muito importante devido a sua adaptatividade para uma expansão em série. O conceito de adaptatividade significa que o sinal original é decomposto em uma base obtida diretamente do sinal original. Com um procedimento matemático

de implementação fácil (Decomposição em Valores Singulares - SVD) a transformada KL é usada para se obter uma decomposição (do sismograma) “ótima” em um número finito de modos de energia (Modos Ortogonais Próprios - POM, Funções Empíricas Ortogonais - EOF ou auto-imagens) os quais são obtidos da matriz de correlação do dado original.

Organização desta Tese

Esta tese está dividida em: **Introdução**, **5 Capítulos**, **Conclusões e Perspectivas** e os **Apêndices A, B, C**. No **Primeiro Capítulo**, *Representação de Funções: aspectos teóricos*, é feita uma introdução sobre a teoria da análise de sinais estacionários, com referência a aplicação da Transformada de Fourier bem como suas falhas na análise de sinais não-estacionários. Sem perda de generalidades, procuramos descrever, também, de forma sucinta a necessidade da análise de sinais não-estacionários baseada na teoria da transformada em ondaletas. Essa transformada é introduzida com o propósito de se obter uma representação melhor para uma função $f(t) \in L^2(\mathbb{R})$ usando a decomposição no domínio tempo-freqüência. A Análise em Multiresolução, que é fundamental no domínio das ondaleta, é apresentada na seção (1.3.1) bem como sua implementação computacional fácil que faz da transformada em ondaletas uma ferramenta muito eficiente em análise de sinais não-estacionários.

O **Capítulo 2** traz a Transformada de Karhunen-Loève. Entre as diferentes abordagens para análise de dados complexos o uso dessa transformada mostra ser bastante útil. Essa ferramenta é uma técnica bem conhecida e largamente usada em

muitos campos, tais como oceanografia, meteorologia e processamento de imagens. Neste capítulo, fazemos uma descrição breve, e mais detalhes sobre essa transformada podem ser encontrados nas referências [Sirovich (1987), Holmes et al. (1996), Jolliffe (2002)].

O método da sondagem sísmica é apresentado no **Capítulo 3** como a técnica indispensável na caracterização do meio geológico para a localização de jazidas de petróleo. Desde os levantamentos até as interpretações, as propriedades e componentes dos registros sísmicos são colocadas por uma visão geral, principalmente, para quem não atua diretamente nesse campo. Como sendo a técnica principal na obtenção das imagens sísmicas, a reflexão sísmica predomina-se na obtenção dos sinais de interesse (reflexões) e, também, de sinais espúrios (*ground roll*). Esse capítulo, também, tem por objetivo descrever como se representa um sismograma no formato $2D$, gerado pela formação ordenada de geofones.

O **Capítulo 4** apresentamos a construção de um filtro para a supressão do ruído *ground roll*. O potencial desse filtro está no incremento das Transformadas em Ondas com Karhunen-Loève, o qual resultou no artigo (que pode ser verificado nas Referências Bibliográficas desta tese) *Optimal wavelet filter for suppression of coherent noise with an application to seismic data*[Leite et al. (2008)].

No **Capítulo 5** expomos o objetivo principal do método de filtragem baseado na Decomposição em Multiescala (veja na referência desta tese *Efficient selective filtering of seismic data using multiscale decomposition*). Usando a Transformada em Ondas Discretas, este método apresenta-se como um filtro seletivo e eficiente, uma vez que o registro sísmico original fica decomposto em padrões semelhantes

(Figuras (5.5)-(5.11)), porém em escalas individuais. Dessa forma, o ruído *ground roll* fica mais localizado e a operação de remoção deste, não afeta as regiões de interesse.

Por último, as Conclusões e Perspectivas são apresentadas seguidas pelos apêndices. O **Apêndice A** traz a ferramenta mais tradicional para análise de funções, que é a Análise de Fourier. No **Apêndice B** as propriedades dos filtros h e g são apresentadas e no **Apêndice C**, as formulações matemáticas, gráficas, e algumas propriedades das ondaletas são expostas. Bem como as ondaletas discretas de Daubechies.

Capítulo 1

Representação de Funções: aspectos teóricos

1.1 Introdução

Uma grandeza física pode ser representada por uma função $f(t)$ e esta notação sugere que a grandeza deve ser conhecida para qualquer t . Contudo, quando se deseja obter informações adicionais, as quais não estão disponíveis em seu formato original (domínio temporal), é necessário realizar uma transformação matemática sobre $f(t)$. Dessa forma, a função $f(t)$ pode ser analisada em um outro domínio que é conseguida pela decomposição (produto interno) da função em uma base formada por *funções de análise* conhecidas. Na Transformada de Fourier (Apêndice A), as informações adicionais ocorrem no domínio da frequência. Porém, uma informação localizada em um tempo específico não pode ser “vista” facilmente devido a extensão infinita das funções senos e co-senos que formam a base. Para cobrir a deficiência na TF e usando uma técnica de “janelamento”, Gabor [Gabor (1946)] adaptou essa transformada (transformada de Fourier-Gabor) para analisar regiões em uma seção

(parte) de uma função. De certa forma, a adaptação de Gabor fornece informação sobre onde e qual frequência de um dado evento da função. Entretanto, esta informação é obtida de forma limitada que depende do tamanho da “janela” usada, isto é, o tamanho fixo escolhido para a “janela” irá percorrer a função $f(t)$ para todas as frequências. Devido a esse desconforto, Grossmann e Morlet [Grossmann et al. (1984)], interessados na análise de funções não-estacionárias, usaram funções de análise na base com parâmetros variáveis¹ σ e τ , cuja principal meta era contornar as limitações da análise de Fourier-Gabor. Agora, as funções de análise são “ajustadas” por esses parâmetros (no tempo em τ e na escala em σ) dependendo da informação na frequência que se deseja analisar. Isto é, para analisar as estruturas de funções $f(t)$ em tamanhos diferentes, é necessário usar uma decomposição em tempo-frequência com suportes diferentes no tempo.

Neste capítulo, a Transformada em Ondaletas é introduzida como a ferramenta matemática utilizada para se obter uma representação adequada para funções $f(t) \in L^2(\mathbb{R})$ no domínio em tempo e escala (tempo-frequência). Sendo uma técnica recente da Física-Matemática e com ampla aplicação na Física Estatística [Arneodo et al. (1994, 1995)], essa transformada tem grande aplicabilidade no estudo das propriedades das funções não-estacionárias bem como nos fenômenos com transientes, como é o caso dos sinais geofísicos [Deighan and Watts (1997)]. A utilidade das ondaletas está na sua possibilidade de atuarem como funções base na decomposição de

¹Os parâmetros σ e τ , como veremos adiante, estão relacionados com frequência e tempo, respectivamente. Essas funções com parâmetros variáveis são as ondaletas.

²Na verdade, essa transformada será utilizada para a decomposição em ondaletas dos sismogramas.

funções $f(t) \in L^2(\mathbb{R})$. As bases em ondaletas cobrem todo o espaço $L^2(\mathbb{R})$, de uma forma mais eficiente que as bases senoidais do método de Fourier. Muitas vezes, a transformada em ondaletas é referenciada a um “microscópico matemático” [Arneodo et al. (1994)], devido a sua versatilidade de localização no tempo e escala. Assim, esta análise é bastante conveniente para funções não-estacionárias, as quais são encontradas e com diversas aplicações em vários ramos da ciência [Kobayashi (2001)].

1.2 Transformada Contínua em Ondaletas

Definição 1.1 (Condição de Admissibilidade) *Seja $\psi \in L^2(\mathbb{R})$ uma função assumindo valores reais. Esta função é uma ondaleta se satisfaz a condição de admissibilidade*

$$C_\psi \equiv \int_0^{+\infty} \frac{|\hat{\psi}(\omega)|^2}{|\omega|} d\omega < +\infty, \quad (1.1)$$

onde C_ψ depende da ondaleta usada e $\hat{\psi}$ é a Transformada de Fourier de ψ .

Para garantir que a integral acima seja finita é necessário que $\hat{\psi}(0) = 0$. Se $\hat{\psi}(0) = 0$ e $\hat{\psi}(\omega)$ é continuamente diferenciável então a condição de admissibilidade é satisfeita. A condição $\hat{\psi}(0) = 0$ impõe que as ondaletas tenham média nula, ou seja; $\int \psi(t) dt = 0$. Esta integral nula significa que as ondaletas mudam, ao menos uma vez, seu sinal ao longo da reta real (caráter oscilatório) e, vão à zero quando $t \rightarrow \infty$. Com a definição e condições apresentadas acima, pode-se definir a Transformada Contínua em Ondaletas em $1D$ para uma função $f(t) \in L^2(\mathbb{R})$.

Definição 1.2 (Transformada Contínua em Ondaletas, [Mallat (1999)]) *A Transformada Contínua em Ondaletas $Wf(\sigma, \tau)$ de uma função $f(t) \in L^2(\mathbb{R})$ é calculada pelo produto interno de f com as ondaletas na posição τ e escala σ :*

$$Wf(\sigma, \tau) = \int_{-\infty}^{+\infty} f(t)\psi_{\sigma, \tau}^*(t)dt, \quad (1.2)$$

onde

$$\psi_{\sigma, \tau}(t) \equiv \frac{1}{\sqrt{|\sigma|}}\psi\left(\frac{t - \tau}{\sigma}\right). \quad (1.3)$$

O produto interno na equação (1.2) fornece os *coeficientes em ondaletas*. Então, de certa forma, os coeficientes representam a correlação entre a função $f(t) \in L^2(\mathbb{R})$ e a base particular formada por ondaletas. A função ψ gera uma base do $L^2(\mathbb{R})$ por translações e dilatações da função ³protótipo $\psi(t)$ (equação (1.3) é uma família de ondaletas). Essa transformada gera uma nova função $Wf(\sigma, \tau)$ a qual é uma projeção da função $f(t)$ em um subespaço do $L^2(\mathbb{R})$ (espaço das ondaletas) e que, agora, depende dos parâmetros de dilatação/contração ⁴ σ e translação ⁵ τ . Como será mostrado na seção (1.3.1), esses parâmetros assumirão valores numéricos especiais.

A Figura (1.1)(topo) está mostrando um exemplo de uma família de ondaletas contínuas para 3 valores diferentes dos parâmetros de escala σ e localização τ no domínio temporal. Esta ondaleta é derivada segunda da gaussiana⁶. A ondaleta protótipo (geradora) é ψ_1 e está localizada em $\tau_0 = 0$ para uma determinada escala σ_0 , ou seja, $\psi_1 = \psi_{\sigma_0, \tau_0=0}$. Observa-se que quando ψ_1 é contraída em $\sigma_0 \rightarrow \frac{\sigma_0}{2}$ e transladada para à direita em $\tau_0 \rightarrow +\tau$ é gerada a on-

³Função semelhante ao “átomo de Gabor” para a Transformada de Fourier-Gabor.

⁴Parâmetro relacionado com a frequência e que não aparece na transformada de Fourier-Gabor

⁵Parâmetro similar ao parâmetro de localização da transformada de Fourier-Gabor

⁶Essa ondaleta também é conhecida como *Chapéu Mexicano* devido seu formato.

ondaleta ψ_2 , caracterizada por $\psi_2 = \psi_{\sigma_0/2,\tau}$. A ondaleta ψ_3 foi gerada por uma contração em $\sigma_0 \rightarrow \frac{\sigma_0}{4}$ e uma translação em $\tau_0 \rightarrow -\tau$ para à esquerda. Sendo assim, ψ_3 é caracterizada por $\psi_3 = \psi_{\sigma_0/4,-\tau}$. O espectro de Fourier, para cada ondaleta, é mostrado na Figura (1.1)(baixo). Observa-se que quando é diminuída a análise no domínio temporal, isto é, a ondaleta possui suporte pequeno no tempo (a ondaleta ψ_1 é mais “espalhada” que ψ_2 e ψ_3), a análise no domínio da frequência é deslocada para frequências maiores. Ao contrário, a ondaleta que possui suporte maior no tempo o seu espectro é localizado em frequências menores.

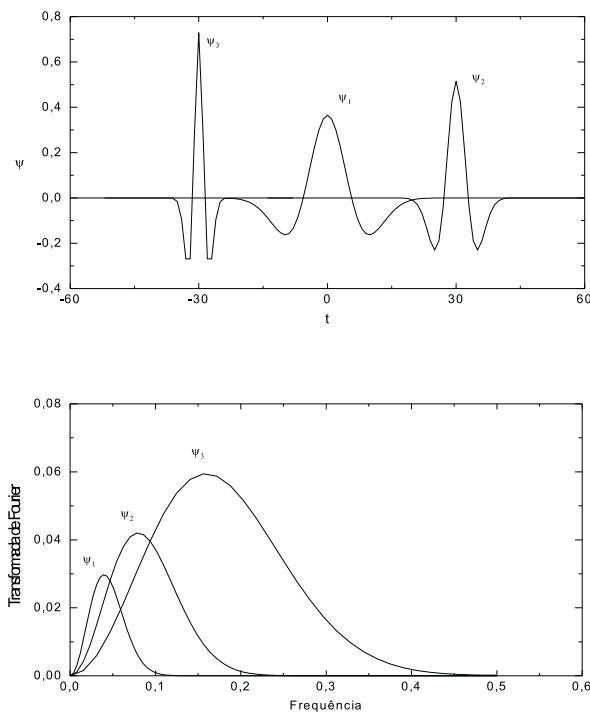


Figura 1.1: Ondaleta normalizada ψ_1 e suas dilatações/translações para valores diferentes de σ e τ . A ondaleta ψ_1 é a derivada segunda da gaussiana. Tem-se $\psi_3 = \psi_{\sigma_0/4,-\tau}$, $\psi_1 = \psi_{\sigma_0,\tau_0=0}$, $\psi_2 = \psi_{\sigma_0/2,\tau}$ e (b) seus espectros de frequências.

1.2.1 Propriedades da Transformada Contínua em Ondaletas

Nesta seção apresento algumas propriedades básicas da transformada contínua em ondaletas.

- **Linearidade.** A linearidade da transformada segue da linearidade do produto interno;

$$W\{\alpha f + \beta g, \psi\}(\sigma, \tau) = \alpha W\{f, \psi\}(\sigma, \tau) + \beta W\{g, \psi\}(\sigma, \tau). \quad (1.4)$$

- **Invariância por Translação.** Se a transformada contínua em ondaletas da função $f(t)$ é dada por $Wf(\sigma, \tau)$, então $f'(t) \rightarrow f(t - \tau')$ tem a seguinte transformada;

$$Wf'(\sigma, \tau) \rightarrow Wf(\sigma, \tau - \tau'). \quad (1.5)$$

- **Invariância por Escala.** Se a transformada contínua em ondaletas da função $f(t)$ é dada por $Wf(\sigma, \tau)$, então $f'(t) \rightarrow \frac{1}{\sqrt{\sigma_0}} f(\frac{t}{\sigma_0})$, com $\sigma_0 > 0$, tem a seguinte transformada;

$$Wf'(\sigma, \tau) \rightarrow Wf\left(\frac{\sigma}{\sigma_0}, \frac{\tau}{\sigma_0}\right). \quad (1.6)$$

- **Conservação da Energia.** A transformada contínua em ondaletas tem a propriedade da conservação da energia;

$$\langle f(t), f(t) \rangle = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \frac{1}{C_\psi} \int_0^{+\infty} \int_{-\infty}^{+\infty} |Wf(\sigma, \tau)|^2 d\tau \frac{d\sigma}{\sigma^2}, \quad (1.7)$$

a qual é similar a fórmula de Plancherel da transformada de Fourier (veja equação (A.9)).

Propriedade de localização em tempo-frequência

A função ψ é de quadrado integrável e o termo $|\sigma|^{-1/2}$, na equação (1.3), corresponde a um fator de normalização. No entanto, sua energia está localizada em uma região do plano (τ, ω) e, portanto, no plano (τ, σ) , uma vez que $\sigma \sim 1/\omega$. Isto significa que as amplitudes de ψ são apreciáveis apenas nesta região. Dessa forma, a equação (1.2) é um produto escalar e que mede as flutuações da função f na vizinhança de τ , cujo tamanho é proporcional a escala σ . Para um melhor entendimento da localização de uma ondaleta $\psi(t)$ são definidas uma localização central τ_0 , uma frequência central ω_0 e as “larguras” de ψ nos dois domínios, σ_t e σ_ω . Estas definições são similares as definições apresentadas no Apêndice A (equações (A.11-A.14)) para as quantidades médias da localização e frequência. Supondo a seguinte normalização $\|\psi(t)\|^2 = \|\hat{\psi}(\omega)\|^2 = 1$ e que $\psi(t)$ é dada para $\tau_0 = 0$ e para ω_0 , σ_t e σ_ω com uma estimativa que pode ser obtida da localização em ambos os domínios por $\sigma_t \sigma_\omega \geq \frac{1}{2}$. Dessa forma, para a ondaleta localizada na origem, a localização no domínio temporal corresponde ao intervalo

$$\left(-\frac{1}{2}\sigma_t, \frac{1}{2}\sigma_t\right), \quad (1.8)$$

e para a localização no domínio da frequência corresponde ao intervalo

$$\left(\omega_0 - \frac{1}{2}\sigma_\omega, \omega_0 + \frac{1}{2}\sigma_\omega\right). \quad (1.9)$$

Com as propriedades de dilatações e translações, os limites acima mudam mas, que ainda, descrevem o comportamento tempo-frequência da transformada em ondaletas. Considere a dilatação $\psi(t) \rightarrow \sigma^{-1/2}\psi(t/\sigma)$. Neste caso, a localização no domínio temporal torna-se

$$\left(-\frac{1}{2}\sigma\sigma_t, \frac{1}{2}\sigma\sigma_t\right), \quad (1.10)$$

e

$$\left(\frac{\omega_0}{\sigma} - \frac{1}{2}\frac{\sigma_\omega}{\sigma}, \frac{\omega_0}{\sigma} + \frac{1}{2}\frac{\sigma_\omega}{\sigma}\right). \quad (1.11)$$

é a localização no domínio da frequência. Se a ondaleta é mais localizada (a energia da ondaleta é concentrada em uma pequena região do espaço), ela fornece uma representação “melhor” da função no plano tempo-frequência (ou tempo-escala). É importante notar que a “forma” da ondaleta permanece a mesma sob dilatação (mudança de escala) e translação, como pode ser vista na Figura (1.1). A Figura (1.2) representa para o particionamento (em blocos) do espaço tempos-escalas (frequências) diferentes. Os limites dos blocos não são exatos mas dão uma indicação da extensão de tempo-frequência da função de base considerada. Mantendo o conceito

de base em mente, pode-se afirmar uma analogia entre a análise em ondaletas e a análise de Fourier. Nas Figuras (1.2) estão as representações das análises dos planos tempo-frequência para Fourier-Gabor e ondaletas, respectivamente. Com Fourier, a decomposição de uma função ocorre em base senoidais e as ondaletas decompõem a função em uma combinação linear de versões transladadas da “ondaleta-mãe” em escalas diferentes.

Outras características importantes das ondaletas contínuas (ou discretas) são as propriedades referentes aos momentos nulos e suavidades. Formalmente, no domínio espectral a suavidade corresponde ao decaimento de $\hat{\psi}(\omega)$ para $\omega \rightarrow \pm\infty$ e o número de momentos nulos corresponde ao decaimento de $\hat{\psi}(\omega)$ para $\omega \rightarrow 0$. Para uma ondaleta com decaimento rápido existe para todo $b \in \mathbb{N}$ uma constante C_b tal que [Mallat (1999)]

$$\forall t \in \mathbb{R}, |\psi(t)| \leq \frac{C_b}{1 + |t|^b}. \quad (1.12)$$

Considerando-se a classe de ondaletas com decaimento rápido, pode-se definir os momentos nulos de $\psi(t)$ como;

Definição 1.3 (Número de momentos nulos [Mallat and Hwang (1992)]). *Uma ondaleta possui q momentos nulos, se e somente se para todo inteiro positivo $p < q$, ela satisfaz*

$$\int_{-\infty}^{+\infty} t^p \psi(t) dt = 0. \quad (1.13)$$

Isto significa que ψ é ortogonal a qualquer polinômio de grau $q - 1$.

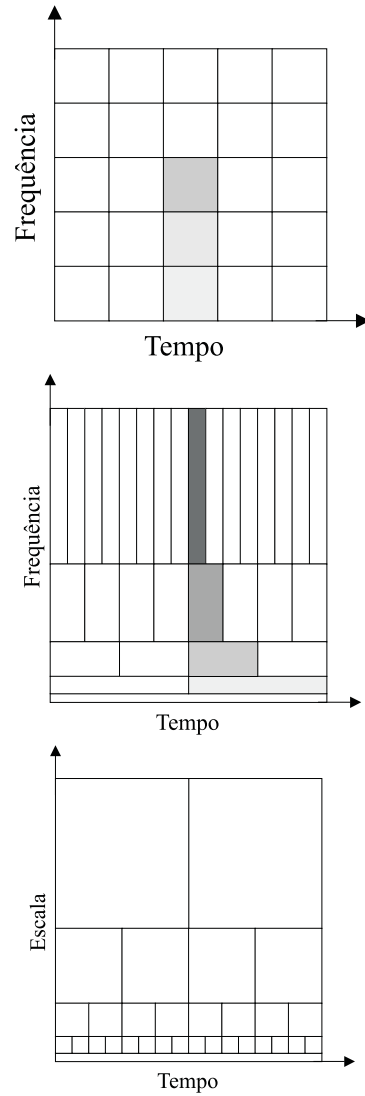


Figura 1.2: Esquema ilustrativo da representação do particionamento do plano $Tempo \times Frequência$ para as transformadas de (topo) Fourier-Gabor, (meio) ondaletas e (baixo) discretização no domínio tempo-escala para as ondaletas.

Como descrito anteriormente e mostrada na equação (1.1), a condição de admissibilidade reflete o fato de que as ondaletas cruzem (zero-crossing) ao menos uma vez a reta real. A quantidade de “cruzamentos” da ondaleta com a reta real pode ser aumentada por aumentar o número de momentos nulos.

1.2.2 Transformada inversa em ondaletas

Ao efetuar a transformação de uma função $f(t)$ para um outro domínio é importante dispor de um processo de recuperação (processo inverso) da função original. A transformada contínua em ondaletas satisfaz a essa propriedade de ser inversível. Portanto, a recuperação da função $f \in L^2(\mathbb{R})$ é possível através dos coeficientes $Wf(\sigma, \tau)$ e é dada pela seguinte forma;

$$f(t) = \frac{1}{C_\psi} \int_0^{+\infty} \frac{d\sigma}{\sigma^2} \int_{-\infty}^{+\infty} Wf(\sigma, \tau) \frac{1}{\sqrt{|\sigma|}} \psi\left(\frac{t-\tau}{\sigma}\right) d\tau. \quad (1.14)$$

A Transformada Inversa é a reconstrução da função original $f \in L^2(\mathbb{R})$ somando as funções da base, que são ponderadas em amplitudes pelos coeficientes $Wf(\sigma, \tau)$ da transformada. Para que a função $f \in L^2(\mathbb{R})$ possa ser reconstruída sem perda de informações, o teorema de Calderón-Grossmann-Morlet [Mallat (1999)], diz ainda que, $\psi(t)$ deve satisfazer a condição de admissibilidade representada pela equação (1.1).

1.3 Transformada discreta em ondaletas

Na seção (1.2) vimos que a transformada contínua em ondaletas utiliza funções dada pela equação (1.3) onde os parâmetros σ e τ controlam a “largura” e a “localização” das *funções de análises* que formam a base. Com os parâmetros contínuos a transformada contínua em ondaletas é uma representação redundante. A diminuição (ou eliminação) da redundância tem um impacto significativo sobre a eficiência dos algoritmos empregados na análise e, uma discretização dos parâmetros τ (tempo) e σ (escala) é suficiente para reduzir a representação redundante a uma base ortonormal. Uma escolha adequada para o parâmetro de escala é $\sigma = \sigma_0^j$ com $j \in \mathbb{Z}$ e $\sigma_0 > 1$ e para o parâmetro de translação é $\tau = k\tau_0$ com $k \in \mathbb{Z}$ e $\tau_0 > 0$. O valor de τ_0 deve ser escolhido de forma que as ondaletas $\psi(t - k\tau_0)$ “cubram” todo o eixo temporal. Deve-se perceber ainda, como está ilustrado na Figura (1.2), que a discretização de τ deve estar relacionada com a discretização de $\sigma = \sigma_0^j$. Portanto, uma escolha conveniente para τ é da forma $\tau = k\sigma_0^j\tau_0$.

Uma classe particular são as ondaletas com valores numéricos dos parâmetros $\sigma_0 = 2$ e $\tau_0 = 1$, tal que, $\sigma \rightarrow 2^j$ e $\tau \rightarrow 2^j k$, com $(j, k) \in \mathbb{Z} \times \mathbb{Z}$. Esta notação conduz a uma estrutura em escalas (índice j) e translações (índice k) chamada diádica, que assemelha-se com a notação musical, em que as potências de 2 estão relacionadas com intervalos (oitavas) e duração das notas. Como será especificado na seção (1.3.3), esse modelo opera com o algoritmo de decomposição diádica que faz uso de banco de filtros, que é um conceito vindo da teoria de processamento de sinais.

Definição 1.4 (Ondaletas discretas) *Uma ondaleta discreta é uma função $\psi(t)$,*

tal que a família de funções

$$\psi_{j,k}(t) = \frac{1}{\sqrt{2^j}} \psi \left(\frac{t - k2^j}{2^j} \right), \quad (1.15)$$

seja uma base ortonormal para $L^2(\mathbb{R})$ com j e k inteiros.

Observação 1.1 Da definição acima, se ψ é uma ondaleta, então, $\psi_{j,k}$ também o será para qualquer $j, k \in \mathbb{Z}$.

Agora são os parâmetros j e k , basicamente, quem controlam, respectivamente, as dilatações e translações das ondaletas. Então, a Transformada Discreta em Ondaletas de uma função $f(t) \in L^2(\mathbb{R})$ terá uma expressão da forma;

$$d_{j,k} = \int_{-\infty}^{+\infty} f(t) \frac{1}{\sqrt{2^j}} \psi \left(\frac{t - k2^j}{2^j} \right) dt, \quad (1.16)$$

onde os coeficientes gerados $d_{j,k}$ são chamados de *coeficientes de detalhes* ou *coeficientes ondaletas*.

1.3.1 Análise em Multiresolução

A estrutura da Análise Multiresolução foi formulada em trabalhos desenvolvidos por Mallat [Mallat (1989a)] onde ele estudou as propriedades de um operador que decompõe uma imagem (ou função) em resoluções diferentes. Mostrou que a diferença de informação entre as aproximações nas resoluções pode ser extraída por decompor o sinal em uma base ortonormal formada por funções ondaletas. Esta decomposição define uma representação multiresolução e é calculada por um algoritmo baseada na convolução com filtros de reconstrução perfeita. Então, nesta

tese, apresento o método da análise multiresolução o qual consiste em representar uma função $f(t) \in L^2(\mathbb{R})$ como um limite de aproximações sucessivas de subespaços $V_j \in L^2(\mathbb{R})$, $j \in \mathbb{Z}$, com resoluções diferentes.

Embora a Análise Multiresolução seja bastante importante, não apenas no tratamento que envolve bases de ondaletas, mas também pelo enorme potencial em termos de desenvolvimento e aplicações, não será descrito nesta tese um tratamento exaustivo. Como sendo um modelo matemático adequado para uma representação por escala de um sistema físico, apresentarei, aqui, alguns conceitos e o formalismo matemático que bastarão para a compreensão do sistema de ondaletas utilizado para cumprir os objetivos deste trabalho.

Definição 1.5 (Análise Multiresolução, [Mallat (1989a)]) *A Análise Multiresolução do $L^2(\mathbb{R})$ é definida como uma sequência de subespaços V_j do $L^2(\mathbb{R})$, com $j \in \mathbb{Z}$, representando os sucessivos níveis de resolução, com as seguintes propriedades;*

1. *Invariância por translação discreta:*

$$\forall k \in \mathbb{Z}, f(t) \in V_0 \Leftrightarrow f(t - k) \in V_0, \quad (1.17)$$

2. *Possibilidade de eliminar detalhes:*

$$\forall j \in \mathbb{Z}, V_{j+1} \subset V_j, \quad (1.18)$$

3. *Cálculo de aproximações menos detalhadas:*

$$\forall j \in \mathbb{Z}, f(t) \in V_0 \Leftrightarrow f\left(\frac{t}{2^j}\right) \in V_j, \quad (1.19)$$

4. *Sabe-se tudo a resolução infinita:*

$$\bigcup_{j \in \mathbb{Z}} V_j = L^2(\mathbb{R}), \quad (1.20)$$

5. *Nada se sabe a resolução nula:*

$$\bigcap_{j \in \mathbb{Z}} V_j = \{0\}, \quad (1.21)$$

6. *Existência de uma base: Existe uma função $\phi \in V_0$ tal que a família $\{\phi_{0,k}(t) = \phi(t - k), k \in \mathbb{Z}\}$, forma uma base ortonormal em V_0 .*

A idéia que está por trás das relações da Análise Multiresolução é escrever uma função $f(t)$ como um limite de aproximações sucessivas em níveis de resolução diferentes. As aproximações podem ser obtidas pela projeção da função original nos subespaços V_j . Cada aproximação pode ser vista como uma versão “suavizada” da função original. As projeções são determinadas através dos elementos da base $\{\phi(t - k)\}_{k \in \mathbb{Z}}$ e serão escritas nos subespaços V_j , os quais serão caracterizados pelas funções ϕ , e nos subespaços W_j , os quais serão caracterizados pelas funções ψ . Os subespaços V_j são os subespaços das escalas e V_0 é o subespaço central. Todos os subespaços V_j são versões escaladas do subespaço central V_0 (propriedade 3). W_j são os subespaços das ondaletas. O fato de que V_j ser espaço de aproximação sucessiva, garante que a aproximação de $f \in L^2(\mathbb{R})$ nesse subespaço (resolução 2^j) é definida como a projeção de f sobre V_j .

Pela propriedade⁷ 4, quando calcula-se uma aproximação de f no subespaço V_j alguma informação é perdida, e essa diferença entre f e sua aproximação pode ser suficientemente pequena para valores de j cada vez menores. Isto é, se P_j é definido como o operador linear projeção ortogonal de f sobre V_j , então $\lim_{j \rightarrow -\infty} P_j f = f$ para todo $f \in L^2(\mathbb{R})$. Por outro lado, a quinta propriedade diz que, para valores

⁷As propriedades enumeradas de 1-6 serão explicadas, de maneira não seqüencial, ao longo deste trabalho.

grandes de j , a projeção de f contém menos e menos detalhes (perde-se informações) e a função converge à zero. As informações adicionais necessárias para passar do subespaço V_j para V_{j-1} (aumentar a resolução de 2^j para 2^{j-1}) são dadas pela projeção de f sobre o complemento ortogonal de V_j em relação a V_{j-1} , denotado por W_j (ou $V_j \oplus W_j = V_{j-1}$).

As aproximações em multiresolução possuem uma implementação bastante simples para a teoria das ondaletas discretas e banco filtros h e g os quais devem possuir a propriedade de reconstrução perfeita.

1.3.2 Transformada discreta em ondaletas

A Função ϕ e o subespaço V_j

Da propriedade 6 surge a necessidade de se definir a função $\phi \in V_0$ (chamada de função de escala) que desempenha um papel importante. As translações inteiras $\{\phi(t - k)\}_k$, com $k \in \mathbb{Z}$, de $\phi(t)$ constituem uma base para o subespaço V_0 . Então, como consequência da segunda e terceira propriedades da multiresolução,

$$\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) = \sum_{k=-\infty}^{+\infty} h[k]\phi(t - k) \quad (1.22)$$

a qual decompõe a função $\frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right) \in V_1$ em uma base de V_0 e a sequência $h[k] = \left\langle \frac{1}{\sqrt{2}}\phi\left(\frac{t}{2}\right), \phi(t - k) \right\rangle$ interpretada como um filtro discreto e satisfazendo as propriedades apresentadas no Apêndice B. Para um valor inteiro j , as funções $\phi_{j,k}(t) = 2^{-j/2}\phi(2^{-j}t - k)$, obtidas através de operações de dilatação e translação,

formam uma base ortonormal para o subespaço V_j . O coeficiente $\sqrt{2^{-j}}$ aparece para normalizar a função no espaço $L^2(\mathbb{R})$. Existe uma única função de escala para cada subespaço de aproximação V_j . A equação acima equivale a relação (B.1) no domínio da frequência que pode ser iterada dando uma relação entre a função escala ϕ e o filtro discreto h .

A Função ψ e o subespaço W_j

Para além da função de escala ϕ definida em V_j , agora há uma função ψ definida em W_j (subespaços das ondaletas), cuja definição é essencial para a geração do sistema completo de ondaletas. Em dois subespaços de escalas consecutivos $V_{j+1} \subset V_j$ propriedade 2 da análise multiresolução, W_{j+1} é o complemento ortogonal de V_{j+1} em V_j . Considerando $j = 0$, uma função $\psi(\frac{t}{2}) \in W_1$ pode ser escrita em termos da base ortonormal de V_0 ;

$$\frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right) = \sum_{k=-\infty}^{+\infty} g[k]\phi(t-k), \quad (1.23)$$

com a sequência $g[k] = \left\langle \frac{1}{\sqrt{2}}\psi\left(\frac{t}{2}\right), \phi(t-k) \right\rangle$ interpretada como um filtro discreto e satisfazendo as propriedades, próprias, apresentadas no Apêndice B. Para que $\{\psi(t-k)\}$ seja uma base ortonormal de ondaletas discretas o filtro discreto g deve satisfazer as condições expostas pelas equações (B.6) e (B.7). Com um valor inteiro j , as funções $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$, obtidas por meio de uma operação de dilatação e uma operação de translação, constituem uma base para o subespaço W_j . O fato de o subespaço W_j ser o complemento ortogonal do subespaço V_{j+1} em V_j ,

conclui-se que $W_{j+1} \perp V_{j+1}$. Deve-se perceber, ainda, que os subspaços W_j são ortogonais entre si: $\dots \perp W_{j+1} \perp W_j \perp W_{j-1} \perp \dots$

1.3.3 Transformada rápida em ondaletas

Pela estrutura da Análise Multiresolução, pode-se definir uma coleção de subspaços W_j tal que $V_j = V_{j+1} \oplus W_{j+1}$. Segue-se que, para $j < J$, pode-se decompor V_j como a soma direta de subspaços $V_j = W_{j+1} \oplus W_{j+2} \oplus \dots \oplus W_J \oplus V_J$. Como consequência, $L^2(\mathbb{R}) = \bigoplus_{j \in \mathbb{Z}} W_j$ representa uma decomposição em ondaletas do espaço $L^2(\mathbb{R})$ em subspaços ortogonais W_j . Portanto, qualquer função de quadrado integrável $f(t) \in L^2(\mathbb{R})$ pode ser decomposta como a soma direta de sua projeção sobre os subspaços das ondaleta mutuamente ortogonais $f(t) = \sum_{j \in \mathbb{Z}} P_{W_j}[f]$. Como $V_{j-1} = V_j \oplus W_j$, segue que para qualquer função $f(t) \in V_j$ pode ser expressa como;

$$P_{V_{j-1}}[f](t) = P_{V_j}[f](t) + P_{W_j}[f](t), \quad (1.24)$$

ou

$$P_{V_{j-1}}[f](t) = \sum_k a_{j,k} \phi_{j,k}(t) + \sum_k d_{j,k} \psi_{j,k}(t), \quad (1.25)$$

onde o $d_{j,k} \equiv \langle f, \psi_{j,k} \rangle$ são obtidos pela projeção de f em W_j ($P_{W_j}[f](t)$) a qual constitui a representação discreta dos detalhes (informação que deve ser adicionada à aproximação em V_j para se obter a aproximação em V_{j-1}) e $a_{j,k} \equiv \langle f, \phi_{j,k} \rangle$ é a

projecção de f em V_j ($P_{V_j}[f](t)$) a qual constitui uma representação aproximada na escala j . Com base nas afirmações e devido à propriedade 4 da análise multiresolução:

$$f(t) = \sum_j \sum_k d_{j,k} \psi_{j,k}, \tag{1.26}$$

que é o equivalente discreto da fórmula de reconstrução de Calderon-Grossmann-Morlet. Na verdade, quando o intervalo torna-se infinito (propriedade 4 da análise multiresolução) a primeira soma da equação (1.25) pode ser omitida e a expansão é puramente em ondaletas, como na equação (1.26).

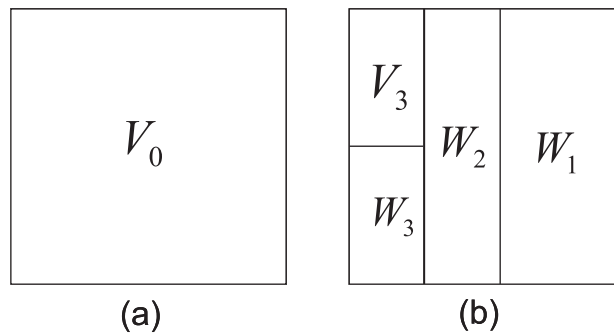


Figura 1.3: Representação da decomposição de f_N nos subespaços V_j e W_j . Em (a) tem-se a representação do subespaço V_0 e em (b) a representação da decomposição na soma direta $V_0 = W_1 \oplus W_2 \oplus W_3 \oplus V_3$.

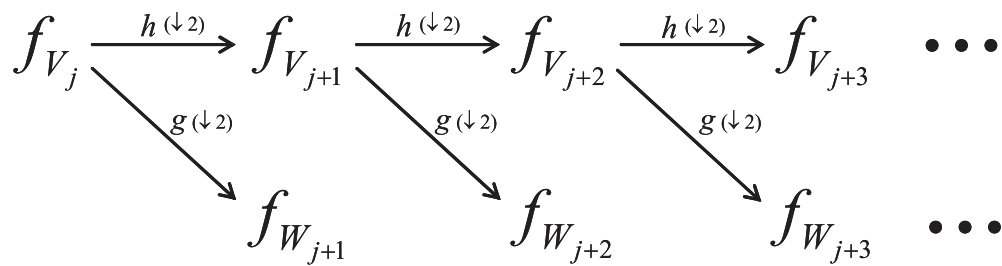


Figura 1.4: Representação do algoritmo de decomposição rápida com os filtros h e g seguida por uma decimação de dois ($\downarrow 2$).

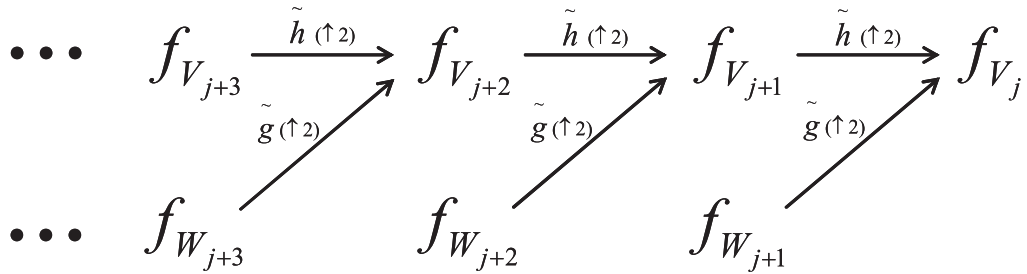


Figura 1.5: Representação do algoritmo de reconstrução rápida com os filtros \tilde{h} e \tilde{g} seguida por uma inserção de dois ($\uparrow 2$).

Como apresentado nesta seção, é possível realizar um produto interno da função $f(t)$ com uma base formada por funções pertencentes ao subespaço V_{j-1} para se obter uma representação de $f(t)$ nos subespaços V_j e W_j . Porém, isso pode ser feito (sem a necessidade de calcular integrais explicitamente) por um método recursivo. Proposto por Mallat [Mallat (1989b)], esse método⁸ é uma forma rápida e computacionalmente eficiente para encontrar os coeficientes $a_{j,k}$ e $d_{j,k}$. Considere uma série temporal f_N formada por N amostras e assumindo que $N = 2^j$ (composta por oito amostras equivale ao índice $j = 3$). O processo começa com a série f_N pertencendo ao subespaço V_j . Isto é, a série $f_N = f_{V_j}$ está associada a uma aproximação contida no subespaço V_j . Essa aproximação é expressa por todos os detalhes $d_{j,k}$ (coeficientes de detalhes ou ondaletas) mais um termo de aproximação (como pode ser vista na Figura (1.4)) na mais alta escala (considerando três níveis);

$$f_N(t) = a_{j+3,k'}\phi_{j+3,k'}(t) + \sum_{j=1}^{j+2} \sum_{k=0}^{k=2^j-1} d_{j,k}\psi_{j,k}(t). \quad (1.27)$$

A decomposição de f_N representada acima dá uma análise em termos de sua projeção sobre os subespaços W_j . Note que, pelas construções apresentadas para os subespaços $W_j \perp W_{j'}$ se $j \neq j'$, $V_j \perp W_j$ e uma vez que esses subespaços estão relacionados, então esta decomposição é única. Uma compreensão lógica é ilustrada na Figura (1.3) onde supõe-se por simplicidade que f_N (formada por oito amostras) está representada em V_0 como no gráfico (a). Pela decomposição de V_0 na soma

⁸Conhecido como decomposição rápida em ondaletas ou ainda algoritmo de decomposição em árvore.

direta dos subespaços $V_0 = W_1 \oplus W_2 \oplus W_3 \oplus V_3$ representada em (b), tem-se uma interpretação dos vários termos da expressão (1.27). O processo começa com $f_{V_0} \in V_0 = V_1 \oplus W_1$ e o primeiro passo é decompor f_{V_0} em $f_{V_1} + f_{W_1}$ onde $f_{V_1} = P_{V_1}[f](t)$ e $f_{W_1} = P_{W_1}[f](t)$. O processo de recursão age sobre $f_{V_1} \in V_1 = V_2 \oplus W_2$, decompondo-o em f_{V_2} e f_{W_2} , onde $f_{V_2} = P_{V_2}[f](t)$ e $f_{W_2} = P_{W_2}[f](t)$. Por fim, a decomposição é realizada sobre $f_{V_2} \in V_2 = V_3 \oplus W_3$, onde $f_{V_3} = P_{V_3}[f](t)$ e $f_{W_3} = P_{W_3}[f](t)$. As componentes $f_{W'}$ são mantidas a parte e no final obtêm-se uma representação em ondaletas para $f(t)$.

Em representação geral, como mostrada na Figura (1.4), o processo de decomposição é repetido, atuando sempre sobre as componentes de aproximações (espaços V_j). Esse processo pode ser pensado como uma transformação de base onde ocorre a mudança $\{\phi_{j,k}\}_{k,j \in \mathbb{Z}} \rightarrow \{(\phi_{j+1,k}, \psi_{j+1,k})\}_{k,j \in \mathbb{Z}}$ e que ambos os conjuntos formam uma base para o subespaço V_j . Ainda na Figura (1.4) $f_{V_{j+1}}$, $f_{V_{j+2}}$ e $f_{V_{j+3}}$ são as componentes de aproximação e $f_{W_{j+1}}$, $f_{W_{j+2}}$ e $f_{W_{j+3}}$ são as componentes de detalhes (domínio das ondaletas) mostradas para três níveis de escala. Este algoritmo corresponde a uma iteração sucessiva dos filtros de reconstrução perfeita h e g (veja Apêndice B). O limite máximo de decomposições corresponde ao nível em que a componente de detalhe consiste de apenas uma amostra. Na decomposição usa-se;

$$a_{j+1,k} = \sum_{k'} h[k' - 2k] a_{j,k'} \quad (1.28)$$

e

$$d_{j+1,k} = \sum_{k'} g[k' - 2k] a_{j,k'}. \quad (1.29)$$

Portanto, os coeficientes $a_{j+1,k}$ e $d_{j+1,k}$ são calculados por convoluções discretas com os filtros h e g , respectivamente, e o fator $2k$ indica uma operação de subamostragem de 2. Por exemplo, se a sequência começa em $a_{0,m}$ com $m = 2^J$ coeficientes, então será decomposto nas sequências $d_{1,m/2}, d_{2,m/4}, \dots, d_{J,m/2^J}$ e $a_{J,m/2^J}$ e a representação em ondaletas terá a mesma quantidade de coeficientes que a representação inicial. A fórmula da reconstrução é descrita como;

$$a_j[k'] = \sum_k h[k' - 2k] a_{j+1}[k] + \sum_k g[k' - 2k] d_{j+1}[k]. \quad (1.30)$$

que corresponde a uma aplicação sucessiva dos filtros de reconstrução perfeita h e g . A Figura (1.5) está ilustrando o algoritmo de reconstrução rápida seguido de uma inserção amostral com os filtros h e g .

Capítulo 2

Transformada de Karhunen-Loève

2.1 Introdução

A Transformada de Karhunen-Loève (ou simplesmente transformada KL) é usada aqui, como um procedimento, que corresponde a conceitos e métodos da Física Matemática, para a análise de processos físicos. Geralmente, um processo físico pode ser representado por uma matriz, como é o caso dos registros sísmicos (Figuras 4.1(a) e 5.1(a)) analisados neste trabalho. A idéia central é a transformação do espaço original para outro com dimensão menor, de maneira que esta representação nova esteja livre de informações redundantes. Com a transformada de Karhunen-Loève é permitido uma representação de um processo físico com um número finito de modos (componentes principais ou valores singulares do processo) utilizando bases formadas pelos autovalores (λ) da matriz correlação (Γ) do processo físico (função, sinal ou imagem) considerada. Em processamento de imagens, a transformada de Karhunen-Loève é conhecida como Transformada de Hostelling e, com um procedimento similar, ela aparece em outras áreas com os nomes: Análise em Componentes Principais (*Principal Component Analysis - PCA*), Decomposição em

Valores Singulares (*Singular Value decomposition* - SVD), Decomposição Ortogonal Própria (*Proper Orthogonal Decomposition* - POD), Funções Ortogonais Empíricas (*Empirical Orthogonal Functions* - EOF). Cada um desses procedimentos (método de aplicação) está especificado e discutido nas referências [Sirovich (1987), Holmes et al. (1996), Jolliffe (2002)] onde é feita uma conexão direta entre esses métodos.

A Transformada KL tem sido usada com muito sucesso em processamento de imagens, oceanografia, meteorologia, [Preisendorfer (1988), Wallace et al. (1992), Thiebaut (1994)]. Ultimamente, a expansão em KL vem sendo utilizada em processamento e caracterização de dados sísmicos [Montagne and Vasconcelos (2006a)-(2006b), López et al. (2005)]. Ela é um método de análise que é descrito essencialmente por matrizes. No escopo deste trabalho, esta metodologia corresponde exatamente na exemplificação do aspecto do dado sísmico (Figura (3.3) ou (4.1(a))) e, portanto, expomos os conceitos da KL na abordagem matricial.

2.2 Definição e propriedades principais

Pode-se pensar em um sistema físico com n medidas (localização espacial) realizadas em cada instante m (localização temporal). Essas medidas podem ser dispostas em uma matriz A de ordem $m \times n$ ($A \in \mathbb{R}^{m \times n}$). A matriz A contém as observações de tal forma que cada elemento a_{mn} , represente uma medida em uma posição para um dado instante. Então, tem-se, para cada coluna de A , um vetor de estado discreto (série temporal de tamanho m). Por exemplo, um registro sísmico (sismograma representado na Figura 3.3) consistindo de n traços com m amostras por traço (com $n < m$ para o caso usual) pode ser representado pela matriz de dados

A . Os elementos $a_{mn} \in A$, agora, correspondem as amplitudes das ondas sísmicas gravadas em certo geofone n em um tempo específico m e, por simplicidade, é assumido que A possui posto $r = n$ (quantidade de autovalores diferentes de zero). Uma matriz simétrica Γ , de ordem $m \times m$, pode ser definida a partir da matriz de dados A , ou seja $\Gamma \equiv AA^t$ ($\Gamma \in \mathbb{R}^{m \times m}$). Aqui é considerada a propriedade da matriz simétrica na qual possui autovalores reais e autovalores ortonormais. A subscrita t no produto AA^t simboliza a matriz transposta de A . Em álgebra linear, tal matriz Γ é definida como matriz de covariância e seus elementos são definidos positivos. Dessa forma, a Transformada de Kahunen-Loève dos elementos da matriz A é definida como sendo a matriz Ψ de ordem $m \times n$ ($\Psi \in \mathbb{R}^{m \times n}$)

$$\Psi = U^t A, \quad (2.1)$$

onde as colunas da matriz U de ordem $m \times m$ ($U \in \mathbb{R}^{m \times m}$) são os autovetores \vec{u}_i da matriz de covariância Γ ;

$$U = (\vec{u}_1 \ \vec{u}_2 \ \dots \ \vec{u}_m), \quad (2.2)$$

correspondendo aos seus autovalores λ_i apresentados na ordem $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_m > 0$. A quantidade $\sum_{i=1}^m \lambda_i$ pode ser analisada como sendo a medida da energia total dos dados, tal que cada λ_i pode ser interpretado como a energia do modo empírico i (modo de energia) representado pelo vetor \vec{u}_i [Aubry et al. (1991)]. Os

elementos originais da matriz A podem ser recuperados pela definição relação inversa da transformada;

$$A = U\Psi. \quad (2.3)$$

A equação acima é uma expansão KL da matriz de dados A . Nesse sentido, uma expansão em KL é “adaptativa”, significando que os dados originais são decompostos em uma base obtida diretamente dos próprios dados empírica (matriz de dados A). Outra propriedade interessante é que, usando somente k^1 autovalores e autovetores pode-se obter a melhor aproximação para a matriz real A . Então, se a matriz Ψ_k for formada por k ($k < m$) linhas de Ψ e as linhas restantes $m - k$ zeradas, então a matriz A_k dada por;

$$A_k = U\Psi_k \quad (2.4)$$

é a melhor aproximação de A por uma matriz de posto k , com $k < r$. Esta propriedade da expansão KL é dita ser “ótima” e conduz à aplicação da compressão de dados, bem como, também, redução de dimensão, por permitir a aproximação da matriz de dados A por uma matriz menor A_k com perda mínima de informações [Holmes et al. (1996)]. Por outro lado, se a matriz de dados considerada contém ruído coerente localizado no espaço e tempo, como é o caso do ruído *ground roll*

¹Uso aqui esse parâmetro, porém não deve ser confundido com o parâmetro k (mesma letra) que caracteriza uma translação para as ondaletas.

presente nos registros sísmicos (veja Figura (3.3)), então a transformada KL pode ser usada para a remoção de tal ruído. Portanto, se a matriz Ψ'_k for formada por zerar as primeiras k linhas da transformada Ψ e as restantes não nulas, então a matriz A'_k dada por

$$A'_k = U\Psi'_k \quad (2.5)$$

será uma versão filtrada da matriz A onde os modos mais coerentes k (os maiores k autovalores da matriz de covariância Γ) foram removidos. Portanto, para uma remoção eficiente é melhor aplicar a KL somente na região do registro sísmico contaminada pelo *ground roll*.

Como já mencionado anteriormente, a transformada de Karhunen-Loève tem sido aplicada de formas diferentes dependendo do contexto onde se queira usá-la. Portanto, é conveniente apresentar aqui uma relação (método computacional) entre a transformada KL e um procedimento algébrico. O método escolhido foi o da Decomposição em Valores Singulares (SVD). A Transformada KL apresenta uma relação direta com a decomposição de valores singulares de matrizes e, esse modelo discute uma analogia de compreensão fácil, como também a sua implementação computacional.

2.3 Transformada de Karhunen-Loève e Decomposição em Valores Singulares

Esta seção envolve o procedimento algébrico para o compto da transformada KL envolvendo a decomposição em valores singulares. Para uma revisão detalhada sobre este método, consulte a referência [Jolliffe (2002)]. A decomposição de valores singulares (SVD) é uma técnica importante em Álgebra Linear, tanto computacional quanto teórica, para análise de matrizes. Aqui, faz-se restrição para matrizes de valores reais. Uma característica importante da SVD é que a decomposição em valores singulares da matriz A é sempre bem condicionada e existem algoritmos estáveis para tal solução deste problema. Portanto, uma decomposição em valores singulares (SVD) da matriz $A \in \mathbb{R}^{m \times n}$ é uma fatorização escrita da forma;

$$A = U\Sigma V^t. \quad (2.6)$$

onde U é uma matriz ortogonal de ordem $m \times m$ como definida na equação (2.2). As colunas de U são chamadas de *vetores singulares à esquerda* de A e formam uma base ortonormal tal que $\vec{u}_i \vec{u}_j = 1$ para $i = j$ e $\vec{u}_i \vec{u}_j = 0$ para $i \neq j$. A matriz $\Sigma \in \mathbb{R}^{m \times n}$ é uma matriz real e diagonal com elementos $\sigma_i = \sqrt{\lambda_i}$ para $i = 1, 2, \dots, r$ sendo r o posto da matriz Σ . Os elementos σ_i são valores não-nulos de Σ e estão sobre a diagonal principal $\Sigma = \sigma_{ij} \delta_{ij}$, onde δ_{ij} é o delta de Kronecker (valor 1 se $i = j$ e 0 se $i \neq j$). Os σ_i são chamados de *valores singulares* da matriz A .

Por convenção, a ordem para os valores singulares é determinada do maior

para o menor, com $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n > 0$ onde o maior valor está no topo esquerdo da matriz Σ . As colunas da matriz $V \in \mathbb{R}^{n \times n}$ correspondem aos n autovetores \vec{v}_i da matriz $A^t A$ (chamados de *vetores singulares à direita* de A) com autovalores σ_i não nulos. Note que a decomposição em valores singulares é equivalente a uma diagonalização (solução) de um problema de autovalores.

2.3.1 Decomposição em Valores Singulares e Aproximação de matrizes

A decomposição em valores singulares também pode ser usada para encontrar a melhor aproximação A_k (de posto $k < r$) para a matriz $A \in \mathbb{R}^{m \times n}$. A expressão (2.6) pode ser escrita da seguinte forma;

$$A_k = U\Sigma V^t = [u_1 \dots u_k] \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_k \end{bmatrix} \begin{bmatrix} v_1^t \\ \vdots \\ v_k^t \end{bmatrix}$$

ou ainda da forma;

$$A_k = \sum_{i=1}^k \sigma_i \vec{u}_i \vec{v}_i^t = \sum_{i=1}^k \sigma_i Q_i. \quad (2.7)$$

Os autovetores \vec{u}_i são chamados de *autovetores empíricos*, *modos ortogonais próprios* ou simplesmente de *modos KL*. No contexto de processamento de imagens as matrizes Q_i são as auto-imagens da matriz A [Andrews et al. (1977)]. Esta

aproximação é útil em problemas de identificação de sistemas, ou identificação de sinais com ruídos.

Agora, comparando as equações (2.3) e (2.6) temos que a transformada de Karhunen-Loève Ψ está relacionada com as matrizes SVD Σ e V pela seguinte relação;

$$\Psi = \Sigma V^t, \quad (2.8)$$

tal que uma expansão de auto-imagens (como em (2.7)) é, precisamente, uma expansão em KL dada pela equação (2.3). Então, a decomposição em valores singulares (SVD) é uma forma computacional eficiente (e equivalente) de calcular a transformada Karhunen-Loève, e este método é usado como um procedimento numérico para o desenvolvimento deste trabalho.

Capítulo 3

Sondagem Sísmica

3.1 Introdução

A sondagem sísmica é um método de exploração do subsolo terrestre, onde se faz uso de aparelhos e técnicas especiais, visando o seu reconhecimento para uma caracterização. Este método tem sido, comumente, utilizado pelo fato de ser capaz de “cobrir” áreas enormes e economicamente viável, caracterizando-se, assim, por uma observação cautelosa do meio terrestre. Nesse sentido, é empregada na localização de jazidas de petróleo e na detecção de falhas geológicas. A sondagem é uma atividade rotineira na indústria do petróleo e tem por objetivo a obtenção de uma representação da geologia de subsuperfície. Para o cumprimento do processo, os dados são submetidos a uma sequência convencional de técnicas que compreende a preparação e tratamento através de técnicas relativamente simples devido ao desenvolvimento de computadores de grande porte. A sondagem pode ser conduzida por métodos diretos, como a perfuração de um poço, que tem um custo bastante elevado ou aplicar métodos indiretos (como a reflexão sísmica), que tem um custo bastante moderado e fornecem informações detalhadas do subsolo, embora sujeitas

a interpretações e visualizações [Yilmaz (2003); Thomas (2001)].

3.2 Prospecção Sísmica

Na prospecção de petróleo, há uma preocupação em obtê-lo de forma economicamente viável. Por isso, antes de perfurar um poço, deve-se garantir uma probabilidade alta de acúmulo de petróleo na região a ser explorada. No entanto, faz-se necessário um estudo elaborado das propriedades dessa área, porque as estruturas geológicas associadas com um reservatório de petróleo são muito complexas e, uma alternativa que não requer a perfuração de um poço, que é uma operação bastante onerosa, consiste na sondagem das camadas do subsolo (representada na Figura (3.1)).

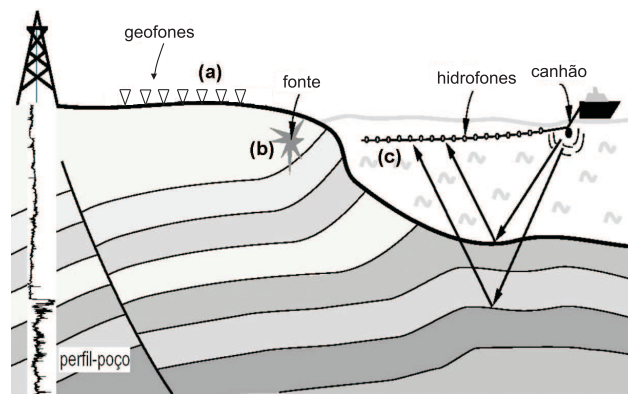


Figura 3.1: Representação da aquisição para a obtenção de dados terrestres e marítimos. Esse processo envolve (caso terrestre) a geração de uma onda sísmica e a sua captação na superfície por um arranjo de receptores (geofones (a)). A fonte de energia mais utilizada na geração das ondas são as fontes explosivas e vibratória.

Esse processo inicia-se com a geração de ondas sísmicas as quais são geradas artificialmente por explosões na superfície ou em pequenas profundidades (representada por (b) na Figura (3.1)). Essas ondas se propagam em todas as direções. Portanto, propagam-se para o interior da Terra. Devido a heterogeneidades do meio, essas ondas sofrem os fenômenos de refração e reflexão, retornando-as à superfície onde são captadas por um arranjo de receptores previamente distribuídos uniformemente (representado por (a) na Figura (3.1)).

A sísmica empregada na exploração de hidrocarbonetos, em geral, está compreendida em três etapas [Yilmaz (2003)]: (i) aquisição dos dados sísmicos que compreende ao levantamento (registro das medidas - dados) dos dados, (ii) processamento sísmico com vista na obtenção das imagens e (iii) interpretação sísmica que compreende a interpretação destas imagens para efetivamente localizar feições de interesse. A aquisição sísmica diz respeito à obtenção dos dados durante um levantamento sísmico, que pode ser terrestre ou marítimo, como está ilustrado na Figura (3.1). Esse processo envolve a geração da onda sísmica (Figura 3.1(b)) e a sua captação em superfície por uma formação previamente ordenada de receptores (geofones para levantamentos terrestres (a) e hidrofones para marítimos (c)). Na modelagem inicial, é preciso definir a geometria de aquisição onde se deve estabelecer a quantidade de receptores, distância entre a fonte e o primeiro receptor bem como a distância entre os demais. Os receptores fazem uma leitura simultânea que transformam as vibrações do solo em sinais elétricos. Em intervalos de tempo dt os receptores verificam se existem sinais retornando à superfície. Cada receptor registra um traço sísmico. Cada traço é uma função do tempo de ida e volta da onda, e

o conjunto de traços gravados (pelo arranjo de geofones) forma um sismograma.

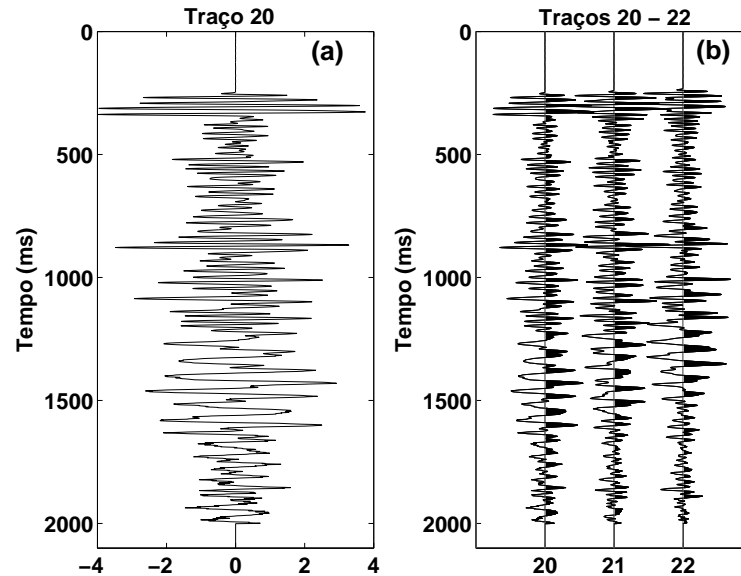


Figura 3.2: Representação para os traços sísmicos. (a) traço sísmico em seu formato convencional (série temporal) onde na horizontal tem-se as amplitudes e na vertical a quantidade de amostras (temporal). Esse traço é o de número 20 do sismograma representado na Figura (3.3). Em (b) tem-se a representação para os traços 20 a 22 (para o mesmo sismograma) em seu formato *wiggle* o qual é muito utilizado para uma visualização adequada.

O sismograma pode ser definido como uma matriz de dados em que os receptores representam as colunas e as linhas são os instantes de tempo, e em seu conteúdo, tem-se armazenamento das amplitudes das ondas capturadas. Na Figura (3.2)(a) tem-se a representação de um traço sísmico como sendo uma série temporal (em seu formato convencional). Já na Figura (3.2)(b) tem-se uma sequência de 3 traços (traços 20 – 22 do sismograma da Figura (3.1)) em seu formato SEG-Y. Essa

representação gráfica é muito utilizada pelas indústrias de petróleo e é chamada de *wiggle*. Nela a série temporal das amostras (traço) é interpolada por uma função de reconstrução (linear ou spline cúbica) que tem suas amplitudes positivas preenchidas (áreas preenchidas). A Figura (3.3) está representando o aspecto (família de traços originados de um único ponto de tiro - Figura 3.3(a)) de um sismograma em geometria $2D$. Esse termo se refere ao arranjo formado por receptores e fonte em uma mesma linha reta. O registro sísmico da Figura (3.3) foi obtido em terra e corresponde ao registro 25 como está descrito na referência [Yilmaz (2003)] e disponível em [Stockwell (2000)].

O formato SEG-Y é um dos diversos padrões de gravação de dados sísmicos desenvolvidos pela *Society of Exploration Geophysicists* (SEG). O pacote computacional utilizado para a leitura e visualização dos sismogramas, aqui analisados, foi o *Seismic Unix* (SU) desenvolvido pelo SEG do Center for Wave Phenomena (CWP) da Colorado School of Mines. Além de ser um programa de distribuição livre [Stockwell (2000)], pode ser utilizado nas plataformas Windows/Unix. Possui grande versatilidade no sentido de os pacotes já serem prontos e na possibilidade de se criar rotinas de implementação para o processamento. Outro aspecto que é bastante atrativo é o fato de o mesmo poder receber contribuições de pesquisadores na forma de pacotes. O método de decomposição em multiescala, desenvolvido por nós, foi implementado ao pacote *Seismic Unix* (mas não disponível ainda) com o intuito do processamento dos registros sísmicos.

Os dados registrados em um levantamento sísmico contem uma quantidade alta de trajetórias possíveis que correspondem a sequências de eventos. Então, o

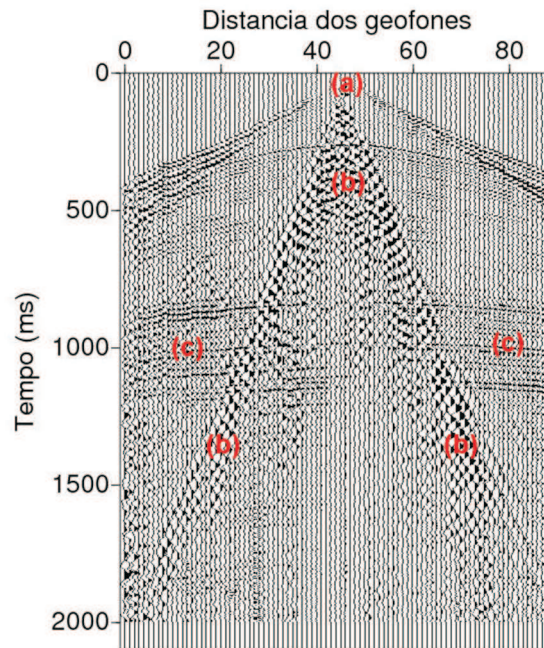


Figura 3.3: Registro sísmico 2D terrestre em representação do tipo *wiggle* com ganho automático [Stockwell (2000)]. O eixo vertical está definido em unidades de tempo (ms). Na horizontal representa o afastamento entre os geofones (m) e cada geofone representa um traço. Em (a) tem-se o local onde ocorreu a geração das ondas sísmicas (fonte). Pode-se observar que a identificação das hipérboles de reflexões (c) ficam prejudicadas pelo “cone” de ruído *ground roll*.

processamento sísmico compreende um conjunto de técnicas que atuam sobre os dados de forma a retirar ou minimizar eventos indesejáveis de forma que as imagens visualizadas sejam uma “fidelidade” máxima possível das camadas geológicas. A visualização é realizada adequando-se as trajetórias possíveis das características

de atributos sísmicos aos seus posicionamentos com o objetivo último de facilitar a interpretação. A interpretação sísmica consiste na representação de um modelo geológico (ou mapas estruturais) compatível com as feições geológicas dos registros sísmicos (sismogramas). É um problema difícil, uma vez que, por natureza, os dados obtidos são não lineares e uma concentração enorme de informações num espaço bem definido. Isto é, diversas configurações geológicas podem ser observadas simultaneamente nos dados e, nesse sentido, há uma dependência enorme do observador (observador humano) em associar estruturas geológicas de interesse com estruturas espúrias. O objetivo principal é determinar a posição e a altitude dos refletores e inferir um modelo para as estruturas geológicas.

3.3 Sísmica de Reflexão

Com a sísmica de reflexão é permitida a visualização da distribuição espacial das camadas do subsolo através do espalhamento (nas interfaces do meio geológico) das ondas sísmicas geradas por fontes controladas. O princípio físico central da sísmica de reflexão está no particionamento da energia, das ondas sísmicas, que acontece nas interfaces geológicas e o fenômeno de reflexão/refração acontece devido ao contraste na densidade dessas interfaces. Como o meio é heterogêneo, então encontrar equações que descrevam a propagação das ondas sísmicas não é fácil. Porém, podem-se admitir condições simplificadoras e favoráveis à propagação. Uma simplificação consiste em considerar que as camadas heterogêneas possam ser modeladas por camadas paralelas planas e de espessuras convenientes de tal forma que a densidade e as propriedades elásticas (de cada camada) sejam realistas às

condições naturais. Das condições acima e pelos conceitos da Óptica Geométrica pode-se provar [Yilmaz (2003); Thomas (2001)] que as reflexões, em uma superfície plana, aparecem nos sismogramas como uma aproximação de uma hipérbole. Essas reflexões estão representadas na Figura (3.2)(c).

Agrupadas as ondas sísmicas geradas pela fonte (explosiva) estão presentes ondas de corpo do tipo P e S (cisalhamento) e as ondas de superfície (ondas do tipo Rayleigh). As ondas que resultam nas hipérbolas são as do tipo P [Gerhardt (1998)], ou ondas compressoriais. O deslocamento deste tipo de onda se dá na mesma direção de propagação da energia da fonte que a emitiu. As ondas do tipo P são velozes (em um sismo são as primeiras a serem detectadas e por isso chamadas, também, de ondas primárias) e se adéquam às principais técnicas da sísmica de reflexão para a construção de imagens da subsuperfície para uma prospecção em hidrocarbonetos.

O método da sísmica de reflexão é o que fornece imagens com maior precisão da subsuperfície e vêm sendo utilizado intensivamente pela indústria do petróleo [Thomas (2001)], graças ao seu poder de resolução. Ele se baseia no estudo de ondas sísmicas, as quais são geradas artificialmente na superfície ou próximo desta. Essas ondas são propagadas para o interior da Terra, chegando a uma profundidade de investigação que depende da energia utilizada na fonte (emissor). Nas discontinuidades do meio geológico, as ondas refletidas que conseguem retornar à superfície são captadas por uma formação de geofones na superfície. Com uma interpretação e visualização, pode-se fazer estimativas dos parâmetros geológicos e prever os locais mais favoráveis para a ocorrência de acúmulos de petróleo.

A utilização deste método se desenvolveu realmente a partir do final da

década de 70, com o desenvolvimento da eletrônica digital e da micro informática. Depois foi possível a construção dos pequenos sismógrafos digitais e sistemas para processamento de sinais sísmicos em micro computadores.

3.3.1 O Ruído de Rolamento Superficial

O sismograma, em sua formação traço-a-traço, apresenta certa variedade de estruturas regulares classificadas como coerentes. E é exatamente nessa quantidade variada de estruturas que se pode revelar informações a respeito da geologia da área explorada. Na representação do sismograma, na Figura (3.3), pode ser observado, visualmente, algumas estruturas. Dentre essas, e bem visível na Figura (3.3)(b), está o ruído de rolamento superficial (*ground roll*). Este sinal está sempre presente nos levantamentos terrestres e representa o principal tipo de ruído coerente. Nos sismogramas, ele aparece, geometricamente, na forma de um “cone” devido a formação dos geofones com o tiro (ponto de energia onde é gerada as ondas sísmicas).

As principais características desse ruído estão presentes nas amplitudes fortes (mais fortes do que o sinal refletido), baixa velocidade e concentração de energia nas frequências baixas [Yilmaz (2003)]. São ondas do tipo Rayleigh que se limitam a uma propagação próxima a superfície da terra sem transmitir energia para o interior da estrutura terrestre. Então, as amplitudes que constituem o ruído coerente *ground roll*, não contêm informações relacionadas com as interfaces pois eles provêm de modos localizados próximos à superfície da Terra. Isso significa que o *ground roll* não carrega informações sobre as estruturas geológicas mais profundas. Então, o *ground roll* torna-se um sinal indesejável pois vai se superpor com as ondas

de amplitudes menores que foram espalhadas (refletidas) nas camadas que pretende-se sondar, mascarando assim, as informações relevantes para a interpretação.

3.3.2 As reflexões

As amplitudes das ondas sísmicas decrescem com a distância à fonte que as emitiu. Porém, a frente de onda torna-se mais plana quanto mais distante da fonte. Com essa aproximação de onda plana é permitida a utilização de um sistema de coordenadas ortogonal para descrever o movimento da frente de onda. Isso constitui uma simplificação conveniente e a geometria definida na modelagem inicial para a aquisição sísmica é importante. Essa importância é bem vista quando da necessidade de se definir sistemas de coordenadas que correspondam na organização dos traços no sismograma. Uma aquisição em coordenadas $(x, z = 0)$ correspondem a fontes e receptores próximos a superfície e a direção ortogonal ao t que mede o tempo de ida e volta da onda. Em uma situação de aquisição ideal, onde um refletor plano se encontra à profundidade z , as reflexões aparecerão no sismograma como uma hipérbole com equação da forma:

$$t^2 = t_o^2 + \frac{x^2}{v^2} = \left(\frac{2z}{v} \right)^2, \quad (3.1)$$

onde t corresponde ao tempo para qualquer posicionamento x , t_o o tempo para fonte e receptor localizados na mesma posição (incidência normal ou afastamento zero) e v a velocidade da onda no meio, supostamente constante. Na prática, a velocidade não é constante e a reflexão será uma aproximação de uma hipérbole.

3.4 Técnicas de Filtragem

Convencionalmente, para a remoção de ruídos são usados filtros, tais como; filtros passa-alta, passa-banda, balanço espectral, multicanal, $f - k$. Os métodos desenvolvidos, baseados nas técnicas acima, são utilizados para a remoção do *ground roll* [Deighan and Watts (1997)]. Porém, esses métodos tem suas limitações pois, só podem ser aplicados ou no domínio da frequência ou no domínio do tempo. Os filtros passa-alta e passa-banda são aplicados no domínio da frequência e são incapazes de separar as amplitudes fortes do *ground roll* e das reflexões. Isso é verdade porque para a mesma banda de frequência as amplitudes do ruído e reflexões se superpõem fortemente [Yilmaz (2003)]. Dessa forma, com essa filtragem, frequências baixas presentes nas reflexões são perdidas. Similarmente, os filtros baseados na transformada $f - k$ são bastante utilizados [Hatton et al. (1986)]. Porém sua desvantagem está na geração de falseamento (distorção das reflexões) criando eventos espúrios quando as amplitudes do *ground roll* são bem maiores que as reflexões.

Algoritmos baseados na Transformada de Karhunen-Loève também tem sido desenvolvidos e implementados como uma técnica para a remoção do *ground roll*. Aqui, cito os trabalhos de Lopez [López et al. (2005)], Montagne [Montagne and Vasconcelos (2006a,b)]. No primeiro trabalho citado, a supressão é feita baseada na afirmação de que as amplitudes do *ground roll* são maiores que as reflexões. Então, faz uso da transformada KL para suprimir as amplitudes maiores sem causar distorções nas reflexões. Porém, o autor comenta, unicamente, que os resultados obtidos são melhores que os resultados obtidos baseados nas técnicas convencionais;

$f - k$, passa-alto, passa-banda. Em seus trabalhos, Montagne e Vasconcelos apresentam um método de supressão baseado na transformada Karhunen-Loève. Foi definido um critério de extremo (máximo da entropia ou mínimo da energia) para identificar a região do registro sísmico a ser filtrada. Então a região é decomposta em *modos de energia KL* pela transformada KL. O *ground roll* fica caracterizado por esses modos. Dessa forma a atenuação é realizada em remover apenas alguns modos KL.

A necessidade da diminuição do custo computacional, da obtenção de resultados mais claros, fazem com que novas ferramentas de filtragem sejam estudadas. Nestes aspectos, a utilização da Transformada em Ondaletas tem sido largamente ampliada. A propriedade de localização em tempo-escala também influi na escolha desta ferramenta para a atenuação do ruído *ground roll*. Para se obter uma análise em tempo-escala, o procedimento é decompor o registro sísmico em uma base formada por ondaletas (essas funções possuem extensão finita no tempo e na frequência) fornecendo uma análise em tempo-frequência limitada pelo Princípio da Incerteza de Heisenberg. A idéia básica da filtragem utilizando as ondaletas é zerar os coeficientes (da transformada) associados ao ruído e preservar os coeficientes (da transformada) que contém as informações desejadas. Os coeficientes da transformada são obtidos ao aplicar algum algoritmo rápido ao registro. Invertendo o algoritmo (Transformada Inversa em Ondaletas) recupera o sinal, porém, com alguns coeficientes atenuados.

Capítulo 4

Supressão do *Ground roll*: um filtro ótimo

4.1 Introdução

A exploração sísmica para prospecção em petróleo realizada por reflexão e refração ocorre devido à propagação de ondas mecânicas no meio heterogêneo do subsolo. As ondas mecânicas são produzidas próximas à superfície de uma área que se deseja explorar (analisar). Devido à diferença de densidade do meio, essas ondas sofrem o fenômeno de reflexão e refração nas interfaces. As ondas que retornam à superfície são registradas por sensores os quais podem ser rearranjados formando uma estrutura padrão: o registro sísmico. Apesar desses registros sísmicos, em princípio, revelarem as feições das estruturas geológicas, eles são, também, caracterizados por sinais ruidosos. Este fato sugere o desenvolvimento de métodos baseados na Física Estatística com objetivo de atenuar o *ground roll* que aparece como a componente principal ruidosa. O sinal *ground roll* é uma denominação comum em geofísica para esse ruído indesejável e que é gerado por ondas dispersivas do tipo Rayleigh. Esses sinais aparecem no registro sísmico mascaram as informações de interesse. Isto é,

atrapalham na leitura/visualização das reflexões que são os sinais de interesse físico. Então, a construção de filtros (para atenuar o ruído *ground roll* eficientemente) é um desafio para os cientistas.

Comumente, os filtros usados agem na eliminação de frequências específicas e a Análise de Fourier é indispensável para a compreensão da maioria dos métodos lineares utilizados. Neste caso, as funções de análise (conjunto de funções que formam a base) são senos e cossenos, as quais têm suporte infinito no tempo. Portanto, a eliminação de uma faixa de frequência afeta o registro por inteiro em todo o tempo. Este conjunto de funções é bem adequado quando o ruído não contém características localizadas no tempo. Devido a essas desvantagens, um procedimento muito utilizado [Deighan and Watts (1997), Miao and Moon (1999), Corso et al. (2003)] é decompor o registro sísmico em uma base formada por ondaletas para suprimir o *ground roll*. A aplicação do filtro é baseada na Transformada em Ondaletas a qual leva em conta a propriedade de localização em tempo-escala. Essa análise é conveniente para sinais não-estacionários. A não-estacionariedade é uma característica dos sismogramas.

Neste capítulo nós apresentamos um método híbrido combinando a Transformada Contínua em Ondaletas e a Transformada de Karhunen-Loève (KL) para atenuar eficientemente o ruído coerente *ground roll* que aparece nos sismogramas. A transformada contínua em ondaletas foi usada na construção do filtro. O sismograma original representado na Figura (4.1)(a), contendo o ruído coerente *ground roll*, foi decomposto (algoritmo para a transformada em ondaletas) na base formada por ondaletas gerando um conjunto de coeficientes $Wf(\sigma, \tau)$ duplamente indexados (escala e tempo). O processo de filtragem consiste em uma atenuação adequada dos

coeficientes (por um fator de atenuação A_f) na escala fortemente “contaminada” pelo ruído *ground roll*. Os coeficientes nas outras escalas, os quais não contêm o *ground roll*, são mantidos. Em seguida o sinal sísmico é reconstruído (algoritmo inverso para a transformada em ondaletas) levando em conta os coeficientes mantidos e os atenuados. A ondaleta escolhida para análise foi a derivada de quinta ordem da gaussiana (Apêndice C). Esta escolha foi baseada em dois critérios: um é que esta derivada da gaussiana satisfaz todas as exigências para a definição de uma base em ondaletas e (outro) é sua forma similar às ondas sísmicas com as ondaletas gaussianas. Existe uma classe enorme de bases para o formalismo das ondaletas. O critério de escolha é variado e não há uma regra fixa [Daubechies (1992)].

A transformada Karhunen-Loève (KL) foi usada no sentido de encontrar o fator de atenuação A_f ótimo através dos espectros de energia KL. Depois do procedimento de filtragem, a transformada KL foi aplicada à região atenuada gerando o espectro de energia KL. O comportamento dos primeiros modos de energia E foi analisado em função do fator A_f . O ponto correspondendo a um mínimo do primeiro modo foi identificado como sendo a máxima extração do ruído coerente *ground roll*. Então, o fator de atenuação ótimo A_f é escolhido para tal valor mínimo.

4.1.1 Ondaletas: passagem do contínuo ao discreto

A transformada contínua em ondaletas tem uma representação redundante. Para diminuir a redundância e aumentar a análise em tempo e frequência as decomposições contínuas são formuladas em pontos discretos baseados nos números de vozes. Grossmann et al. (1989) sugeriram escolhas para os parâmetros σ e τ da

seguinte forma: escolhendo-se $\tau = k2^{j/\nu}$ com $\sigma_0 = 1$ e $\sigma = a^j$ com $a = 2^{1/\nu}$, tem-se uma divisão na escala pelo número de vozes ν . Uma mudança na escala de 2^j para 2^{j+1} é caracterizado por uma oitava. Então, essa escolha significa uma divisão da oitava em ν vozes. Com esta consideração a equação (1.3) assume a forma;

$$\psi_{j,k}^\nu(t) = 2^{-j/2\nu} \psi(2^{-j/\nu}t - k) \quad (4.1)$$

com j e k inteiros. Em situações reais, o sinal sísmico é representado em sua forma discreta $f[n]$, $n = 0, 1, 2, \dots, N-1$, onde N é o número de amostras. A função $f[n]$ e a ondaleta $\psi_j^\nu[n]$ são consideradas como funções de período N e a ondaleta $\psi_j^\nu[n] = \frac{1}{\sqrt{2^{j/\nu}}} \psi\left(\frac{n}{2^{j/\nu}}\right)$ assumindo valores diferentes de zero no intervalo $[-N/2, N/2]$. Portanto, a transformada em ondaletas assumirá a expressão;

$$Wf[n, 2^j] = \sum_{m=0}^{N-1} f[m] \psi_j^\nu[m-n] \quad (4.2)$$

Essa decomposição em multivozes diminui o passo (oitavas) para cada escala, dando uma representação mais contínua. Aqui, é utilizado $\nu = 4$ para a ondaleta derivada quinta da gaussiana. Deve-se perceber que é possível a reconstrução do sinal (transformada inversa em ondaletas) baseada na discretização redundante de Grossmann que equivale a discretização da equação (1.14).

4.1.2 Transformada de Karhunen-Loève

Como se viu na seção (2.2), o sismograma pode ser considerado como uma matriz $M \in \mathbb{R}^{m \times n}$ de valores reais formada por n traços (cada traço um geofone) e m^1 amostras por traço. Dessa forma, o dado sísmico aqui analisado (Figura 4.1(a)) corresponde a uma matriz M consistindo de $n = 96$ traços (posições na horizontal) com $m = 1001$ amostras por traço (tempo na vertical) com $n < m$. Sendo assim, os elementos M_{ij} correspondem à amostra i no traço j^2 . A idéia central da transformada KL é utilizar as informações contidas na matriz de correlação Γ dos dados usando seus autovetores \vec{u} e autovalores λ . Com esse procedimento o dado sísmico ficará representado por seus *autovetores empíricos* (\vec{u}) e em um número finito de *modos KL* (modos de energia KL λ). A utilização da transformada KL é adaptativa no sentido de o sismograma original ser decomposto por uma base que é obtida diretamente do dado empírico.

4.2 O filtro ótimo em ondaletas

Nesta seção expomos o procedimento para o método ótimo híbrido onde fazemos uso das transformadas em ondaletas e KL. Rigorosamente falando, o método consiste em três etapas principais: (1) a seleção da região, (2) o filtro em ondaletas e (3) a escolha ótima para o fator de atenuação. A seleção da região pode ser feita por inspeção visual, pois, o *ground roll* está bem localizado. Pode-se usar

¹Este índice m (índice que caracteriza a dimensão da matriz) não deve ser confundido com o índice m do somatório da equação (4.2).

²Este índice, representando uma posição na matriz M , não deve ser confundido com o índice de escala para as ondaletas.

também, alguma ferramenta de reconhecimento de padrão, tal como o método de otimização desenvolvido por Montagne e Vasconcelos [Montagne and Vasconcelos (2006a,b)] onde é usado um critério de extremo. Em seguida, o filtro baseado em ondaletas foi aplicado à região selecionada. Os coeficientes que representam esta região são reduzidos (em valor) pelo fator de atenuação A_f . O valor para A_f deve ser escolhido de acordo com a terceira parte do método. Para um fator de atenuação nulo $A_f = 0$, não há filtragem e no limite oposto $A_f = 1$ a região é completamente removida. Então, existe algum fator A_f entre esses dois limites que deve ser o *fator ótimo* que caracteriza a melhor atenuação.

1. *Seleção da região.* Como já mencionado anteriormente, a assinatura do *ground roll* no dado sísmico segue um padrão bem localizado no espaço e no tempo. Mais precisamente (Figura 4.1(a)), ele aparece como uma forma geométrica descendente trapezoidal e não está afetando as demais regiões. Portanto, é natural aplicar o filtro somente nesta região do sismograma. A região foi selecionada de acordo com o critério de extremo utilizando a decomposição (transformada KL) em modos de energia KL (autovalores λ). Com esta decomposição, a maior parte da energia fica presente nos primeiros modos os quais caracterizam o *ground roll*. Eliminando os primeiros modos mais energéticos, a energia do sinal sísmico deve diminuir. Isto sugere um critério de energia mínima para selecionar a região ótima que será atenuada.
2. *O filtro baseado nas ondaletas.* A transformada em ondaletas foi aplicada ao registro sísmico, traço a traço, gerando um conjunto de *coeficientes ondale-*

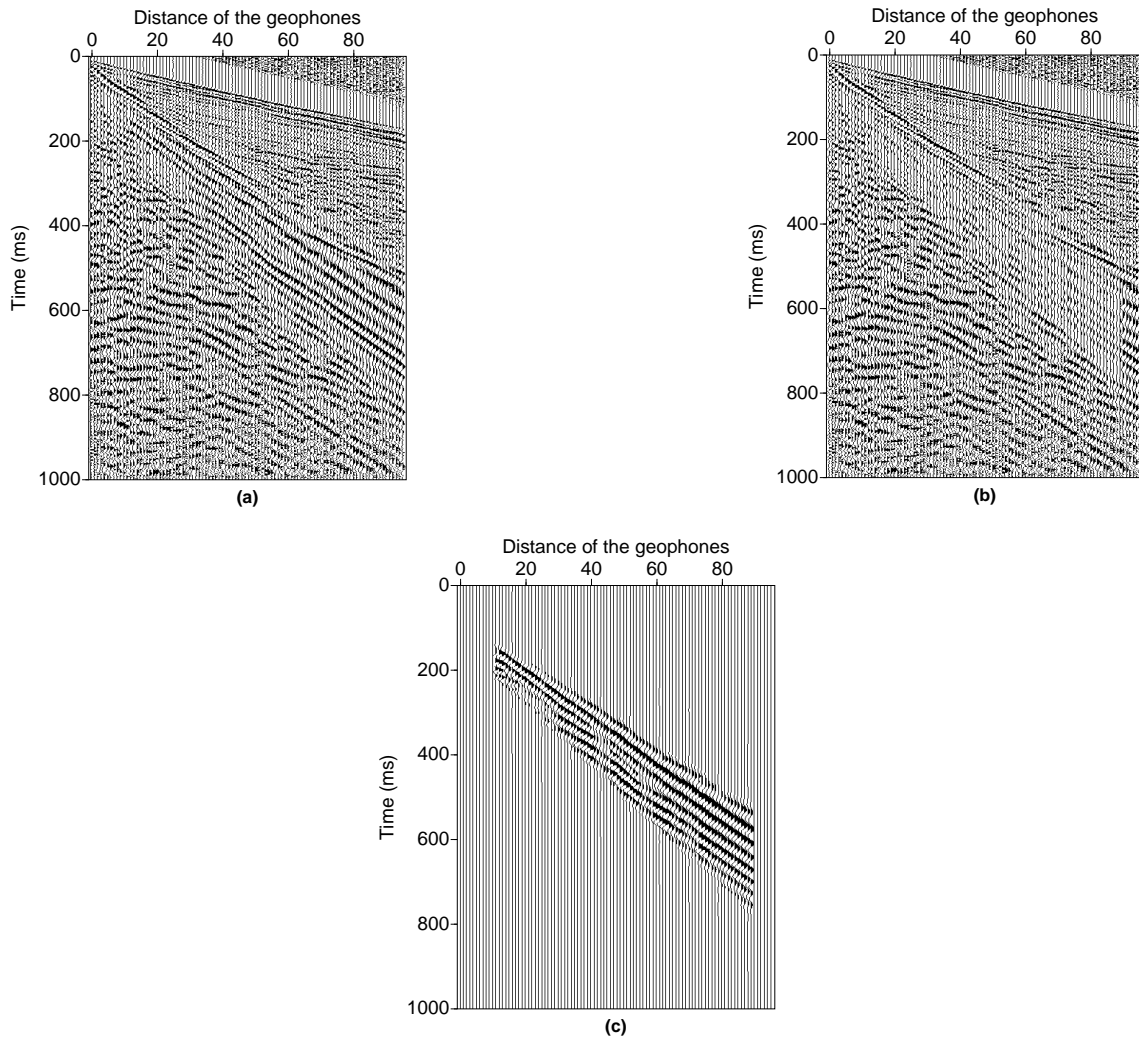


Figura 4.1: (a) O sismograma original formado por 96 traços (horizontal) e 1001 amostra (vertical) por traço. O *ground roll* aparece (bem visível) em uma região que vai do canto superior esquerdo ao canto inferior direito. Em (b) temos o sismograma atenuado pelo fator ótimo de $A_f = 0.80$. Em (c) temos a porção que foi removida do sismograma original. A região atenuada está nos limites (horizontal) dos traços 12 – 90 e (vertical) amostras de 125 – 230 (traço 12) e 470 – 780 (traço 90).

tas $Wf[n, 2^j]$. Como discutido no trabalho de Corso [Corso et al. (2003)], a base escolhida foi a derivada quinta da gaussiana e os cálculos mostraram que a ordem da derivada não é relevante para a filtragem. O filtro em ondaletas atuou na região selecionada e a atenuação foi feita de acordo com a frequência baixa presente no ruído coerente *ground roll*. Para cada traço, um limiar na frequência ω_f (limiar de corte) foi calculado como sendo a frequência média entre as duas maiores amplitudes no espectro de Fourier. A frequência abaixo de ω_f (o *ground roll* deve está relacionado com frequência abaixo deste limiar) são atenuadas de A_f . Então o processo de filtragem depende do fator de atenuação A_f . Os coeficientes em ondaletas atenuados de A_f são definidos como $A_f \cdot W^*f[n, 2^j]$ e os coeficientes sem ruídos (sinal limpo) $\tilde{W}f[n, 2^j]$. Similarmente, pode-se escrever o sinal limpo como; $\tilde{W}f[n, 2^j] = Wf[n, 2^j] - A_f \cdot W^*f[n, 2^j]$.

3. *Seleção do fator de atenuação ótimo.* Os valores de A_f , em princípio, foram escolhidos convenientemente para se obter uma atenuação eficiente. Em seguida a decomposição KL foi realizada no registro sísmico filtrado (parte removida representada na Figura 4.1(c)) para obter o espectro de energia KL. Para calcular a transformada Karhunen-Loève sobre a região considerada (na forma trapezoidal) foi necessário primeiro realizar uma transformação linear. Como descrita por Montagne e Vasconcelos [Montagne and Vasconcelos (2006a)], essa transformação é realizada para se obter uma região retangular que possa ser representada por uma matriz $\tilde{M} \in \mathbb{R}^{\tilde{m} \times \tilde{n}}$. Deve-se perceber que com a

variação do fator de atenuação, os valores da matriz \tilde{M} mudam. Porém, sua dimensão permanece inalterada sendo que, para cada valor de A_f , é obtida uma matriz de dados \tilde{M}_{A_f} e os autovalores $\tilde{\lambda}_i$ são obtidos da matriz de correlação correspondente $\tilde{\Gamma} = \tilde{M}_{A_f} \tilde{M}_{A_f}^t$. A soma dos autovalores da matriz Γ é definida como a energia total $E = \sum_i \lambda_i$ contida na decomposição KL. Então, λ_i é a energia contida no autovetor empírico \vec{u}_i , tal que a energia relativa E_i contida em um modo KL é dada por;

$$\tilde{E}_i(A_f) = \frac{\tilde{\lambda}_i}{\sum_{i=1}^{\tilde{r}} \tilde{\lambda}_i}, \quad (4.3)$$

onde \tilde{r} é o posto (rank) da matriz $\tilde{M} \in \mathbb{R}^{\tilde{m} \times \tilde{n}}$. O critério adotado para selecionar o melhor fator de atenuação (A_f) foi o seguinte: as energias relativas dos modos KL foram graficados contra os fatores de atenuações A_f ($\tilde{E}_i \times A_f$ e representados na Figura (4.2)). O valor ótimo de A_f foi selecionado onde os modos KL (primeiros modos) apresentam um mínimo. Este critério faz sentido físico (para atenuar o *ground roll*) porque o mínimo do primeiro modo representa o ponto onde o modo mais energético (ou coerente) é atenuado eficientemente. Também, apresenta resultados práticos muito bons.

4.2.1 Resultados

A transformada em ondaletas foi aplicada ao dado sísmico original (Figura 4.1(a)) sendo que o filtro atuou somente na região destacada para um certo valor A_f .

Os valores usados para o fator de atenuação A_f foram $A_f = 0.5 - 0.95$ com intervalos de 0.05. Então, tem-se uma sequência de vinte aplicações gerando os padrões de dados sísmicos. Os padrões mostrados nas Figuras 4.1(b) e 4.1(c), obtidas pela transformada inversa em ondaletas (atenuados de A_f ótimo), respectivamente, são o atenuado e a porção que foi removida do registro original. Perceba que a região escolhida tem limites na horizontal do traço 12 – 90. Os limites na vertical são as amostra de 125 – 230 (para o traço 12) e 470 – 780 (para o traço 90).

Depois da transformação linear, a região trapezoidal pode ser representada por uma matriz $\tilde{M}_{A_f} \in \mathbb{R}^{\tilde{m} \times \tilde{n}}$ com $\tilde{m} = 310$ e $\tilde{n} = 78$. Para cada matriz de dados \tilde{M}_{A_f} foi obtido o espectro de energia KL (como na equação (4.3)). A Figura (4.2) mostra a energia relativa \tilde{E}_i para os cinco primeiros modos KL como função do fator A_f . Pode-se ver claramente nesta figura que o primeiro modo de energia \tilde{E}_1 atinge um mínimo em $A_f = 0.80$. Isto significa que para este valor de A_f , o modo mais energético da região selecionada é atenuado eficientemente. Note, também, que o mínimo de \tilde{E}_1 é bastante preciso no sentido de que há um aumento rápido de \tilde{E}_1 a partir do valor ótimo $A_f = 0.80$. Isso indica que usando o mínimo do primeiro modo de energia KL é um critério eficiente para selecionar o melhor fator de atenuação A_f . Ainda na mesma figura, pode ser visto que o segundo modo de energia KL apresenta um mínimo também em $A_f = 0.80$. Este fato adiciona reforços a escolha de $A_f = 0.80$ como o fator de atenuação ótimo. Em outras palavras, o melhor fator de atenuação A_f (em um filtro baseado nas ondaletas para atenuar o *ground roll* dos dados sísmicos) pode ser escolhido por procurar o mínimo dos primeiros modos de energia como função do fator de atenuação A_f .

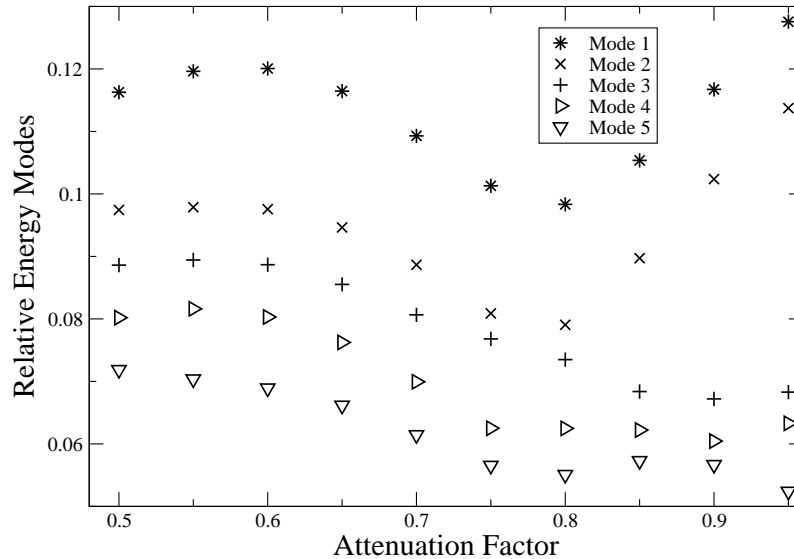


Figura 4.2: A energia relativa \tilde{E} dos 5 primeiros modos KL da região atenuada mostrada como uma função do fator de atenuação A_f . O mínimo dos 2 primeiros modos caracterizam o fator de atenuação ótimo.

A Figura (4.3) está mostrando a distribuição dos modos de energia (para a região considerada) do sismograma sísmico original (linha sólida) e para o sismograma atenuado (linha tracejada) para o valor $A_f = 0.80$. A característica geral vista nesta figura é que depois da filtragem os 10 modos menores (maiores) contêm relativamente menos (mais) energia que os modos do dado original. Este foi um resultado desejado no sentido de que pretende-se reduzir a energia do ruído coerente *ground roll*. Na parte em detalhe (*inset*) da Figura (4.3) estão visualizados os 10 primeiros modos da região atenuada (barras) e do dado original (*). Deve-se notar, em particular, que enquanto os dois primeiros modos KL contem, aproximadamente, 30% da energia total da região do registro original, esses dois primeiros modos responde por

menos que 20% da energia da região filtrado. Isto indica uma redução significativa dos modos mais coerentes.

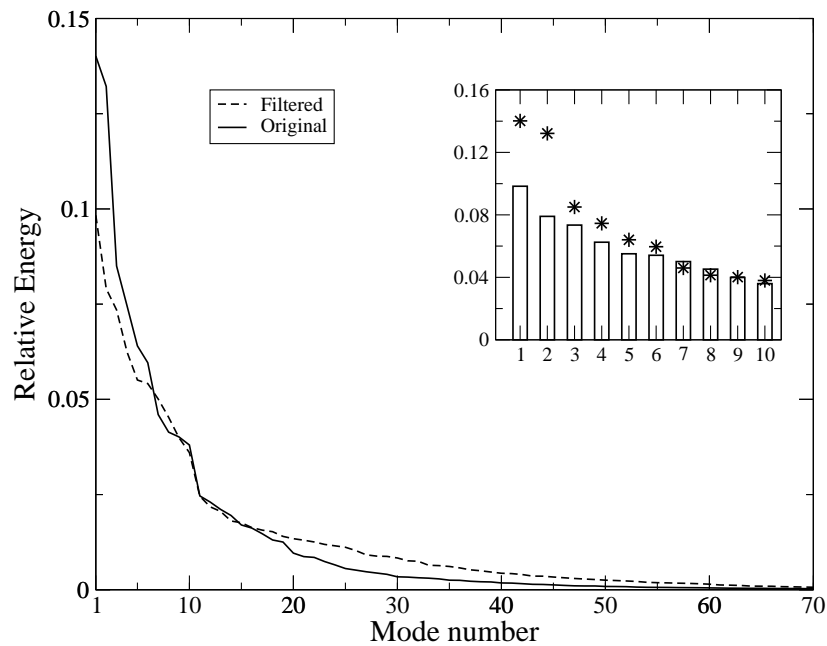


Figura 4.3: A distribuição de energia KL em função dos modos antes e depois de aplicar o filtro ondaleta ótimo com fator de atenuação $A_f = 0.80$. A linha sólida (tracejada) corresponde ao espectro de energia da região original (atenuada). No gráfico em detalhe (*inset*) os 10 primeiros modos são mostrados com barras (*star*) indicando os espectros da região atenuada (original).

4.3 Comentários finais

Uma vez que o registro sísmico original é decomposto em uma base apropriada de ondaletas, o filtro atua somente na região considerada. Esta região é definida previamente e, portanto, o processo de filtragem não afeta o restante do sinal. Os coeficientes da escala com frequência dominante na faixa do ruído *ground*

roll são atenuados pelo fator de atenuação A_f , sendo seu valor o mesmo para todos os traços. O fato de que este filtro, assim, construído, atua localmente é uma consequência direta da localização em tempo-escala das ondaletas. Como as ondaletas tem suporte limitado no tempo, torna-se possível a ação do filtro estritamente à região delimitada. Esta propriedade torna este método de filtragem uma ferramenta bastante adequada no processamento de sinais sísmicos.

A eficiência do método leva em conta o fator de atenuação ótimo, o qual foi encontrado pelo espectro de energia KL (\tilde{E}) do dado filtrado. Onde o primeiro modo apresentar um mínimo o fator A_f deve ser escolhido como sendo ótimo. É interessante notar que, usando o mesmo registro sísmico e a mesma base em ondaletas, Corso [Corso et al. (2003)] tinha obtido o mesmo valor de atenuação $A_f = 0.80$. Isto é, o sinal na região escolhida (identificada, por exemplo, por inspeção visual) foi atenuado de 80%. Porém, agora, nesse trabalho foi usado um critério quantitativo (critério de extremo) para encontrar o fator ótimo de atenuação. Essa é a vantagem maior desse método comparado com a inspeção visual descrita por Corso [Corso et al. (2003)]. Embora o direcionamento do método seja a supressão do *ground roll*, este não está restrito à aplicações geofísicas. A mensagem principal é que este método pode ser implementado em processos de análise baseados em tempo e frequência.

Capítulo 5

Supressão do *Ground roll*: decomposição em multiescala

5.1 Introdução

Nos capítulos anteriores, foram descritas transformações matemáticas, muito utilizadas na Física Estatística, com uma atenção especial voltada para a transformada em ondaletas. Das equações (1.2)-(1.14) é fácil perceber como a transformada em ondaletas fornece um mecanismo para uma representação em análise tempo-escala. Por estas equações, percebe-se que a variação de σ causa o efeito de dilatação (para $\sigma > 1$) ou contração (para $\sigma < 1$) para a função ondaleta de análise considerada. Então, quando σ aumenta, a função ondaleta “dilata-se” para uma análise das características de $f(t)$ em escalas maiores (comportamento de f em tempos longos). As características em tempos curtos aparecem quando σ decresce. Por outro lado, a variação em τ significa uma análise translacional (posição da ondaleta no tempo) da função f em pontos distintos τ . A compreensão destas conexões é importante mesmo que a transformada contínua em ondaletas provê uma representação redundante do espaço amostral. Contudo, para analisar dados discre-

tos, como foi colocado anteriormente, é necessário uma discretização das ondaletas nos parâmetros σ e τ . Essa discretização está descrita na seção (1.3). A expansão em série dada pela equação (1.26) é similar a série de Fourier. Sendo que a expansão em ondaletas é duplamente indexadas e com propriedade de localização em tempo-escala (equivalente a tempo-frequência). Com essa propriedade a transformada em ondaletas é uma ferramenta bastante flexível para processamento de dados sísmicos.

Neste capítulo apresentamos um método eficiente para a remoção do ruído coerente *ground roll* baseado nas ondaletas discretas. A eficiência desse método está na *Decomposição em Multiescala* do dado sísmico original utilizando a Transformada Discreta em Ondaletas. Com essa decomposição, o sismograma original fica representado por escalas individuais em padrões semelhantes (aspecto *2D*). Assim, as características que não são detectadas em uma determinada escala, podem ser localizadas em outras. Como as escalas da transformada em ondaletas estão relacionadas com a frequência, então uma decomposição em multiescala representa uma “separação” nas frequências do sinal.

A ocorrência da heterogeneidade subsolo terrestre é devido as suas camadas, falhas, fraturas e, isso insinua que os modelos desenvolvidos para uma análise devem levar em conta a dependência nas escalas. Desde que os eventos com propriedades físicas de interesse (como as hipérbolas da Figura 3.3(c)) e os indesejáveis (como o ruído *ground roll* representado na Figura 3.3(b)) são não-estacionários, os espectros do sinal de interesse e o ruído podem se superpor. Porém, os intervalos na frequência para o ruído *ground roll* e eventos de interesses são diferentes. Então, uma decomposição em multiescala significa uma distribuição da energia (por es-

calas) nos coeficientes ondaletas $d_{j,k}$. Dessa forma, cada escala está relacionada com características diferentes e será representada com mesmo padrão (mesma quantidade de traços e amostras) que o sinal sísmico original. Assim, é permitida uma identificação (que pode ser visual) adequada de eventos particulares. Essa técnica nova de decompor o sinal sísmico, traço a traço, em escalas individuais, aumenta o controle para filtragem do ruído *ground roll*.

O sismograma utilizado para a remoção do *ground roll* está representado na Figura 5.1(a). Ele foi obtido pela página eletrônica da base de dados do *Center for Wave Phenomena* da *Colorado School of Mines* e está descrito na referência [Yilmaz (2003)]. É um dado obtido de um levantamento sísmico terrestre. A fonte geradora das ondas sísmicas é do tipo dinamite. A localização de fonte é central e os geofones com distribuição uniforme. Esse dado é composto por 90 traços e 2048 amostras por traço. No eixo horizontal está a representação do espaçamento entre os geofones (alguns metros) e cada geofone representa um traço sísmico (uma série temporal discreta). Na vertical tem-se a quantidade das amostras (nesse caso 2048) em representação temporal (ms). Para a decomposição em multiescala, o índice de escala j varia de 1 a 11. Pois o registro sísmico possui 2048 amostras por traço e essa quantidade equivale a 2^{11} para transformada discreta que faz uso da potência de dois no processamento (transformações diádicas). Na representação em escala, o padrão sísmico para o índice de escala $j = 1$ (índice de escala menor) significa uma representação na mais fina resolução (frequências maiores) e índice $j = 11$ tem a mais grossa resolução com o mais baixo conteúdo na frequência, onde as amplitudes do ruído *ground roll* devem ser dominante.

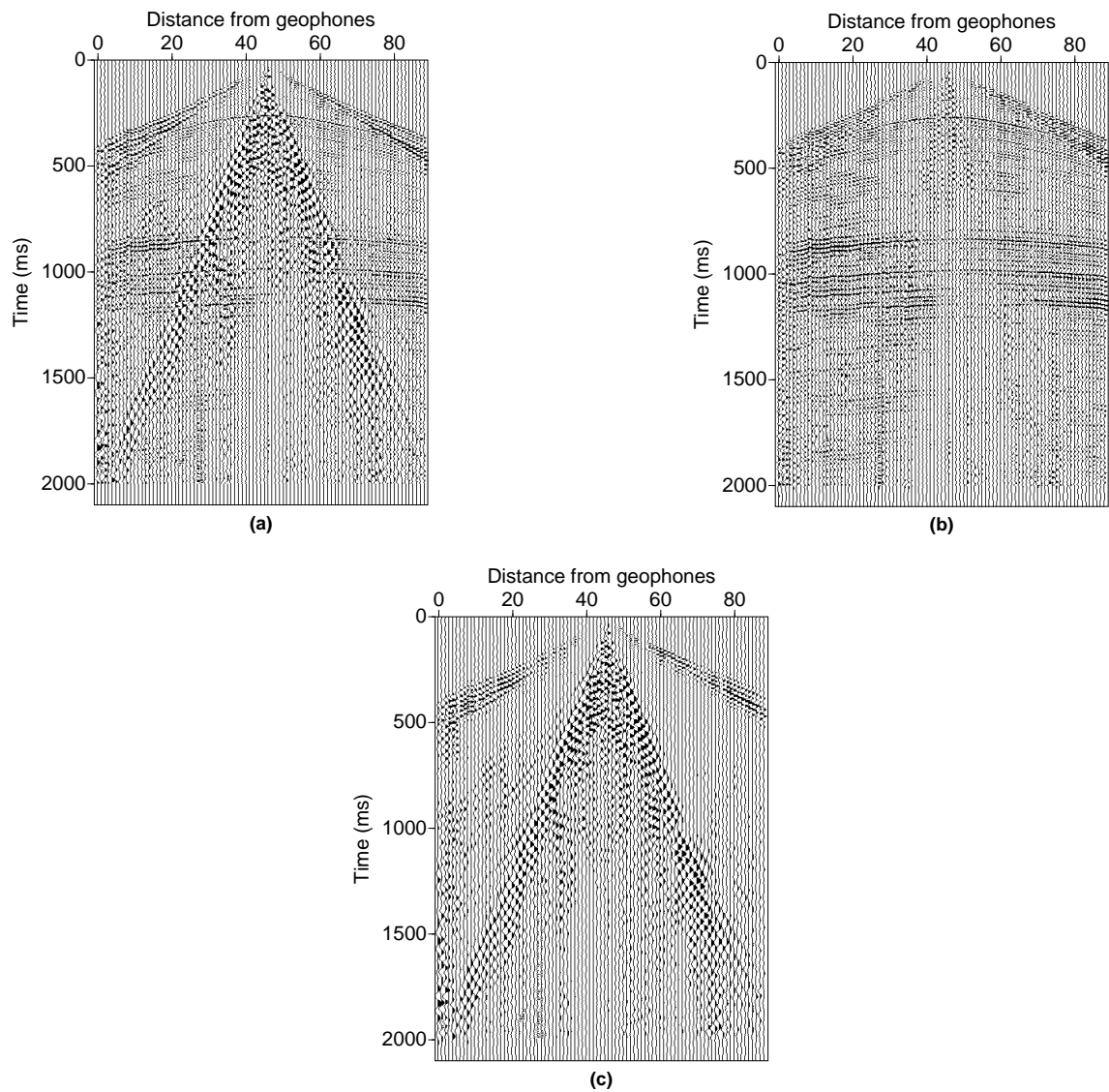


Figura 5.1: (a) O sismograma original onde o ruído *ground roll* está presente e atrapalhando as reflexões. Ele é composto por 90 traços e 2048 amostras por traço. Em (b) temos o sismograma atenuado. Agora as reflexões são dominantes. Em (c) temos a porção que foi removida do sismograma original.

5.1.1 Remoção do *ground roll*

A transformada em ondaletas foi utilizada, anteriormente, no desenvolvimento de métodos para a remoção do ruído coerente *ground roll* de registros sísmicos [Corso et al. (2003), Deighan and Watts (1997)]. No filtro desenvolvido por Corso et al. (2003), o *ground roll* foi identificado em uma região (espaço-tempo) e a atenuação ocorreu em uma escala escolhida. Como a região escolhida do dado sísmico pode não corresponder, exclusivamente, as componentes do ruído *ground roll*, então um fator de atenuação A_f foi definido. Se $Wf[n, 2^j]$ são os coeficientes da transformada em ondaletas para o sismograma original e $A_f W^* f[n, 2^j]$ os coeficientes em ondaletas atenuados por A_f , então defini-se o sismograma limpo por $\tilde{W}f[n, 2^j] = Wf[n, 2^j] - A_f W^* f[n, 2^j]$. Para o valor $A_f = 0$ não há filtragem (atenuação) e, no caso limite oposto, $A_f = 1$ a região considerada do sismograma é completamente atenuada. A reconstrução dos coeficientes $\tilde{W}f[n, 2^j]$ é possível (transformada inversa em ondaletas) baseada na discretização redundante de Grossmann que equivale a discretização da equação (1.14).

Como mencionado anteriormente, o método apresentado aqui para a supressão do *ground roll* é baseado na decomposição em multiescala. Com a decomposição em multiescala, o ruído *ground roll* ficará restrito aos coeficientes das escalas maiores (frequência baixa). Dessa maneira, é possível uma localização visual (como pode ser visto na Figura (5.8)) deste ruído. Portanto, o procedimento de filtragem consiste em atenuar os coeficientes em ondaletas na região que corresponde a energia do ruído. Esse processo é realizado por traço-a-traço do sismograma.

5.2 Remoção do *ground roll*: decomposição em multiescala

O método consiste em uma decomposição em multiescala do registro sísmico considerando as propriedades de localização tempo-escala da transformada em ondaletas discretas. O procedimento é realizado por três etapas: (i) decomposição em multiescala do registro sísmico, (ii) reconstrução por escala individual e (iii) atenuação do ruído *ground roll* na escala onde suas amplitudes são dominantes e, em ambos, usando ondaletas discretas. Agora descrevo em mais detalhes cada etapa. Na primeira etapa (i) o dado sísmico original foi decomposto, traço a traço, no espaço das ondaletas (em sua forma discreta) pelo método da transformada rápida em ondaletas. Como representado na Figura (5.2) para os onze níveis de escalas, o sinal sísmico fica caracterizado pelos coeficientes $d_{j,k}$ em suas respectivas escalas individuais. O sinal de entrada pertence ao subespaço V_0 e corresponde aos coeficientes $a_{j=0}$. Para cada traço, o sinal possui $n = 2048$ amostras por traço. A aplicação dos filtros h e g (veja Apêndice B) correspondem a uma decomposição dos dados, gerando coeficientes a_1 e d_1 . As duas partes ficam representadas nos subespaços V_1 e W_1 , respectivamente, possuindo $\frac{n}{2}$ amostras cada. Os coeficientes $d_1 \in W_1$ são mantidos e o processo de decomposição, para o próximo nível, é repetido sobre os coeficientes $a_1 \in V_1$. Novamente, a ação dos filtros h e g gera os coeficientes a_2 e d_2 , ambos com $\frac{n}{4}$ amostras e caracterizados nos subespaços V_2 e W_2 , respectivamente. Os coeficientes $d_2 \in W_2$ são mantidos e o processo é repetido sobre os a 's até o nível $j = 11$ gerando os coeficientes $a_{11} \in V_{11}$ e $d_{11} \in W_{11}$ os quais terão apenas uma

amostra. Como em cada nível, a soma da quantidade das amostras é a mesma que a quantidade das amostras no nível anterior, então o sinal original ficará caracterizado pelos coeficientes d_1, d_2, \dots, d_{11} e a_{11} , onde a soma é n amostras.

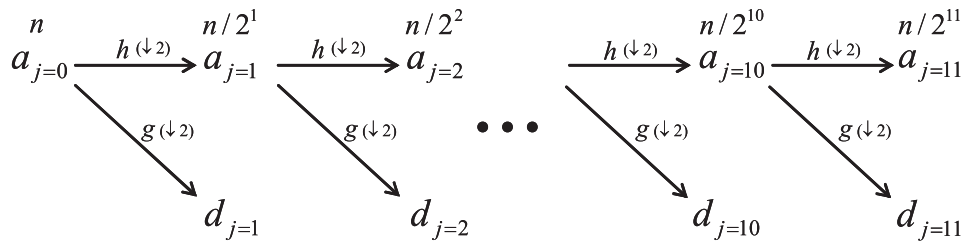


Figura 5.2: Representação do método de decomposição rápida em ondaletas para o registro sísmico representado na Figura 5.1(a). Supõe-se que o traço sísmico inicial está representado em V_0 com n amostras e o processo ocorre da interação com o filtro h e g seguido por uma decimação de fator 2.

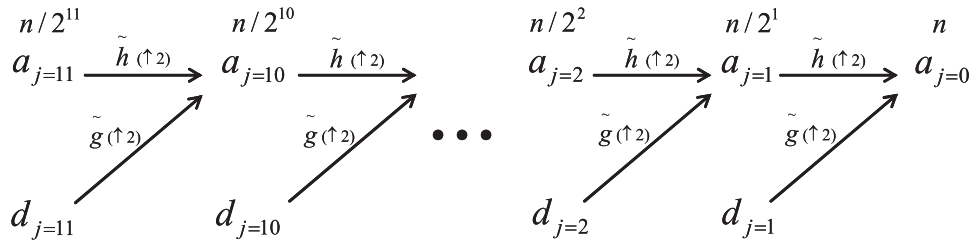


Figura 5.3: Representação do método de reconstrução rápida em ondaletas. O processo começa com os coeficientes a_{11}, d_{11} e ocorre a interação com os filtros \tilde{h} e \tilde{g} seguidos por uma inserção de fator 2.

Na etapa segunda (ii) os coeficientes d_j são reconstruídos (transformada inversa em ondaletas discretas) pelo método da transformada inversa rápida (como mostrado na Figura (5.3)) para cada índice j individual. De acordo com a equação

(1.26), para uma determinada escala (por exemplo índice $j = 1$) a reconstrução desse padrão sísmico pode ser representada pela fórmula $f_{j=1}(t) = \sum_k d_{j=1,k} \psi_{j=1,k}$ com todos os coeficientes $d_{j \neq 1,k}$ zerados. Para a reconstrução do padrão sísmico na escala seguinte segue o mesmo procedimento, porém, com os coeficientes $d_{j \neq 2,k}$ zerados. A fórmula da reconstrução para essa escala pode ser representada por $f_{j=2}(t) = \sum_k d_{j=2,k} \psi_{j=2,k}$. Para se obter a mesma quantidade de amostras que o registro original, a fórmula de reconstrução deve ser aplicada sobre todas as escala. Porém, a notação acima consiste em dizer que apenas na escala considerada os coeficientes são não-nulos. Como os índices de escalas são $j = 1, \dots, 11$, tem-se então 11 padrões sísmicos como os representados nas Figuras (5.5)-(5.11).

O objetivo exposto na etapa terceira (iii) é a atenuação do ruído *ground roll* pelo método desenvolvido por Corso et al. (2003), só que aqui é utilizado ondaletas na sua forma discreta. Com a decomposição em multiescala, percebe-se que o ruído *ground roll* torna-se mais localizado no domínio tempo-escala. Portanto, pode ser facilmente, sujeito a uma remoção cirúrgica a qual acontecerá em cada escala. Nas escalas onde o *ground roll* e reflexões estão presentes (superpostos), é usado um fator de atenuação A_f . Então, o filtro em ondaletas trabalha reduzindo os coeficientes (representando o *ground roll* em escala) por um fator A_f . A escolha do fator A_f depende da quantidade visual do ruído *ground roll* presente no dado sísmico (razão reflexões-ruído). Por exemplo, para o padrão sísmico representado na Figura (5.8) (onde o *ground roll* aparece em uma pequena região superior central) o valor do fator de atenuação A_f deve ser escolhido menor que para o padrão representado na Figura (5.9), onde as amplitudes do *ground roll* dominam o sismograma inteiro. Para as

escalas onde somente o *ground roll* está presente, o padrão sísmico é desconsiderado. O fator de atenuação não é usado nas escalas onde o *ground roll* não é visível e esse padrão sísmico será considerado para a reconstrução do sismograma o qual será o sinal atenuado. A reconstrução é feita por uma soma simples dos padrões sem o *ground roll*.

O resultado da decomposição em multiescala está mostrado nas Figuras (5.5)-(5.11). A Figura (5.5) está mostrando um padrão sísmico para o índice de escala $j = 1$. Esta escala contém as frequências maiores, para o dado analisado, as quais são dominadas por ruídos não coerentes. Os padrões de frequências relativamente altas (porém menor que anterior) das Figuras (5.6) e (5.7) são para os índices intermediários $j = 2$ e $j = 3$, respectivamente. Eles são dominados por eventos de reflexões (hipérboles) e nenhum ruído coerente *ground roll* está presente. Portanto, esses padrões sísmicos (nessas escalas) representam partes do sinal com estruturas geofísicas da subsuperfície (sinais de interesse). Logo, o fator de atenuação A_f não é usado nestas escalas e estes padrões são mantidos para a reconstrução do registro sísmico onde a energia do ruído *ground roll* não deve estar presente. O índice de escala aumenta, e o *ground roll* começa a aparecer. Isso torna-se evidente com a visualização da Figura (5.8). O padrão mostrado nesta figura é para o índice intermediário de escala $j = 4$ e pode-se ver que o *ground roll* está presente em uma pequena região localizada no topo central do sismograma. Com isso, torna-se fácil a eliminação do *ground roll* sem afetar as outras regiões. O filtro baseado nas ondas discretas é utilizado para atenuar o *ground roll* com um fator A_f somente nessa região e, conseqüentemente, não afetando as hipérboles localizadas na parte central

do padrão sísmico.

A Figura (5.9) corresponde ao índice intermediário $j = 5$. É perceptível que as amplitudes do *ground roll* são dominantes e as reflexões estão praticamente invisíveis. Portanto, apesar de poucas, porém de interesse físico, o padrão sísmico nesta escala não pode, simplesmente, ser desconsiderado. Contudo, para preservar a predominância do sinal de relevância física (hipérboles), foi utilizado o fator de atenuação A_f nessa escala de tal forma que seja eliminado o máximo de ruído e mantendo o máximo das reflexões. A região escolhida foi a que forma um cone, sendo exatamente a região do *ground roll*. As Figuras (5.10) e (5.11) representam os padrões sísmicos correspondendo aos índices de escalas $j = 6$ e $j = 7 - 11$, respectivamente. Para a Figura (5.11) tem-se um conjunto de índice de escala ($j = 7 - 11$) sendo representado por um único padrão sísmico. Isso porque, como pode-se perceber já na figura anterior, os padrões nestas escalas correspondem somente ao ruído coerente *ground roll* e nenhuma informação física relevante está presente. Então, não há necessidade de uma representação individual desses padrões nos índices de escalas $j = 8, 9, 10, 11$, pois já se sabe que contêm informações referente somente ao *ground roll*.

O resultado para todo o processo de filtragem do ruído *ground roll*, de acordo com as etapas descritas acima e envolvendo a técnica eficiente de decomposição em multiescala, está mostrado nas Figuras 5.1(b) e 5.1(c). Na figura 5.1(a) tem-se a representação do sismograma em seu padrão original, onde as amplitudes do *ground roll* atrapalham na visualização (interpretação) das hiperboles que são sinais de interesse com estruturas geológicas. A figura 5.1(b) está representando o sismograma

reconstruído após a remoção do *ground roll* (sismograma atenuado). Percebe-se visivelmente, que o ruído foi suprimido de tal forma que as reflexões de interesse foram recuperadas. Isso torna-se evidente quando se observa a figura 5.1(c), a qual está ilustrando a “porção” que foi removida. Pode-se perceber, ainda neste padrão sísmico, comparado com o original e atenuado, apenas o sinal referente ao *ground roll* (forma de um cone como no sismograma original) está presente. Dessa forma, a parte removida do sismograma original não afetou o sinal de interesse. Portanto, a energia que caracteriza as reflexões é preservada.

Para ilustrar em mais detalhes o efeito da ação do filtro na frequência, apresentamos um traço sísmico juntamente com sua Transformada de Fourier. Na Figura 5.4 temos a representação para o (topo) sismograma original, (meio) sismograma atenuado e (baixo) porção removida do sismograma original. O traço escolhido foi o representado pelo geofone de número $l = 65$. Para cada sismograma mostrado nas Figuras (5.5)-(5.11) tem-se, abaixo, o seu traço na representação convencional (série temporal) e seu espectro de Fourier. Este tipo de informação deve estar presente, pois, dá informações sobre as frequências por escalas. O espectro de Fourier mostra o conteúdo nas frequências e percebe-se que o *ground roll* corresponde às características nas frequências baixas. Olhando a figura (5.4) torna-se evidente que as frequências atenuadas são compatíveis com o ruído coerente *ground roll*.

5.3 Comentários finais

Existe uma diversidade de técnicas e algoritmos baseados nas ondaletas para a remoção do *ground roll*. No início da minha pesquisa nesta área, a preocupação maior foi entender todas as técnicas e ferramentas. A maioria dos métodos é comercializada e utilizada pela indústria do petróleo. Isso, sugere ao autor que as técnicas devem ser empregadas às situações reais com confiabilidade nas interpretações dos conceitos dentro do contexto geofísico. Portanto, procurei descrever de forma clara para os usuários, detalhes da técnica e algoritmo utilizados no processamento do registro sísmico aqui analisado.

Neste capítulo, nós apresentamos um método alternativo de remoção do *ground roll*. A ferramenta utilizada foi a transformada discreta em ondaletas baseada na técnica da *decomposição em multiescala*. As duas abordagens foram verificadas para um registro sísmico (contendo o *ground roll*) e suas validades foram comprovadas no sentido de que a remoção aconteceu de forma conveniente. Os resultados indicam as propriedades bem particulares do registro. Esta perspectiva foi utilizada para identificar as diversas regiões do registro com a presença do *ground roll*. E por fim, a região foi atenuada de A_f escala por escala onde o ruído é dominante.

A decomposição em multiescala tornou-se capaz de localizar e identificar as características do dado sísmico em uma dada escala, verificando se o *ground roll* está presente. Esse procedimento pode ser feito em uma forma eficiente, filtrando o ruído coerente somente na região do espaço-tempo e escalas na qual o ruído está presente e guardando as outras regiões e escalas não filtradas. Este método de filtragem é

muito mais seletivo e cirúrgico porque é realizado em escala por escala. Esta é a razão porque ele é mais eficiente. Um outro detalhe importante das propriedades das ondaletas e que deve ser notado, é que os padrões sísmicos mantêm sua posição no tempo enquanto o índice de escala varia.

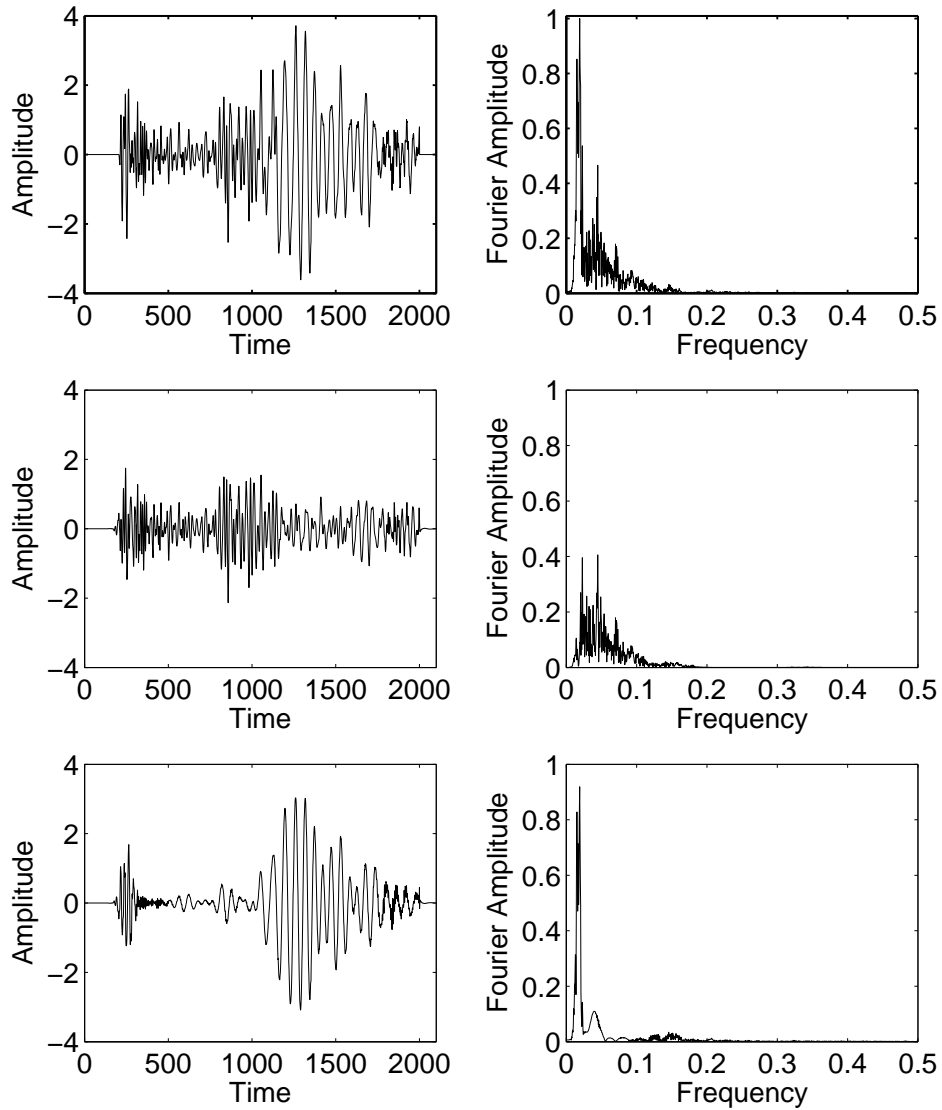


Figura 5.4: Um único traço representado em seu formato convencional (série temporal) e seu espectro de Fourier. (topo) temos o traço 65 do sismograma original. (meio) temos o mesmo traço para o sismograma atenuado e (baixo) o traço da porção que foi removida do sismograma original.

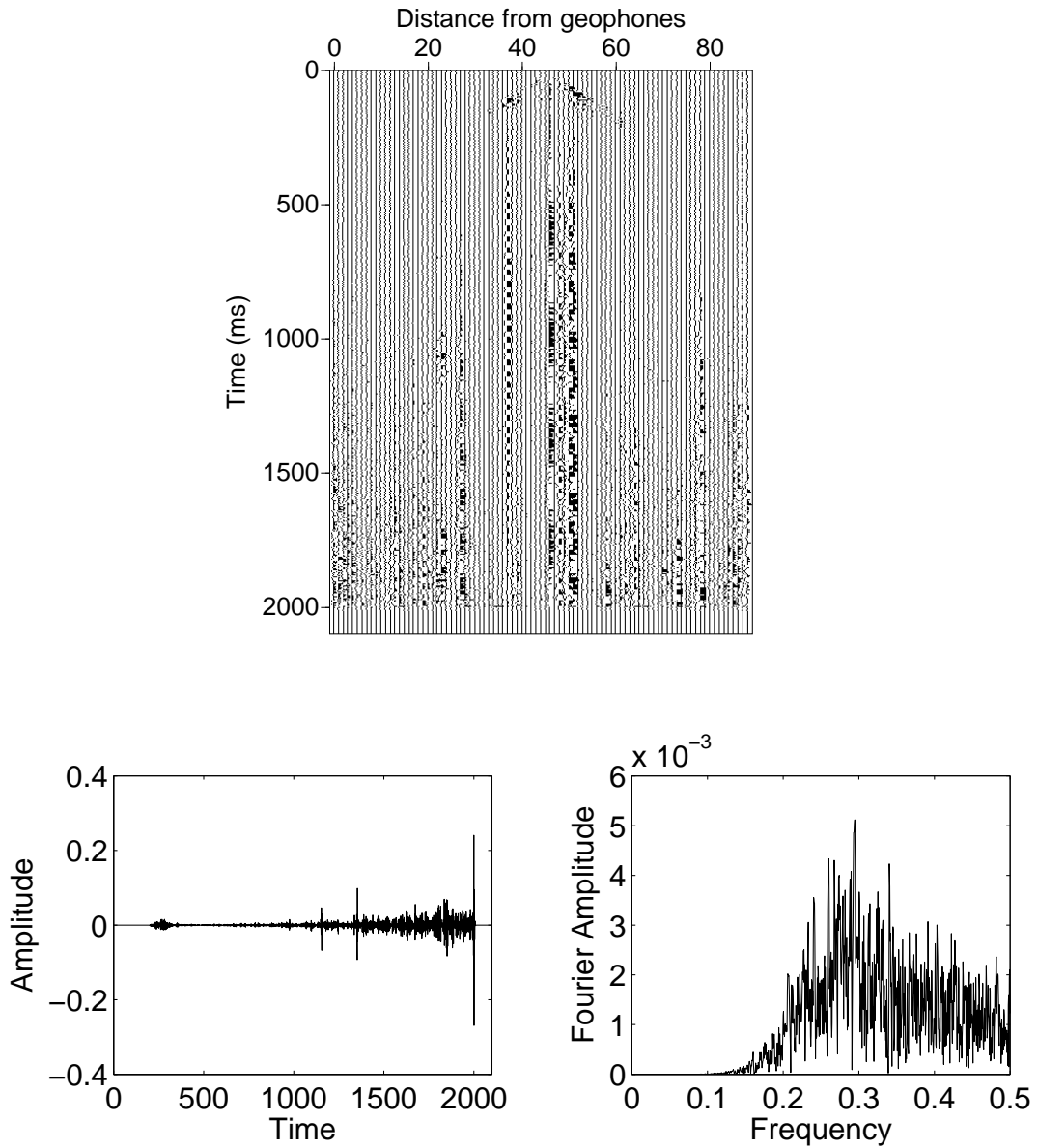


Figura 5.5: O padrão sísmico correspondendo ao menor índice de escala $j = 1$. Este registro contém as mais altas componentes de frequências e o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier.

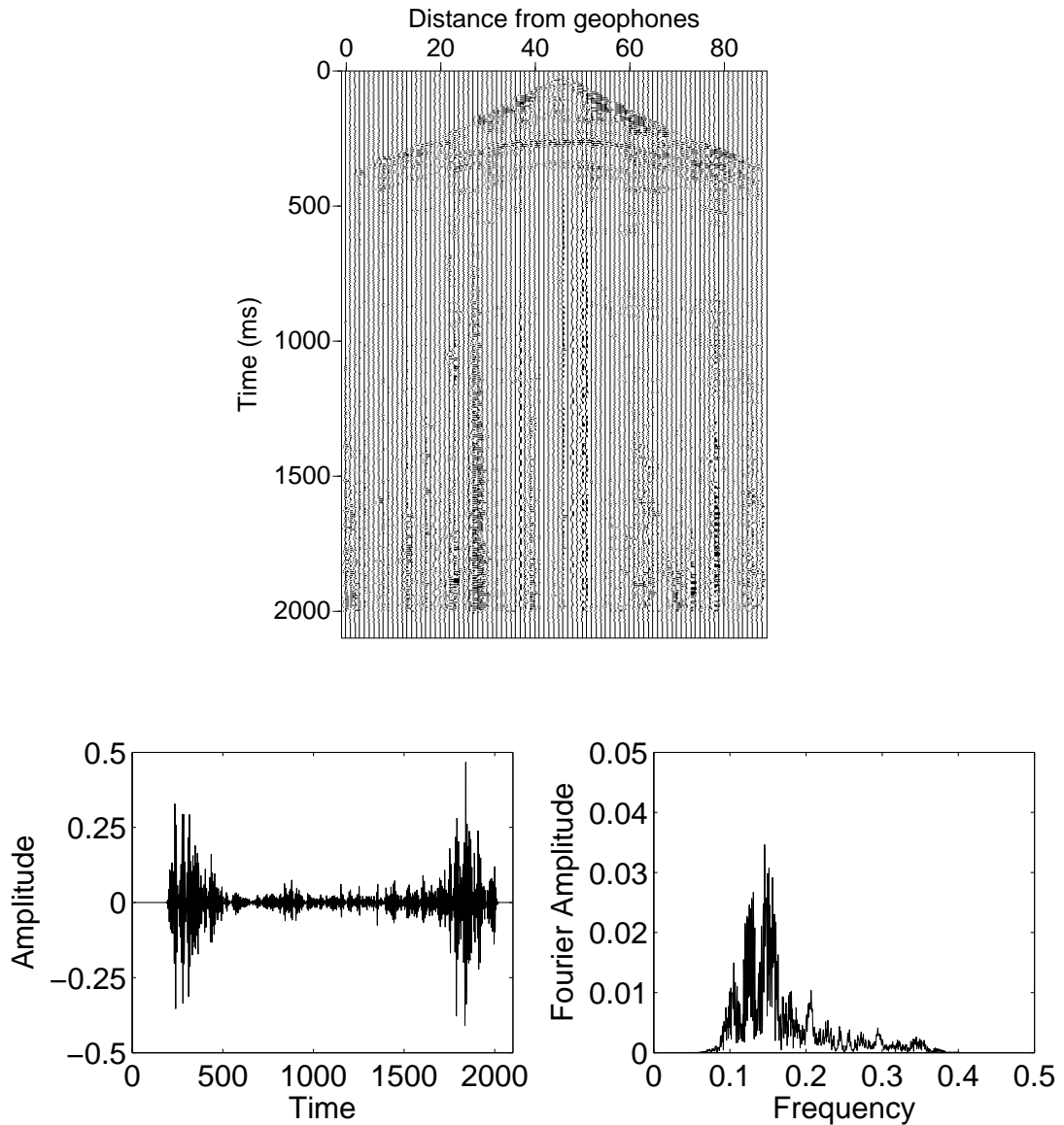


Figura 5.6: O padrão sísmico correspondendo ao índice intermediário de escala $j = 2$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta somente eventos de reflexões. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier.

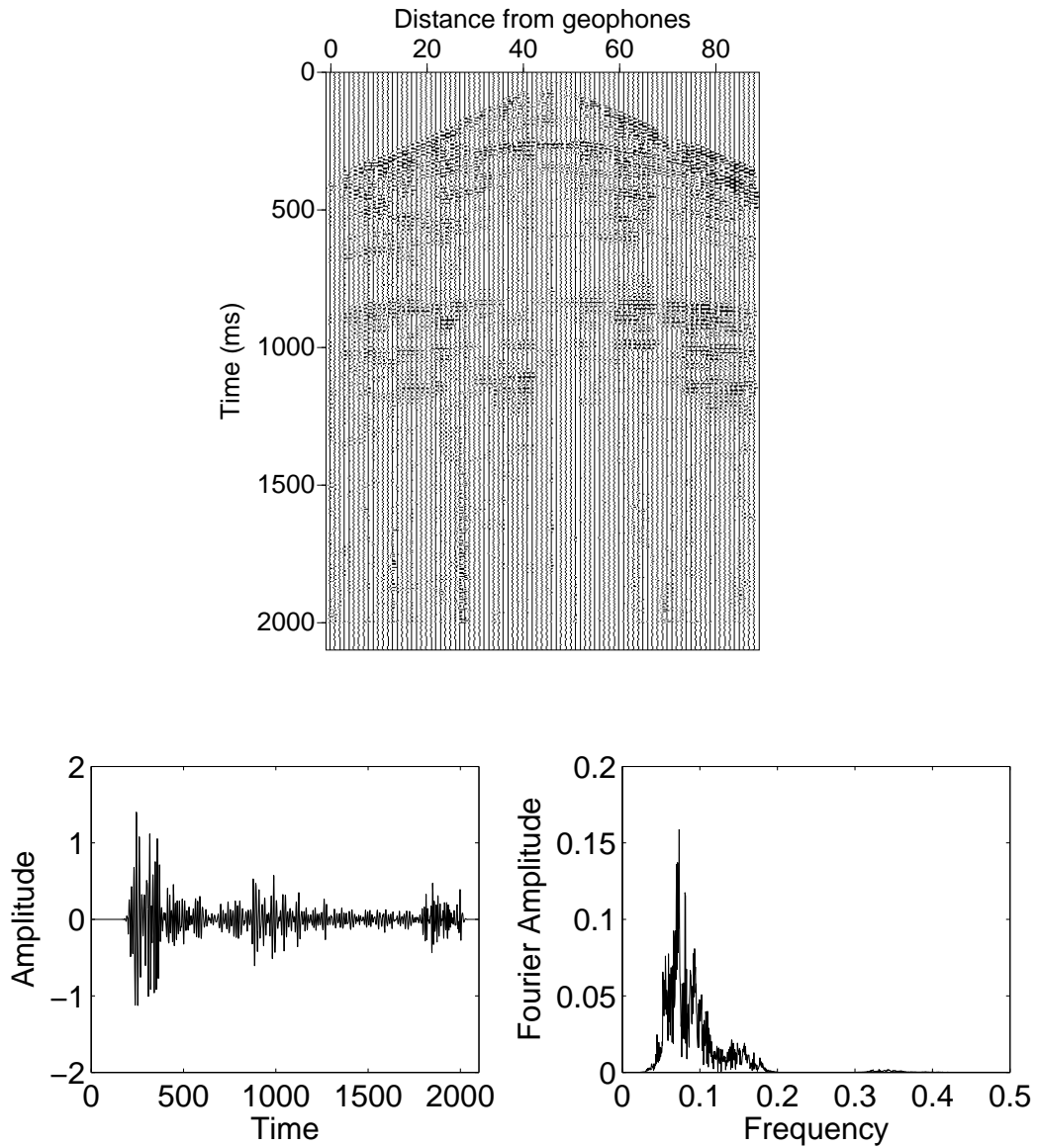


Figura 5.7: O padrão sísmico correspondendo ao índice intermediário de escala $j = 3$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta somente eventos de reflexões. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier.

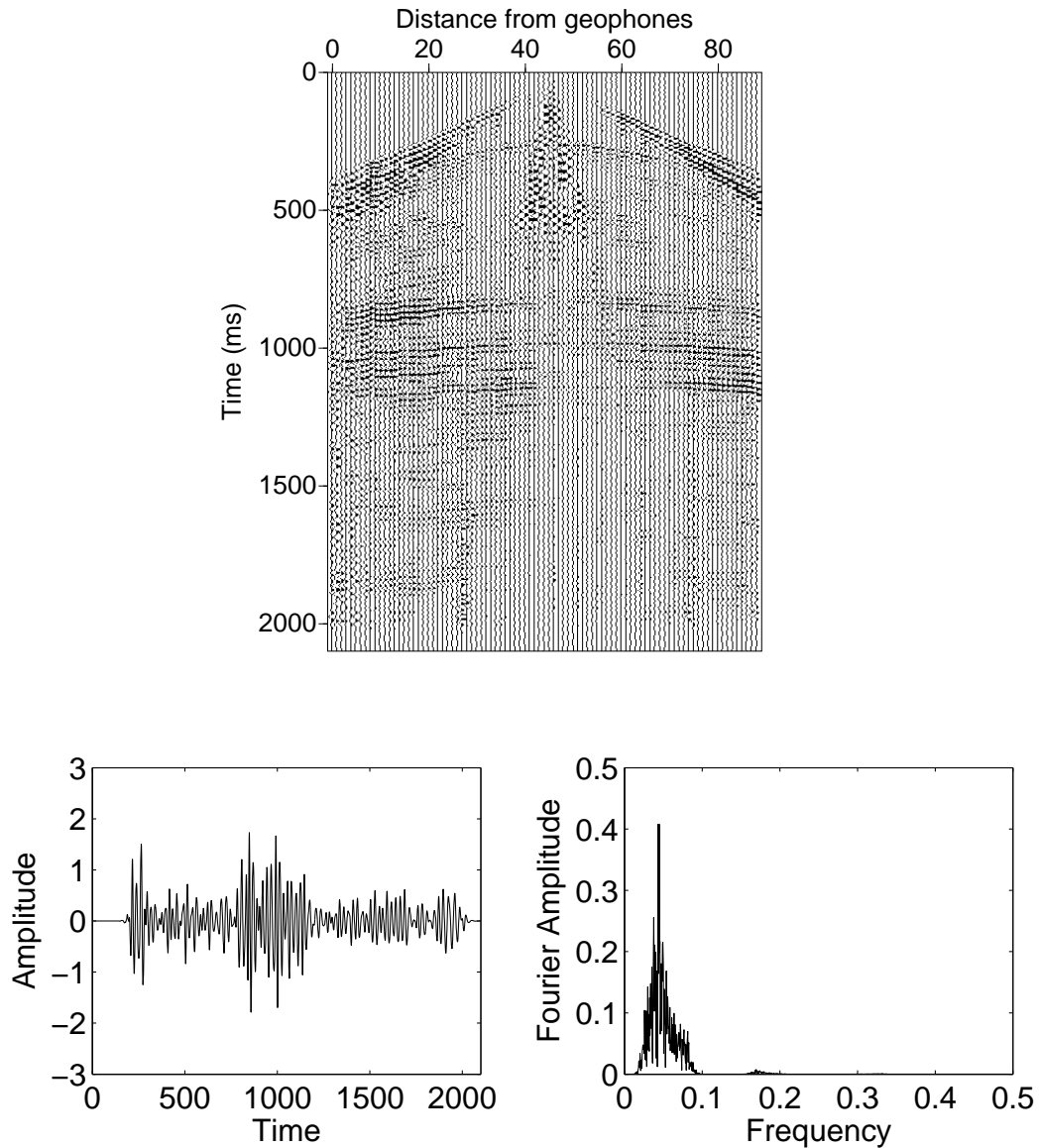


Figura 5.8: O padrão sísmico correspondendo ao índice intermediário de escala $j = 4$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta uma região onde o ruído *ground roll* começa aparecer. Porém as reflexões são dominantes na maior parte do sinal. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier.

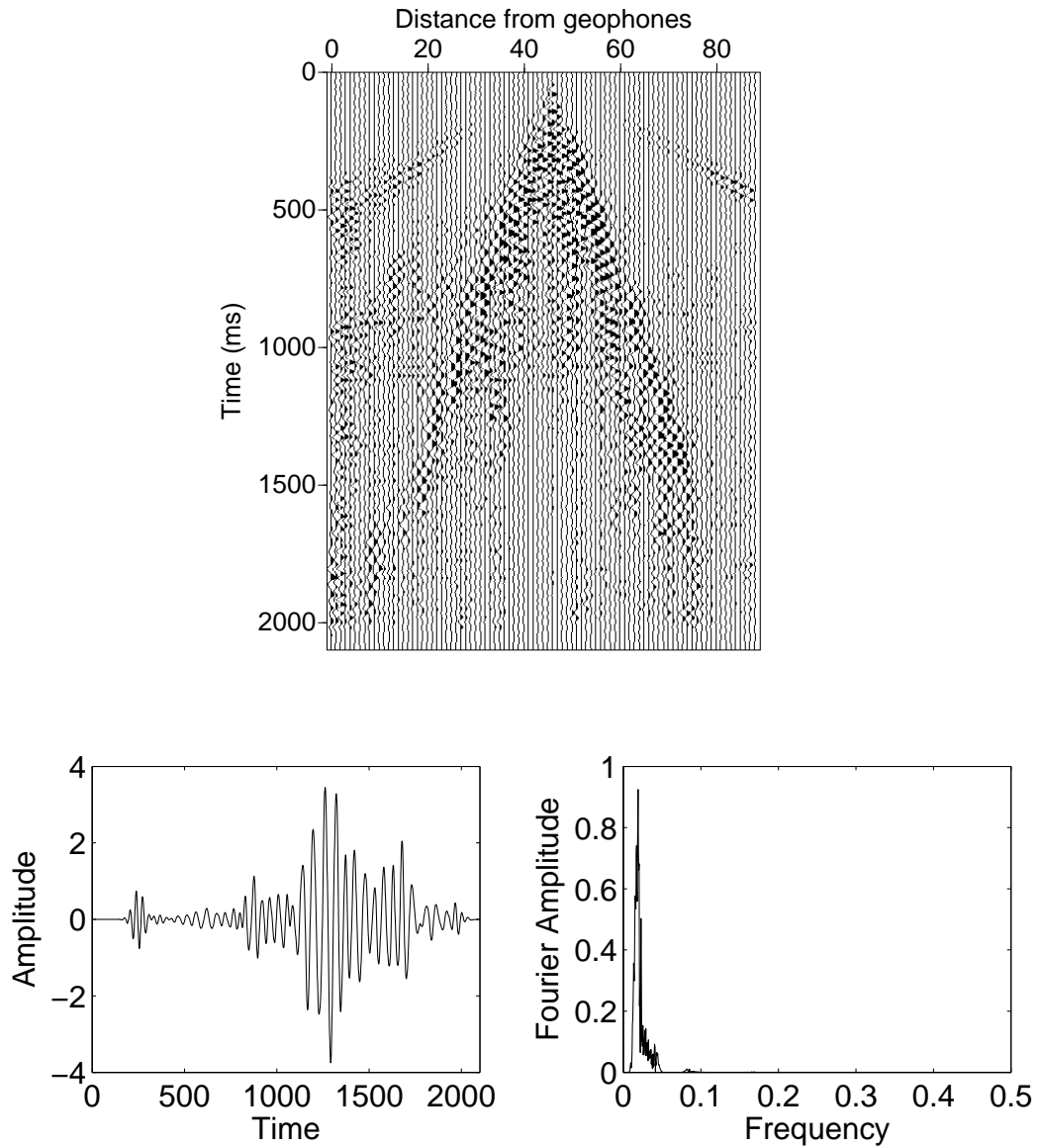


Figura 5.9: O padrão sísmico correspondendo ao índice intermediário de escala $j = 5$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta o ruído *ground roll* como o evento dominante e as reflexões quase invisíveis. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier.

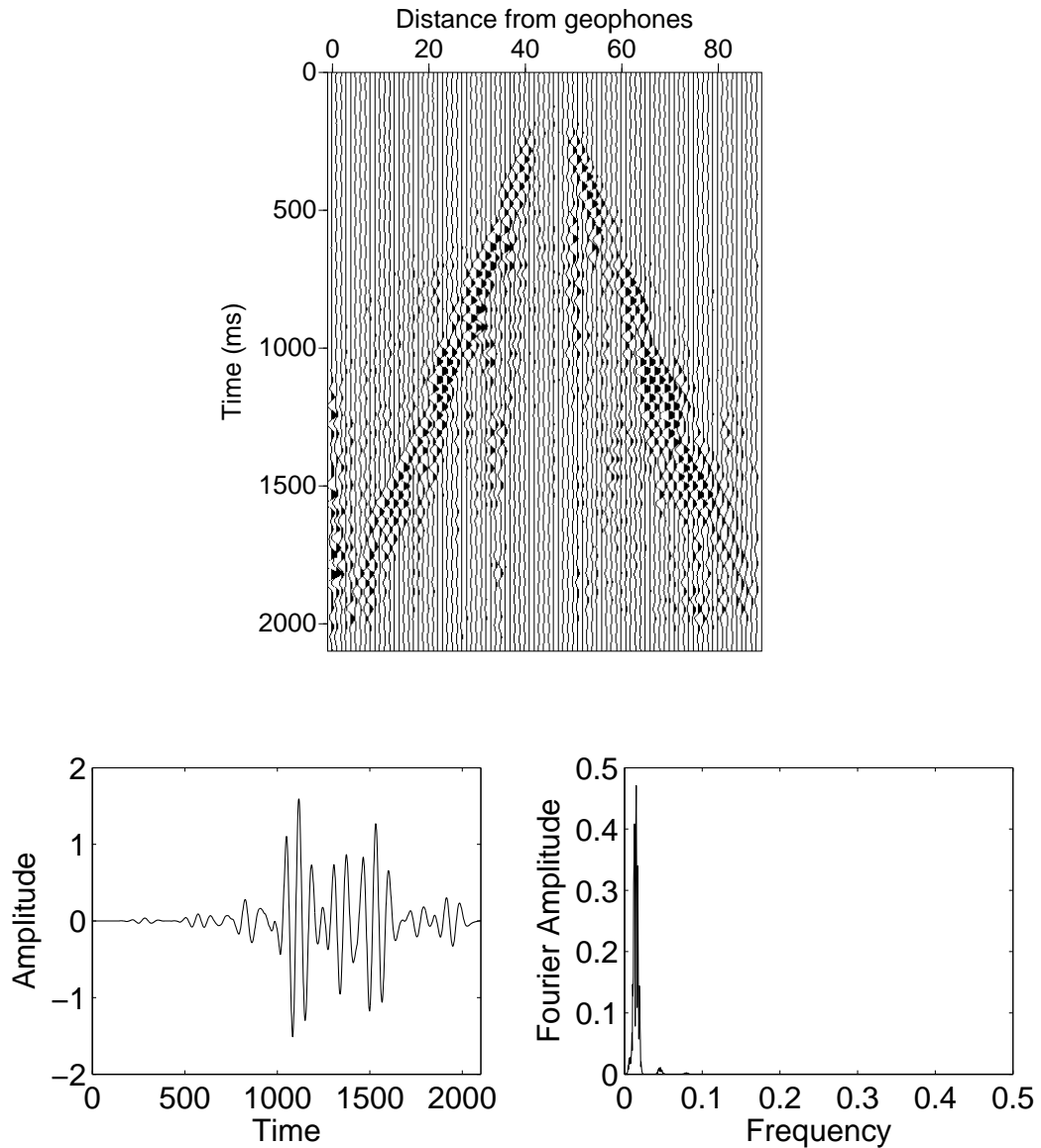


Figura 5.10: O padrão sísmico correspondendo ao índice intermediário de escala $j = 6$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta o ruído *ground roll* como único evento. As reflexões não existem. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier.

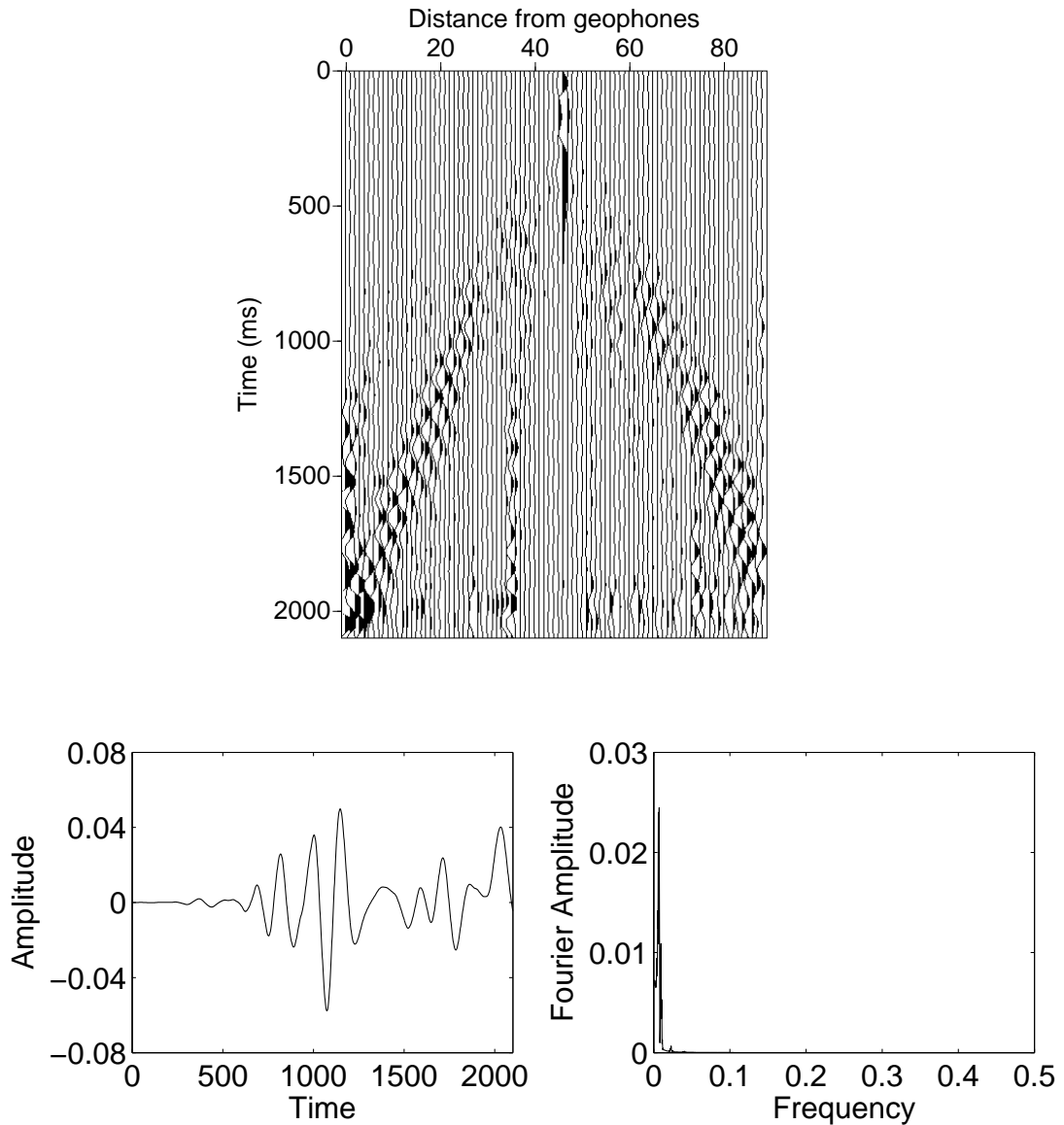


Figura 5.11: O padrão sísmico correspondendo aos índices de escalas $j = 7-11$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta o ruído *ground roll* como único evento. As reflexões não existem. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier.

Conclusões

O contexto principal desta tese se caracteriza pela busca de uma análise dos sinais sísmicos, os quais são bastante complexos, visando a remoção do ruído coerente *ground roll*. As técnicas tradicionais (Fourier) deixam a desejar e o método novo, aqui implementado, se baseia na utilização da Transformada em Ondaletas. A contribuição principal consiste em duas técnicas: (i) na Decomposição em Multiescala, a qual é apresentada como uma técnica robusta e (ii) a Transformada de Karhunen-Loève (decomposição KL) para selecionar o melhor fator de atenuação A_f .

Para melhorar a qualidade das reflexões, primeiro o registro sísmico foi submetido a uma decomposição em multiescala. Com esta decomposição foi capaz de localizar e identificar características do registro em uma dada escala, verificando se o *ground roll* está presente nesta escala. Uma vez que o ruído foi localizado numa região de uma escala (como na Figura (5.8)) este é atenuado somente nesta região e nessa única escala. Tanto as outras regiões do registro quanto as outras escalas não são afetados pela filtragem. Este procedimento é feito de forma cirúrgica, significando que os coeficientes que caracterizam uma certa região em tempo-espaco serão atenuados de A_f localmente. Esta operação preserva o máximo de energia

contida no sinal que carrega as informações de interesse (reflexões). Com este método o processo de filtragem é muito mais seletivo e preciso porque é realizado em escala por escala.

O método de decomposição em multiescala mostrou-se ser valioso na preservação das informações de interesse dentro do escopo da geofísica. Isto é particularmente pertinente a situações nas quais as características do *ground roll* podem ser bem discriminadas. Com este método obtivemos melhorias, especialmente para o alto grau de adaptabilidade para os sinais sísmicos, como também para a melhoria das propriedades de localização do filtro.

A comprovação do método, geralmente, é feita visualmente e depende muito do observador (interpretação visual humana). Porém, o espectro de frequência é mostrado para um único traço dos sismogramas original, atenuado, porção removida (ruído *ground roll*) e para os padrões sísmicos em escalas individuais. Percebe-se que não há uma constatação significativa por inspeção visual do sinal reconstruído e do sinal original. Isso significa que ambos apresentam conteúdo de frequências similares. Essa afirmação diz que o filtro não distorceu as informações de interesse, removendo assim, somente o ruído *ground roll*, que é o sinal indesejável. Ainda relacionado ao conteúdo de frequência, observamos que o espectro de frequência do ruído é formado, principalmente, por frequências baixas.

Apêndice A

Análise de Fourier

A.1 Introdução

Neste Apêndice nós resumimos as propriedades matemáticas relevantes que resulta da transformada de Fourier e também, a transformada de Fourier-Gabor a qual fornece uma análise “melhor” que a representação de Fourier. Nenhuma demonstração será apresentada, por isso o leitor deve consultar trabalhos originais neste domínio. A formulação é introduzida para funções $f(t) \in L^2(\mathbb{R})$ em uni-dimensional, não sendo necessário generalizá-las para o espaço bi-dimensional. Do ponto de vista prático, uma transformada de Fourier é uma integral de Fourier de uma função f , definida em \mathbb{R} . Como a sua informação espectral ocorre em termos de frequência, então o domínio de definição da transformada de Fourier, \hat{f} , definida em \mathbb{R} , é o domínio das frequências.

A importância da transformada de Fourier não vem apenas do significado físico de suas interpretações, tais como análise em tempo-frequência de sinais, também, pelo fato que as técnicas de análise de Fourier serem extremamente poderosas, como por exemplo, no estudo em análise de ondaletas.

A.1.1 Notação

Os conjuntos de números inteiros e reais serão representados por \mathbb{Z} e \mathbb{R} , respectivamente, e $L^p(\mathbb{R})$ é o espaço de Hilbert de funções $f(t)$ mensuráveis, tal que $\int_{-\infty}^{+\infty} |f(t)|^p dt < +\infty$. Para as funções $f(t)$ e $g(t)$ de quadrado integráveis em 1-dimensão, pertencentes ao $L^2(\mathbb{R})$, o produto interno de $f(t)$ com $g(t)$, no domínio do tempo, é definido por;

$$\langle f, g \rangle = \int_{-\infty}^{+\infty} f(t)g^*(t)dt, \quad (\text{A.1})$$

onde $g^*(t)$ significa o complexo conjugado de $g(t)$. A norma (energia) sobre o $L^2(\mathbb{R})$ da função $f(t)$ é induzida pela definição do produto interno;

$$\|f\|^2 = \int_{-\infty}^{+\infty} |f(t)|^2 dt. \quad (\text{A.2})$$

A convolução de duas funções $f(t) \in L^2(\mathbb{R})$ e $g(t) \in L^2(\mathbb{R})$ é definida por

$$f * g(t) = \int_{-\infty}^{+\infty} f(\tau)g(t - \tau)d\tau. \quad (\text{A.3})$$

Para qualquer função $f(t)$, a dilatação de $f(t)$ por um fator de escala σ é representada por

$$f_\sigma(t) = \frac{1}{\sigma} f\left(\frac{t}{\sigma}\right). \quad (\text{A.4})$$

A.2 Transformada de Fourier

A Transformada de Fourier decompõe a função $f(t)$ em uma combinação das funções harmônicas senos e co-senos. Essas funções harmônicas são combinadas na forma de uma exponencial complexa;

$$\begin{aligned} \varphi_\omega(t) &= \cos(\omega t) + i \operatorname{sen}(\omega t) \\ &= e^{i\omega t}. \end{aligned} \quad (\text{A.5})$$

A Transformada de Fourier da função $f(t)$ é dada pela integral

$$\begin{aligned} \hat{f}(\omega) &= \langle f(t), \varphi_\omega(t) \rangle \\ &= \int_{-\infty}^{+\infty} f(t) e^{-i\omega t} dt. \end{aligned} \quad (\text{A.6})$$

A função transformada $\hat{f}(\omega)$ também pertence a $L^2(\mathbb{R})$. Nesse espaço funcional é inversível e é dada por;

$$\begin{aligned}
 f(t) &= \langle \hat{f}(\omega), \varphi_t(\omega) \rangle \\
 &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega) e^{i\omega t} d\omega.
 \end{aligned} \tag{A.7}$$

Com a definição do produto interno, usando-se a Transformada de Fourier, conduz à relação de Parseval;

$$\langle f(t), g(t) \rangle = \int_{-\infty}^{+\infty} f(t) g^*(t) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega) \hat{g}^*(\omega) d\omega = \frac{1}{2\pi} \langle \hat{f}(\omega), \hat{g}(\omega) \rangle, \tag{A.8}$$

à qual tem como caso particular a relação de Plancherel

$$\langle f(t), f(t) \rangle = \int_{-\infty}^{+\infty} f(t) f^*(t) dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \hat{f}(\omega) \hat{f}^*(\omega) d\omega = \frac{1}{2\pi} \langle \hat{f}(\omega), \hat{f}(\omega) \rangle, \tag{A.9}$$

ou

$$\langle f, f \rangle = \int_{-\infty}^{+\infty} |f(t)|^2 dt = \frac{1}{2\pi} \int_{-\infty}^{+\infty} |\hat{f}(\omega)|^2 d\omega. \tag{A.10}$$

Da relação de Plancherel pode-se definir as quantidades Mallat (1999);

$$u = \frac{1}{\|f\|^2} \int_{-\infty}^{+\infty} t |f(t)|^2 dt = \langle t \rangle, \quad (\text{A.11})$$

$$\xi = \frac{1}{2\pi \|f\|^2} \int_{-\infty}^{+\infty} \omega |\hat{f}(\omega)|^2 d\omega = \langle \omega \rangle, \quad (\text{A.12})$$

$$\sigma_t^2 = \frac{1}{\|f\|^2} \int_{-\infty}^{+\infty} (t - u)^2 |f(t)|^2 dt, \quad (\text{A.13})$$

e

$$\sigma_\omega^2 = \frac{1}{2\pi \|f\|^2} \int_{-\infty}^{+\infty} (\omega - \xi)^2 |\hat{f}(\omega)|^2 d\omega. \quad (\text{A.14})$$

A variância temporal σ_t e a variância na frequência σ_ω da função $f(t)$ satisfazem;

$$\sigma_t^2 \sigma_\omega^2 \geq \frac{1}{4\pi}. \quad (\text{A.15})$$

A.2.1 Transformada de Fourier-Gabor

A Transformada de Fourier-Gabor Mallat (1999) ou *Short Time Fourier Transform* relaciona a função $f(t)$ com uma “função de análise” $g_{u,\xi}(t)$, conhecida como “átomo de Gabor”;

$$\tilde{f}'(u, \xi) = \int_{-\infty}^{+\infty} f(t)g_{u,\xi}^*(t)dt = \int_{-\infty}^{+\infty} f(t)g(t-u)e^{-i\xi t}dt, \quad (\text{A.16})$$

onde $g_{u,\xi}(t)$ são os “átomos” de Gabor e são construídos por translação no tempo;

$$g_{u,\xi}(t) = g(t-u)e^{i\xi t}. \quad (\text{A.17})$$

A energia de $g_{u,\xi}$ é concentrada na vizinhança de u sobre um intervalo de tamanho σ_t . Sua transformada de Fourier é uma translação por ξ ;

$$\hat{g}_{u,\xi}(\omega) = \hat{g}(\omega - \xi)e^{-iu(\omega - \xi)}. \quad (\text{A.18})$$

A Transformada de Fourier-Gabor é inversível e sua transformada inversa é dada por;

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \tilde{f}'(u, \xi)g(\xi - t)e^{i\xi u}dud\xi. \quad (\text{A.19})$$

A.3 Comentários

A equação (A.8) prova que o produto interno de f e g e a norma sobre o $L^2(\mathbb{R})$ são conservados pela transformada de Fourier a menos do fator constante 2π . Para o caso onde $f = g$ tem-se a relação (A.9) (ou A.10) que pode ser pensada como

a energia da função $f(t)$. Então, a energia da função f pode ser analisada em ambos os domínios. Na transformação do sinal do domínio do tempo para o domínio da frequência, a informação sobre a localização temporal é perdida.

As equações (A.11) e (A.12) indicam que a função $f(t) \in L^2(\mathbb{R})$ tem uma posição média no plano (t, ω) , dada pelas coordenadas $(\langle t \rangle, \langle \omega \rangle)$. A dispersão da energia em torno desta média é caracterizada pelos desvios padrões σ_t e σ_ω . Estes definem um retângulo, chamado “caixa de Heisenberg”, onde a maior parte da energia está concentrada. A localização da energia no tempo está relacionada com a localização da energia na frequência pelo *Princípio da Incerteza de Heisenberg*. Isto é, a concentração da energia no tempo e frequência está restrita ao *Princípio da Incerteza de Heisenberg*.

O Princípio da Incerteza de Heisenberg tem uma interpretação particular importante em Mecânica Quântica quanto a incerteza da posição e momento de uma partícula livre. O estado de uma partícula pode ser descrita por uma função de onda $f(t) \in L^2(\mathbb{R})$. Então, a equação (A.11) é a localização média da partícula e seu momento médio é dado pela equação (A.12). As variâncias desses valores médios são dadas pelas equações (A.13) e (A.14).

Na transformada de Fourier-Gabor a função de análise $g_{u,\xi}(t)$ é conhecida como “átomo de Gabor” e definida pela equação (A.16) por uma translação no tempo. A função $g_{u,\xi}(t)$ é delimitada por uma região, e por isso, também, é conhecida como ser uma “função janelada”. A energia de $g_{u,\xi}(t)$ está concentrada na vizinhança de u em um intervalo de tamanho σ_t . A energia de $\hat{g}_{u,\xi}(t)$ é localizada na frequência ξ em um intervalo de tamanho σ_ω . No plano tempo-frequência (t, ω) , o espalhamento da

energia do “átomo” $g_{u,\xi}$ é representado simbolicamente pela “caixa de Heisenberg”, como ilustrado na Figura (A.1). Esta “caixa” está centrada em (u, ξ) e tem uma largura σ_t no tempo e uma largura σ_ω na frequência. Apesar de fornecer informação no tempo-frequência, a Fourier-Gabor tem a desvantagem de usar um tamanho fixo da “janela” ($g_{u,\xi}$ tem o mesmo suporte) que irá percorrer sobre todo o sinal (mesma janela para todas as frequências).

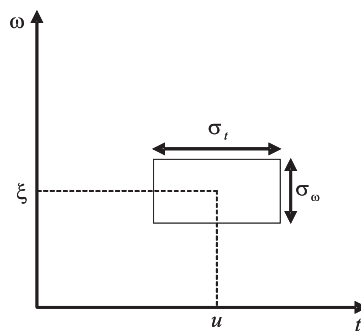


Figura A.1: Caixa de Heisenberg representando a localização tempo-frequência da energia para um “átomo de Gabor”.

Apêndice B

Propriedades dos Filtros h e g

B.1 O Filtro de aproximação h

É possível mostra Mallat (1989a,b) que a multiresolução está associada a uma função de escala $\phi(t)$ e o filtro h . A equação (1.11) pode ser expressa no domínio das frequências (transformada de Fourier de ambos os lados). Isto equivale à;

$$\hat{\phi}(2\omega) = \frac{1}{\sqrt{2}} \hat{h}(\omega) \hat{\phi}(\omega). \quad (\text{B.1})$$

com $\hat{h}(\omega) = \sum_{k=-\infty}^{+\infty} h[k] e^{-ik\omega}$. Na equação (1.11) a sequência $h[k] = \langle \frac{1}{\sqrt{2}} \phi(\frac{t}{2}), \phi(t-k) \rangle$ é interpretada como filtro discreto. Uma consequência direta da ortogonalidade de $\{\phi(t-k)\}_{k \in \mathbb{Z}}$ é que o filtro h satisfas as seguintes propriedades;

$$|\hat{h}(\omega)|^2 + |\hat{h}(\omega + \pi)|^2 = 2. \quad (\text{B.2})$$

e

$$\hat{h}(0) = \sqrt{2} \quad (\text{B.3})$$

que segue diretamente da propriedade 4 da multiresolução. A equação (B.1) pode ser iterada (no domínio das frequências), dando uma relação entre a função de escala ϕ e o filtro h . Devido as propriedades de h (equações B.2 e B.3), é possível mostrar;

$$\hat{\phi}(\omega) = \hat{\phi}(0) \prod_{m=1}^{+\infty} \frac{1}{\sqrt{2}} \hat{h}(2^{-m}\omega) \quad (\text{B.4})$$

que é sempre válida em uma multiresolução. O período de $\hat{h}(\omega)$ foi considerado ser 2π .

B.2 O Filtro de Detalhe g

A equação (1.12) representa a decomposição de $\psi(t/2) \in W_1 \subset V_0$ em uma base ortonormal $\{\phi(t - k)\}_{k \in \mathbb{Z}}$ de V_0 . A sua expressão no domínio das frequências (transformada de Fourier) dá;

$$\hat{\psi}(2\omega) = \frac{1}{\sqrt{2}} \hat{g}(\omega) \hat{\phi}(\omega). \quad (\text{B.5})$$

As condições que o filtro discreto g deve satisfazer para que $\{\psi(t - k)\}_{k \in \mathbb{Z}}$ seja uma base ortonormal de W_0 são;

$$|\hat{g}(\omega)|^2 + |\hat{g}(\omega + \pi)|^2 = 2 \quad (\text{B.6})$$

e

$$\hat{g}(\omega)\hat{h}^*(\omega) + \hat{g}(\omega + \pi)\hat{h}^*(\omega + \pi) = 0. \quad (\text{B.7})$$

Desde que, $|\hat{h}(\pi)| = |\hat{h}(-\pi)| = 0$ e $|\hat{h}(0)| = |\hat{h}(2\pi)| = \sqrt{2}$, segue da equação (B.7) que $|\hat{g}(0)| = |\hat{g}(2\pi)| = 0$ e $|\hat{g}(\pi)| = |\hat{g}(-\pi)| = \sqrt{2}$. A função que satisfaz a equação as equações (B.6) e (B.7) devido a equação (B.8) é

$$\hat{g}(\omega) = e^{-i\omega}\hat{h}^*(\omega + \pi). \quad (\text{B.8})$$

Fazendo-se a transformada inversa de Fourier da equação (B.8)

$$g[k] = (-1)^{1-k}h[1 - k]. \quad (\text{B.9})$$

Uma vez que os coeficientes $h[n]$ são conhecidos, então é possível encontrar $g[n]$ com a equação (B.9). Da equação (B.9) vemos que os coeficientes de filtro de g são completamente caracterizados pelos coeficientes de h . Então é a função de escala que determina a análise multirresolução.

Apêndice C

Funções Ondaletas

C.1 Introdução

As ondaletas exigem certas condições de existência com uma fundamentação matemática bastante flexível. Diferentemente da análise de Fourier, onde as funções da base são senoidais, a transformada em ondaletas tem como funções de base “ondas pequenas” (onda pequena vem do termo “wavelet” que pode ter essa tradução). Esse conceito tem característica de oscilação (variação de frequência) e limitação no tempo (energia concentrada), permitindo uma análise em tempo-frequência.

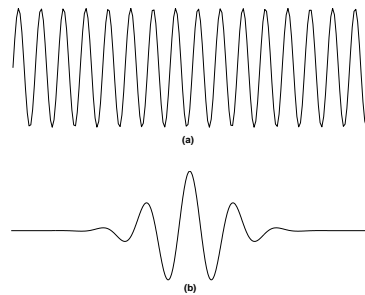


Figura C.1: (a) representação de uma onda senoidal e (b) a ondaleta de Morlet.

A noção de uma ondaleta (oscilatória e limitada no tempo) já pode ser percebida pela representação da figura acima. A Figura (a) está ilustrando uma onda senoidal (oscilatória e ilimitada no tempo) com amplitude uniforme sobre $-\infty \leq t \leq +\infty$ (e, portanto tendo energia infinita) e (b) uma ondaleta possuindo energia bem limitada em uma região do espaço.

C.2 Exemplo de ondaletas

Devido a flexibilidade de escolha, muitas ondaletas tem sido utilizadas como bases. Aqui é exemplificada algumas destas ondaletas.

C.2.1 Ondaletas derivadas da Gaussiana

Essas funções são formadas das derivadas da função gaussiana $g(t) = \frac{1}{\pi^{1/4}\sqrt{3\sigma}} e^{-\frac{t^2}{2\sigma^2}}$ e pertence a classe de ondaletas não ortogonais. A figura abaixo mostra as ondaletas geradas pelas derivadas da gaussiana de ordem 2 – 5, respectivamente, com $\sigma = 1$.

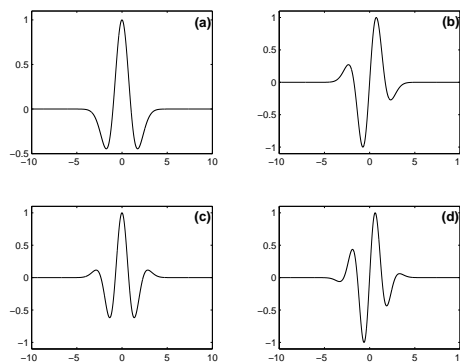


Figura C.2: Ondaletas geradas das derivadas da função gaussiana. Derivadas de ordem (a) segunda, (b) terceira, (c) quarta e (d) quinta.

C.2.2 Ondaletas de Morlet

A ondaleta de Morlet é complexa e definida pela equação $\psi(t) = \pi^{-1/4}(e^{i\omega_0 t} - e^{-\frac{\omega_0^2}{2}})e^{-t^2/2}$. A figura abaixo mostra essa ondaleta para $\omega_0 = 5$.

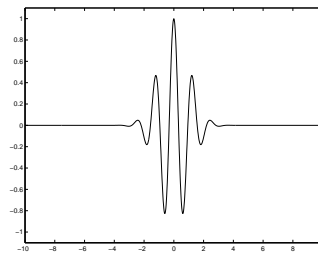


Figura C.3: Representação da parte real da ondaleta de Morlet.

C.2.3 Ondaletas de Shannon

Essa ondaleta não tem suporte compacto no tempo, porém possui um decaimento assintótico lento no tempo. Sua expressão analítica é $\psi(t) = \frac{\sin(\frac{\pi t}{2})}{\frac{\pi t}{2}} \cos(\frac{3\pi t}{2})$

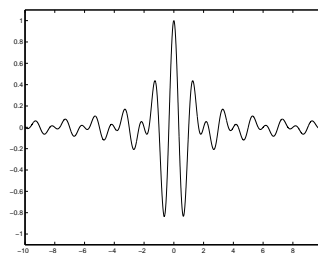


Figura C.4: Representação da ondaleta de Shannon.

C.2.4 Ondaletas de Daubechies

As ondaletas construídas por Daubechies formam bases completas para o $L^2(\mathbb{R})$ e são organizadas em famílias com o parâmetro de identificação N . A função de escala para uma determinada família é,

$$\phi_N(t) = \sum_{k=0}^{2N-1} h[k] \sqrt{2} \phi_N(2t - k), \quad (\text{C.1})$$

A função ondaleta $\psi(t)$ é definida por;

$$\psi_N(t) = \sum_{k=0}^{2N-1} g[k] \sqrt{2} \phi_N(2t - k) \quad (\text{C.2})$$

A função $\psi(t)$ tem N momentos nulos se

$$\int_{-\infty}^{+\infty} t^p \psi(t) dt = 0, \quad \text{com } 0 \leq p < N. \quad (\text{C.3})$$

C.2.5 Coeficientes $h[k]$ das ondaletas de Daubechies

Propriedades da função de escala e da função ondaleta Daubechies (1992);

Normalização da função de escala

$$\int_{-\infty}^{+\infty} \phi(t) dt = 1. \quad (\text{C.4})$$

Usando-se a relação (C.1) chega-se a um sistema da forma

$$\sum_{k=0}^{2N-1} h[k] = \sqrt{2}. \quad (\text{C.5})$$

Ortogonalidade da função de escala

$$\int_{-\infty}^{+\infty} \phi(t)\phi(t-k)dt = \delta_{0,k}. \quad (\text{C.6})$$

Usando-se a relação (C.1) chega-se a para os coeficientes

$$\sum_{n=0}^{2N-1-2k} h[n]h[n+2k] = 0 \quad \text{com } k = 1, \dots, N-1.. \quad (\text{C.7})$$

Ortogonalidade da ondaleta com a função de escala

$$\int_{-\infty}^{+\infty} \psi(t)\phi(t-k)dt = 0. \quad (\text{C.8})$$

Usando as relações (C.1) e (C.2) chega-se a uma combinação dos coeficientes $g[k]$ com $h[k]$ pela seguinte relação

$$g[k] = (-1)^{1-k}h[1-k], \quad (\text{C.9})$$

garantindo a ortogonalidade entre $\psi_N(t)$ e $\phi_N(t)$.

Ortogonalidade da função ondaleta à polinômio

$$\int_{-\infty}^{+\infty} t^\alpha \psi(t)dt = 0, \quad \text{com } \alpha = 0, 1, \dots, N-1. \quad (\text{C.10})$$

Usando-se a relação (C.2) chega-se a um sistema da forma

$$\sum_{k=0}^{2N-1} (-1)^k h[k] k^\alpha = 0 \quad \text{com } \alpha = 0, 1, \dots, N-1.. \quad (\text{C.11})$$

Com a condição (C.5) é permitido escrever uma equação

$$h_0 + h_1 + h_2 + h_3 + h_4 + h_5 + h_6 + h_7 = \sqrt{2} \quad (\text{C.12})$$

Com a condição (C.7) é permitido escrever $N-1$ equações

$$\begin{aligned} h_0 h_2 + h_1 h_3 + h_2 h_4 + h_3 h_5 + h_4 h_6 + h_5 h_7 &= 0 \\ h_0 h_4 + h_1 h_5 + h_2 h_6 + h_3 h_7 &= 0 \\ h_0 h_6 + h_1 h_7 &= 0 \end{aligned} \quad (\text{C.13})$$

Com a condição (C.11) é permitido escrever N equações

$$\begin{aligned} -h_1 + h_2 - h_3 + h_4 - h_5 + h_6 - h_7 &= 0 \\ -h_1 + 2h_2 - 3h_3 + 4h_4 - 5h_5 + 6h_6 - 7h_7 &= 0 \\ -h_1 + 4h_2 - 9h_3 + 16h_4 - 25h_5 + 36h_6 - 49h_7 &= 0 \\ -h_1 + 8h_2 - 27h_3 + 64h_4 - 125h_5 + 216h_6 - 343h_7 &= 0 \end{aligned} \quad (\text{C.14})$$

Resolvendo-se o sistema formado pelas equações (C.12), (C.13) e (C.14) obtém-se os coeficientes h 's. Os coeficientes h 's estão escritos na Tabela (C.1). Esses coeficientes (filtros) definem a ondaleta ψ_4 da família de ondaletas ortonormais de Daubechies

Tabela C.1: Coeficientes de Filtros de Daubechies para ondaleta com $N = 4$ momentos nulos.

N	k	$h[k]$
4	0	0.230377781330889
	1	0.714846570552912
	2	0.630880767929853
	3	-0.027983769416859
	4	-0.187034811719093
	5	0.030841381835560
	6	0.032883011666885
	7	-0.010597401785069

com suporte compacto finito. O suporte dessas ondaletas é de $2N-1$. Para ondaletas de ordem maiores ($N' > N$), os coeficientes podem ser obtidos de maneira análoga. Para $N = 1$ tem-se que as ondaletas de Daubechies recaem nas ondaletas de Haar, que podem ser encaradas como sendo as ondaletas de menor grau e a regularidade destas ondaletas aumentam à medida que o valor de N cresce. Na Figura (C.5) estão ilustradas as funções ondaletas de Daubechies para $N=1,2,3,4,6,8$, respectivamente.

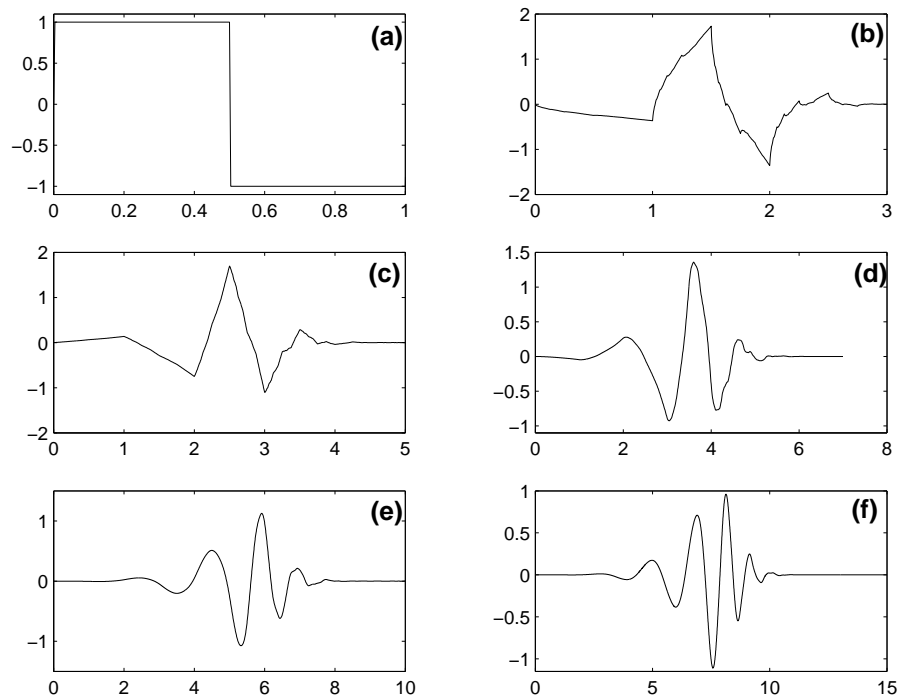


Figura C.5: Representação das famílias de ondaletas de Daubechies para os correspondentes $N = 1, 2, 3, 4, 6, 8$.

Lista de Figuras

1.1	Ondaleta normalizada ψ_1 e suas dilatações/translações para valores diferentes de σ e τ . A ondaleta ψ_1 é a derivada segunda da gaussiana. Tem-se $\psi_3 = \psi_{\sigma_0/4, -\tau}, \psi_1 = \psi_{\sigma_0, \tau_0=0}, \psi_2 = \psi_{\sigma_0/2, \tau}$. e (b) seus espectros de frequências.	16
1.2	Esquema ilustrativo da representação do particionamento do plano <i>Tempo</i> \times <i>Frequência</i> para as transformadas de (topo) Fourier-Gabor, (meio) ondaletas e (baixo) discretização no domínio tempo-escala para as ondaletas.	21
1.3	Representação da decomposição de f_N nos subespaços V_j e W_j . Em (a) tem-se a representação do subespaço V_0 e em (b) a representação da decomposição na soma direta $V_0 = W_1 \oplus W_2 \oplus W_3 \oplus V_3$	30
1.4	Representação do algoritmo de decomposição rápida com os filtros h e g seguida por uma decimação de dois ($\downarrow 2$).	31
1.5	Representação do algoritmo de reconstrução rápida com os filtros \tilde{h} e \tilde{g} seguida por uma inserção de dois ($\uparrow 2$).	31

- 3.1 Representação da aquisição para a obtenção de dados terrestres e marítimos. Esse processo envolve (caso terrestre) a geração de uma onda sísmica e a sua captação na superfície por um arranjo de receptores (geofones (a)). A fonte de energia mais utilizada na geração das ondas são as fontes explosivas e vibratória. 44
- 3.2 Representação para os traços sísmicos. (a) traço sísmico em seu formato convencional (série temporal) onde na horizontal tem-se as amplitudes e na vertical a quantidade de amostras (temporal). Esse traço é o de número 20 do sismograma representado na Figura (3.3). Em (b) tem-se a representação para os traços 20 a 22 (para o mesmo sismograma) em seu formato *wiggle* o qual é muito utilizado para uma visualização adequada. 46
- 3.3 Registro sísmico 2D terrestre em representação do tipo *wiggle* com ganho automático [Stockwell (2000)]. O eixo vertical está definido em unidades de tempo (ms). Na horizontal representa o afastamento entre os geofones (m) e cada geofone representa um traço. Em (a) tem-se o local onde ocorreu a geração das ondas sísmicas (fonte). Pode-se observar que a identificação das hipérbolas de reflexões (c) ficam prejudicadas pelo “cone” de ruído *ground roll*. 48

- 4.1 (a) O sismograma original formado por 96 traços (horizontal) e 1001 amostra (vertical) por traço. O *ground roll* aparece (bem visível) em uma região que vai do canto superior esquerdo ao canto inferior direito. Em (b) temos o sismograma atenuado pelo fator ótimo de $A_f = 0.80$. Em (c) temos a porção que foi removida do sismograma original. A região atenuada está nos limites (horizontal) dos traços 12–90 e (vertical) amostras de 125–230 (traço 12) e 470–780 (traço 90). 61
- 4.2 A energia relativa \tilde{E} dos 5 primeiros modos KL da região atenuada mostrada como uma função do fator de atenuação A_f . O mínimo dos 2 primeiros modos caracterizam o fator de atenuação ótimo. 65
- 4.3 A distribuição de energia KL em função dos modos antes e depois de aplicar o filtro ondaleta ótimo com fator de atenuação $A_f = 0.80$. A linha sólida (tracejada) corresponde ao espectro de energia da região original (atenuada). No gráfico em detalhe (*inset*) os 10 primeiros modos são mostrados com barras (*star*) indicando os espectros da região atenuada (original). 66
- 5.1 (a) O sismograma original onde o ruído *ground roll* está presente e atrapalhando as reflexões. Ele é composto por 90 traços e 2048 amostras por traço. Em (b) temos o sismograma atenuado. Agora as reflexões são dominantes. Em (c) temos a porção que foi removida do sismograma original. 71

- 5.2 Representação do método de decomposição rápida em ondaletas para o registro sísmico representado na Figura 5.1(a). Supõe-se que o traço sísmico inicial está representado em V_0 com n amostras e o processo ocorre da interação com os filtro h e g seguido por uma decimação de fator 2. 74
- 5.3 Representação do método de reconstrução rápida em ondaletas. O processo começa com os coeficientes a_{11} , d_{11} e ocorre a interação com os filtros \tilde{h} e \tilde{g} seguidos por uma inserção de fator 2. 74
- 5.4 Um único traço representado em seu formato convencional (série temporal) e seu espectro de Fourier. (topo) temos o traço 65 do sismograma original. (meio) temos o mesmo traço para o sismograma atenuado e (baixo) o traço da porção que foi removida do sismograma original. 81
- 5.5 O padrão sísmico correspondendo ao menor índice de escala $j = 1$. Este registro contém as mais altas componentes de frequências e o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier. . . . 82
- 5.6 O padrão sísmico correspondendo ao índice intermediário de escala $j = 2$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta somente eventos de reflexões. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier. . . . 83

- 5.7 O padrão sísmico correspondendo ao índice intermediário de escala $j = 3$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta somente eventos de reflexões. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier. 84
- 5.8 O padrão sísmico correspondendo ao índice intermediário de escala $j = 4$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta uma região onde o ruído *ground roll* começa aparecer. Porém as reflexões são dominantes na maior parte do sinal. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier. 85
- 5.9 O padrão sísmico correspondendo ao índice intermediário de escala $j = 5$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta o ruído *ground roll* como o evento dominante e as reflexões quase invisíveis. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier. 86

- 5.10 O padrão sísmico correspondendo ao índice intermediário de escala $j = 6$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta o ruído *ground roll* como único evento. As reflexões não existem. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier. 87
- 5.11 O padrão sísmico correspondendo aos índices de escalas $j = 7 - 11$. Este registro tem o mesmo padrão (90 traços e 2048 amostras por traço) que o registro original e apresenta o ruído *ground roll* como único evento. As reflexões não existem. (baixo) O traço 65 (para este registro) representado em formato convencional (série temporal) e seu espectro de Fourier. 88
- A.1 Caixa de Heisenberg representando a localização tempo-frequência da energia para um “átomo de Gabor”. 98
- C.1 (a) representação de uma onda senoidal e (b) a ondaleta de Morlet. . 102
- C.2 Ondaletas geradas das derivadas da função gaussiana. Derivadas de ordem (a) segunda, (b) terceira, (c) quarta e (d) quinta. 103
- C.3 Representação da parte real da ondaleta de Morlet. 104
- C.4 Representação da ondaleta de Shannon. 104
- C.5 Representação das famílias de ondaletas de Daubechies para os correspondentes $N = 1, 2, 3, 4, 6, 8$ 109

Referências Bibliográficas

Andrews, H., Hunt, B.,(1977); *Digital Image Restoration*. Prentice Hall.

Arneodo, A., Muzy, J. F., Bacry, E.(1994); *The Multifractal Formalism Revisited with Wavelets*. International Journal of Bifurcation and Chaos **4**(2), 245 – 302.

Arneodo, A., Muzy, J. F., Bacry, E.(1995); *The thermodynamics of fractals Revisited with Wavelets*. Physica A **213**, 232 – 275.

Aubry, N., Guyonnet, R., Lima, R., (1991); *Spatiotemporal analysis of complex signals: theory and applications*. Journal of Statistical Physics. **64**, 683 – 739.

Battle, G. (1987); *A block spin construction of ondelettes. Part I: Lemarié functions*. Comm. Math. Phys., **110**, 601 – 615.

Coffeen, J. A. (1978); *Seismic Exploration Fundamentals*, Pennwell Publishing Company.

Corso, G., Kuhn, P. S., Lucena, L. S., Thomé, Z. D. (2003); *Seismic Ground Roll Time-Frequency Filtering using the Gaussian Wavelet Transform*. Physica A, **318**, 551 – 561.

- Daubechies, I. (1988); *Orthonormal Bases of Compactly Supported Wavelets*. Communications on Pure and Applied Mathematics, **41**, 909 – 996.
- Daubechies, I. (1990); *The wavelet transform, time-frequency localization and signal analysis*. IEEE Transactions on Information Theory, Vol. **36**, pp. 961 – 1005.
- Daubechies, I. (1992); *Ten Lectures on Wavelets*. Society for Industrial and Applied Mathematics. Philadelphia.
- Daubechies, I. (1996); *Where Do Wavelets Come From? A Personal Point of View*. Proceedings of the IEEE, Special Issue on Wavelets **84**(4), 510 – 513.
- Deighan, A. J., Watts, D. R. (1997); *Ground-roll supression using the wavelet transform*. Geophysics, **62**(6), 1896 – 1903.
- Droujinine, A. (2006); *Multi-scale geophysical data analysis using the eigenimage discrete wavelet transform*. Journal of Geophysics and Engineering, **3**, 59 – 81.
- Gabor, D. (1946); *Theory of Communication*. Journal of the Institute of Electrical Engineering **93**(3) 429 – 457.
- Galvão, R. K. H., Araújo, M. C. U., Pimentel, M. F. (2001); *Estudo comparativo Sobre Filtragem de Sinais Usando Transformadas de Fourier e Wavelet*. Q. Nova, **24**, n^o6, 874 – 884.
- Gerhardt, A. L. B. (1998); *Aspectos da Visualização Volumétrica de Dados Sísmicos*. Dissertação de Mestrado, PUC-Rio.

Gottlieb, D.; Shu, C. W. (1997); *Siam Review*, **39**, 644.

Grossmann, A., Morlet, J. (1984); *Decomposition of Hardy Functions into Square Integrable wavelets of Constant Shape*. SIAM Journal on Mathematical Analysis, **15**(4), 723 – 736.

Grossmann, A., Kronland-Martinet, R., and Morlet, J., 1989. *Reading and understanding continuous wavelets transforms, in Wavelets: Time-Frequency Methods and Phase Space*, edited by J. Combes, A. Grossmann and P. Tchamitchian, pp. 2 – 20, Springer-Verlag, New York.

Haar, A. (1910); *Zür Theorie der Orthogonalen Funktionensysteme*. Mathematische Annalen, **69**, 331 – 371.

Hatton, L., Worthington, M. H., Makin, J. (1986); *Seismic Data Processing: Theory and Practice*, Blackwell Scientific Publications, 1^a edition.

Herrmann, F. J., Oct. 1998. *Multiscale analysis of well and seismic data*. In: Hassanzadeh, S. (Ed.), Proc. SPIE Vol. 3453, p. 180-208, Mathematical Methods in Geophysical Imaging V, Siamak Hassanzadeh; Ed. Vol. 3453 of Presented at the Society of Photo-Optical Instrumentation Engineers (SPIE) Conference. pp. 180 – 208.

Hoekstra, E. V. (1996); *Multiscale Analysis of Seismic Data by The Wavelet Transform*. M.Sc. Thesis, Delft University of Technology, July.

- Holmes, P., Lumley, L., and Berkooz, G. (1996); *Turbulence, Coherent Structures, Dynamical Systems and Symmetry*. Cambridge University Press, Cambridge.
- Jolliffe, I. (2002); *Principal Component Analysis*. Springer Series in Statistics, (Springer New York, 2nd ed.
- Kobayashi, M. (2001); *Wavelets and their Applications in Industry*. Nonlinear Analysis, **47**, 1749 – 1760.
- Lemarié, P. G. (1988); *Une nouvelle basis d'ondelettes de $L^2(\mathbb{R})$* . Journal de Mathématiques Pures et Appliquées. **67**, 227 – 236.
- Leite, F. E. A., Montagne, R., Corso, G., Vasconcelos, G., Lucena, L. S., (2008). *Optimal wavelet filter for suppression of coherent noise with an application to seismic data*. Physica A, **387**, 1439 – 1445.
- Leite, F. E. A., Tavares, D. M., Lucena, L. S., (2007b). *Efficient selective filtering of seismic data using multiresolution decomposition*. Physica A. Submitted.
- López, L. C., Londoño, E. G., Kazmierczak, T. S., 2005. *Using the Karhunen-Loève transform to suppress ground roll in seismic data*. Earth Sciences Research Journal. Vol. **9**, n° 2, 139 – 147.
- Mallat, S. (1989); *A Theory for Multiresolution Signal Decomposition: The Wavelet Representation*. IEEE Transactions on Pattern Analysis and Machine Intelligence. **11**(7), 674 – 693.

- Mallat, S. (1989); *Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$* . Transactions of American Mathematical Society, **315**(1), 69 – 87.
- Mallat, S. (1989); *Multifrequency channel decomposition of images and wavelet models*. IEEE Transaction in Acoustic, Speech and Signal Processing, **37**, 2091 – 2110.
- Mallat, S. and Hwang, W. L. (1992); *Singularity detection and processing with wavelets*. IEEE Transaction in Information Theory, **38**(2), 617 – 643.
- Mallat, S. and Zhang, Z. (1993); *Matching pursuit in a time-frequency dictionary*. IEEE Transactions on Signal Processing, 3397 – 3415.
- Mallat, S. (1999); *A Wavelet Tour of Signal Processing*. Academic Press, New York.
- Meyer, Y. (1986); *Principe d'incertitude, bases Hilbertiennes et algèbres d'opérateurs*. Séminaire Bourbaki. **662**, n°662, 1985 – 1986.
- Meyer, Y. (1990); *Ondelettes*. Hermann. New York.
- Meyer, I. (1992); *Wavelets and Operators*. Advanced mathematics. Cambridge University Press.
- Meyer, Y. (1993); *Wavelets: Algorithms and Applications*. Society for Industrial and Applied Mathematics. Philadelphia. SIAM.
- Miao, X. G., Moon, W. M. (1999); *Application of wavelet transform in reflection seismic data analysis*. Geosciences Journal, vol.3(**3**), 171 – 179.

- Montagne, R., Vasconcelos, G. (2006a); *An optimized filter for seismic data using the Karhunen-Loève transform and a minimum-energy criterium*. Physical Review E, **74**, 01016213.
- Montagne, R., Vasconcelos, G. (2006b); *Thermodynamic criteria for optimal suppression of coherent noise in seismic data using the Karhunen-Loève transform*. Physica A, **371**, pp. 122 – 125.
- Morlet, J., Arens, G., Fourgeau, E., Giard, D. (1982); *Wave propagation and sampling theory; Complex signal and scattering in multilayered media*. Geophysics, **47**(2), 203 – 221.
- Preisendorfer, R. W., (1988); *Principal component analyses in Meteorology and Oceanography*. Elsevier.
- Qian, S.; Chen, D. (1996); *Joint Time-Frequency Analysis-Methods and Applications*. Prentice Hall PTR; Upper Saddle River.
- Raz, S. (1987); *Beam Stacking: A Generalized Processing Technique*. Geophysics, Vol. **52**, pp. 1199 – 1210.
- Richart, F. E., Hall, J. R., Woods, R. D. (1970); *Vibration of Soils and Foundations*. Prentice-Hall.
- Sirovich, L. (1987); *Turbulence and the dynamics of coherent structures*. Quarterly of Applied Mathematics. **45**, 561 – 590.

- Souza, J. R. M. (2002); *Compressão de Dados Sísmicos Utilizando a Transformada Wavelet*. Dissertação de Mestrado, UFRN, Setembro.
- Stockwell, J. (2000); *Seismic Unix Release 34*. Center for Wave Phenomena, Colorado School of Mines. [<http://www.cwp.mines.edu/cwpcodes/index.html>].
- Strang, G., Nguyen, T. (1996); *Wavelets and Filter Banks*. Wellesley Cambridge Press, Wellesley.
- Thiebaux, H. J., (1994); *Statistical analyses for ocean and atmospheric sciences*. Academic Press.
- Thomas, J. E. (2001); *Fundamentos de Engenharia de Petróleo*. Rio de Janeiro; Interciência; Petrobrás.
- Wallace, J. M., Smith, C., Bretherton, C. S. (1992); *Singular value decomposition of wintertime sea surface temperature*. Journal of Climate, **5**(6), 561 – 576.
- Walker, J. S. (1997); *Fourier Analysis and Wavelet Analysis*. Notices of the AMS, **44**, n^o6, 658 – 670.
- Yilmaz, O. (1987); *Seismic Data Processing*. Society of Exploration Geophysicists, Tulsa (USA).

Livros Grátis

(<http://www.livrosgratis.com.br>)

Milhares de Livros para Download:

[Baixar livros de Administração](#)

[Baixar livros de Agronomia](#)

[Baixar livros de Arquitetura](#)

[Baixar livros de Artes](#)

[Baixar livros de Astronomia](#)

[Baixar livros de Biologia Geral](#)

[Baixar livros de Ciência da Computação](#)

[Baixar livros de Ciência da Informação](#)

[Baixar livros de Ciência Política](#)

[Baixar livros de Ciências da Saúde](#)

[Baixar livros de Comunicação](#)

[Baixar livros do Conselho Nacional de Educação - CNE](#)

[Baixar livros de Defesa civil](#)

[Baixar livros de Direito](#)

[Baixar livros de Direitos humanos](#)

[Baixar livros de Economia](#)

[Baixar livros de Economia Doméstica](#)

[Baixar livros de Educação](#)

[Baixar livros de Educação - Trânsito](#)

[Baixar livros de Educação Física](#)

[Baixar livros de Engenharia Aeroespacial](#)

[Baixar livros de Farmácia](#)

[Baixar livros de Filosofia](#)

[Baixar livros de Física](#)

[Baixar livros de Geociências](#)

[Baixar livros de Geografia](#)

[Baixar livros de História](#)

[Baixar livros de Línguas](#)

[Baixar livros de Literatura](#)
[Baixar livros de Literatura de Cordel](#)
[Baixar livros de Literatura Infantil](#)
[Baixar livros de Matemática](#)
[Baixar livros de Medicina](#)
[Baixar livros de Medicina Veterinária](#)
[Baixar livros de Meio Ambiente](#)
[Baixar livros de Meteorologia](#)
[Baixar Monografias e TCC](#)
[Baixar livros Multidisciplinar](#)
[Baixar livros de Música](#)
[Baixar livros de Psicologia](#)
[Baixar livros de Química](#)
[Baixar livros de Saúde Coletiva](#)
[Baixar livros de Serviço Social](#)
[Baixar livros de Sociologia](#)
[Baixar livros de Teologia](#)
[Baixar livros de Trabalho](#)
[Baixar livros de Turismo](#)