A wavelet transform approach to seismic processing

P r o e f s c h r i f t

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof. dr. ir. J. Blaauwendraad,
in het openbaar te verdedigen ten overstaan van een commissie,
door het College van Dekanen aangewezen,
op maandag 24 november te 16:00 uur

door

Franciscus Joseph DESSING

doctorandus in de sterrenkunde
geboren te Rotterdam

Dit proefschrift is goedgekeurd door de promotor:
Prof. dr. ir. A. J. Berkhout

Toegevoegd promotor:
Dr. ir. C. P. A. Wapenaar

Samenstelling Promotiecommissie:

Rector Magnificus, voorzitter	
Prof. dr. ir. A. J. Berkhout	*Technische Universiteit Delft, promotor*
Dr. ir. C. P. A. Wapenaar	*Technische Universiteit Delft, toegevoegd promotor*
Prof. dr. ir. J. T. Fokkema	*Technische Universiteit Delft*
Prof. dr. R. K. Snieder	*Universiteit Utrecht*
Prof. dr. ir. M. A. Viergever	*Universiteit Utrecht*
Prof. ir. K. F. Wakker	*Technische Universiteit Delft*
Dr. T. J. H. Smit	*Shell U.K. Exploration and Production, London*

ISBN 90-9011084-4

SUPPORT
The research for this thesis has been financially supported by the DELPHI consortium.

Typesetting system: LaTeX 2ε
Printed in The Netherlands by: Beeld en Grafisch Centrum, Technische Universiteit Delft.

COVER
The earth unlimited by Joossien Nap and Frank Dessing

Aan mijn moeder

en

Aan Joossien

Contents

Chapter 1

Introduction

1.1 Survey

The subject treated in the present monograph belongs to a specific discipline in the earth sciences, known as exploration geophysics. An exploration geophysicist commits himself to discovering and characterizing natural resources, such as oil, coal and gas. For this purpose he often resorts to the seismic method, which is an elastodynamic remote sensing technique. A steady increase of the energy demand in a globally competitive business causes a continuous search for techniques to improve the seismic method, in terms of efficiency and effectiveness. A potentially interesting analysis technique has been developed in the late eighties: the wavelet transformation. The wavelet transformation is a mathematical tool; it makes the scale and time or space dependency, present in almost any measurement of physical phenomena, explicitly manifest. In this thesis, the significance of the wavelet transformation in various stages of the seismic method will be explored.

The current chapter will be utilized to concisely introduce the main aspects of the thesis, and to point out the elements which comprise innovations. In section 1.2 the seismic method and the associated conceptual physical model will be dealt with. In section 1.3 attention will be paid to the role of representations and transformations for the seismic method, more specifically to the role of the wavelet transformation. In section 1.4 the outline of the thesis will be discussed with an emphasis on its novel aspects. Finally, in section 1.5 a number of notational and mathematical conventions applicable to the complete thesis are put together.

1.2 The seismic method

The field of earth sciences sets itself the target to acquire extensive knowledge of the structure and material properties of the earth, and of the dynamic processes

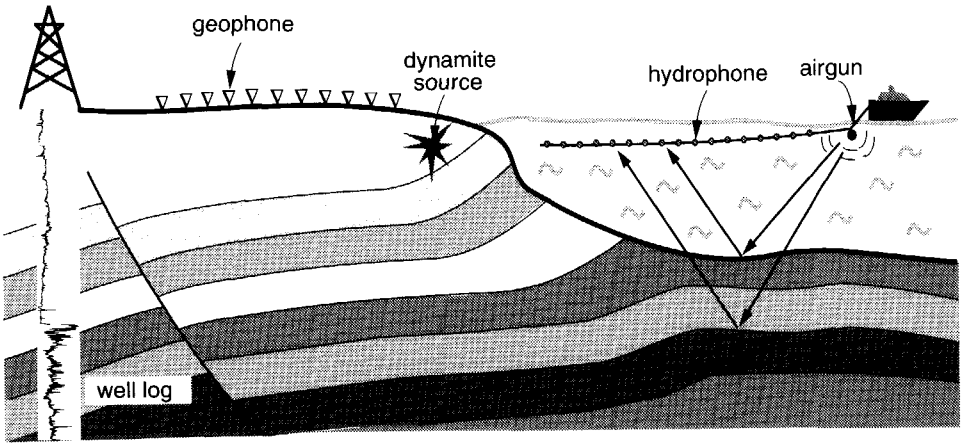


Fig. 1.1 *An artist impression of the surface seismic reflection method. Generally a distinction is made between a land survey (left-hand side) and a marine survey (right-hand side)*

taking place in the earth. This knowledge serves a wide range of goals. For the exploration geophysicist the goal is to locate natural resources in order to meet the expected worldwide growth in demand for energy (EC-report, 1996). An exploration geophysicist does not only want to know where the fossil fuels are, but also how to get them out. Consequently, he is interested in both the structure and the material properties of the subsurface.

A naive method to reveal the two of them is to excavate the earth. Besides the accurateness of this method, it is quite impractical due to its low success rate and due to the fact that it severely disturbs the balance of the living creatures in and on the earth. Fortunately, the exploration geophysicist can investigate the earth more efficiently, either by probing it remotely and globally, or by probing it directly and locally. The remote and global method comprises various techniques, such as gravimetry and seismic exploration. The direct and local method comprises all logging techniques, such as electromagnetic and sonic logging. If the earth is considered to be a system, than these methods have in common that they provide information on the system earth (both its structure and material properties) via a process, that is generally externally initiated.

Among the global methods the method of seismic exploration, also referred to as the seismic method, is by far the most powerful. In the seismic method, elastic wave fields are generated by localized sources. The elastic wave fields interact with the system earth and after the interaction they are measured with localized receivers. The seismic method turns out to be an adequate method to infer, via the measured

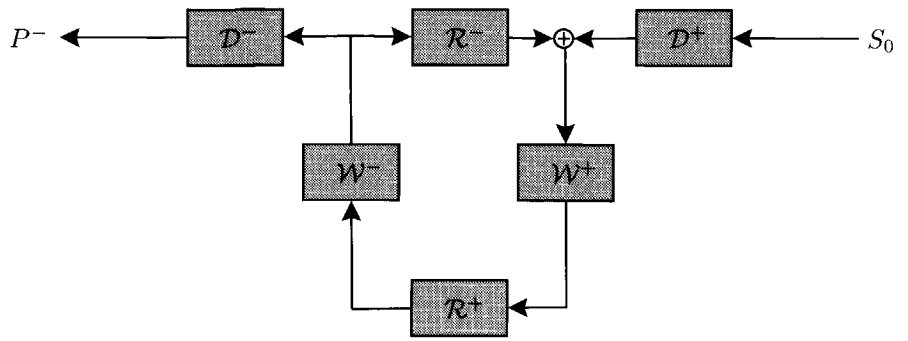


Fig. 1.2 *Conceptual model of a surface seismic reflection experiment.*

wave fields, both the structure and the material properties of the earth, i.e. both to locate the natural resources and to characterize them.

The seismic exploration method is not just a single unique method. It constitutes a number of methods which can be discriminated on the basis of their specific source and receiver configurations. The surface seismic reflection method, which is most often used, is characterized by the fact that both sources and receivers are located at or near the surface. Figure 1.1 presents an artist impression of the surface seismic reflection method. Albeit an artist impression, Figure 1.1 gives a good idea what kind of interaction is referred to above. An elastodynamic wave, also referred to as an elastic wave, is propagating from the source down into the subsurface, it is partially reflected at points where the material properties relevant for wave propagation vary, and after this interaction part of the wave propagates up to the surface and will be measured by a set of receivers.

On a conceptual level the surface seismic reflection method can be represented by a set of operators related to the distinctive features of the method. Together the operators form the conceptual model of Figure 1.2. The quantity S_0 represents a source wave field. The operator $\mathcal{D}^+$ converts the source wave field into a downgoing wave field. Moreover, it can be used to represent the configuration of the sources at or near the surface. The propagation from the surface down into the subsurface is represented by the propagation operator $\mathcal{W}^+$. Reflection in the subsurface at a specific depth level is described by the reflection operator $\mathcal{R}^+$. The propagation up to the surface is given by the propagation operator $\mathcal{W}^-$. The upgoing wave field at the surface is fed into the earth again via a reflection at the surface. The reflection at the surface is represented by $\mathcal{R}^-$. It converts upgoing wave fields into downgoing wave fields. The measuring process at the surface is represented by the operator $\mathcal{D}^-$. It converts the upgoing wave fields into wave fields measured by a hydrophone or geophone. Besides, $\mathcal{D}^-$ represents the receiver configuration. Hence,

the action of the operator $\mathcal{D}^-$ yields the seismic reflection data P^-. The presented model is referred to as the WRW-model (Berkhout, 1982; Berkhout and Wapenaar, 1990); it is a forward model, making manifest the most relevant aspects of a seismic experiment.

To infer the structure and the material properties of the subsurface from the seismic reflection data, an inversion procedure is required. Generally, a distinction is made between two phases of inversion. In imaging the aim is to estimate the reflection operator $\mathcal{R}^+$ by getting rid of the effects of the surface related operators $\mathcal{D}^+$ and $\mathcal{D}^-$ (decomposition) and $\mathcal{R}^-$ (multiple elimination), and by removing the effects of the propagation operators $\mathcal{W}^+$ and $\mathcal{W}^-$ (downward and upward extrapolation). The imaging procedure reveals the structure of the earth in terms of the reflectivity $\mathcal{R}^+$. As was recently pointed out by Berkhout (1997a), seismic imaging does not require a prior knowledge of $\mathcal{R}^+$. The latter is very important as geological boundaries may represent complex reflectors. Imaging is followed by lithologic inversion. The aim of lithologic inversion is to estimate the material properties relevant for the recovery of the natural resources, such as the type of the rock, its porosity, its porefill and its permeability. The fully estimated angle-dependent reflection operator $\mathcal{R}^+$ is input to the lithologic inversion procedure.

The trend in the oil industry has shifted from a wish to increase the oil and gas production "at any cost" in the late 1970s and 1980s to a more cost-sensitive way of operating in the 1990s (EC-report, 1996). Hence, just roughly pinpointing the location of the natural resource reservoirs does not suffice anymore; the exploration geophysicist has to provide information that really improves the production of the reservoirs, especially in complex areas and/or for complex subsurface structures, where other exploration techniques are difficult and relatively expensive. The exploration geophysicist resorts to various solution techniques to meet this requirement. With the help of the conceptual model of the previous paragraph, the solution techniques can be easily ordered.

The described requirement necessitates, first of all, larger and more flexible, sometimes even repeatable, acquisitions (represented by the operators $\mathcal{D}^\pm$), yielding larger data sets (represented by P^-). Secondly, it requires a more accurate preprocessing to separate desired and undesired signals, for example primaries and multiples (removal of $\mathcal{R}^-$). Thirdly, it calls for a more accurate and more efficient incorporation of the effects of complex structures on the propagation operators (inversion for $\mathcal{W}^\pm$). Fourthly, it asks for a better understanding of the relation between the complex subsurface parameters and the effective angle-dependent reflection behavior (understanding of $\mathcal{R}^+$). Finally, it necessitates the development of techniques to automatically recover potentially interesting areas from the output images via image analysis techniques (analysis of $\mathcal{R}^+$). Divers roads can be followed to fulfill the presented goals. In this thesis, I will present the opportunities provided by the wavelet transformation.

1.3 Representations and transformations

A representation of a particular field or operator is the projection of this field or operator on a particular set of functions. The advantage of choosing a particular representation is expressed by Marr (1982): "a representation is a formal system for making explicit certain entities or types of information, together with a specification of how the system does this". Any representation highlights certain aspects of the data at the expense of information pushed into the background. Hence, some representations might be appropriate for efficient computations, while others are for example more appropriate to filter undesired signals. A transformation is a rule prescribing how to go from one representation of a field or operator to another representation of the same field or operator. In section 1.3.1 the attention will be focussed on the wavelet representation, and in section 1.3.2 the opportunities provided by the wavelet transformation will be discussed in relation to the surface seismic reflection method.

1.3.1 The wavelet transformation

Among the many possible transformations and associated representations the wavelet transformation, yielding the wavelet representation, is by far the most recently developed. In 1982, the geophysicist Morlet introduced the wavelet transformation as a tool to carry out a joint time-frequency analysis. Instead of a decomposition into frequency components as the Fourier transform[1] does, or instead of a decomposition into frequency components within a window of fixed size, as the windowed Fourier transform does, the wavelet transformation is a decomposition of a field or operator into components with fixed shape, but with variable location and size. By adjusting the size of the decomposing blocks while keeping the shape fixed, the wavelet transformation acts as a mathematical microscope. Whereas the wavelet transformation in its basic form is continuous and highly redundant, its popularity increased significantly after the formulation of the so-called discrete wavelet transformation, which is an orthogonal and efficient decomposition with basically the same properties as the continuous wavelet transformation. The multiresolution decomposition scheme (Mallat, 1989a) and the advent of wavelets of compact support (Daubechies, 1988) were essential to this development.

[1] A subtle difference exists between the meaning of the words *transform* and *transformation*. Webster (1988) assigns the following meaning to the word *transformation*: "the operation of changing one configuration or expression into another in accordance with a mathematical rule...". The word *transform* is assigned two meanings in Webster: (1) "a mathematical element obtained from another by transformation" and (2) "transformation". Hence, a transform refers either to the result or to the action itself. Both meanings will be used here.

1.3.2 Applications of the wavelet transformation

The wavelet transformation as a tool to improve the seismic reflection method went through a development cycle consisting of naive and broad enthusiasm in the first stage, skepticism with respect to the possibilities of the new tool in the second stage, and dedicated enthusiasm in the third stage. At the moment that this thesis is written, the seismic community is in my opinion moving on from the second to the third stage. The third stage is characterized by a well-developed awareness of the possibilities and limitations, and a good perception of the specific opportunities for the seismic method.

From different fields of study the following successful application areas crystallized out in the course of time: (1) compression, (2) time-frequency signal processing, (3) operator representation, (4) singularity analysis, and (5) image analysis. If a tool is applicable to one field, it does not necessarily mean that the application will be successful in another field as well. The operators, the data, or the image to which the transformation has to be applied, might be of just a different nature, causing the tool to be less effective. Considering the five application areas, looking at the major steps in the surface seismic reflection method of Figure 1.2, and facing the challenges discussed at the end of section 1.2 on page 4, the following potentially interesting applications for the wavelet transformation can be formulated[2], which I list here together with their current status:

1. *Compression*

   - Application: compression of seismic reflection data.
   - Principle: decorrelation of reflection events.
   - Status: successfully applied with compression ratios up to 100 (Bosman and Reiter, 1993; Reiter, 1996).
   - Remark: the complete compression algorithm consists of three main steps: the wavelet transformation, a quantization step, and a coding step on bit level (Vetterli and Kovačević, 1995; Chen, 1995).

2. *Time-frequency signal processing*

   - Application: preprocessing of seismic reflection data, for example multiple suppression, ground-roll removal or noise suppression.
   - Principle: local subdivision of the time-frequency plane, and sensitivity to local changes.
   - Status: there is no clear advantage over existing methods yet.
   - Remark: especially the trace-to-trace coherency of a seismic reflection event is not easily taken into account.

[2]Note that only the major application areas are given.

3. *Operator representation*

   - Application: representation of propagation operators for efficient migration schemes.
   - Principle: wavelet transformation is "blind" to slowly varying parts in the operator (Beylkin et al., 1991).
   - Status: in development; increased interest due to success of compression.
   - Remark: chapter 4 deals with this subject.

4. *Singularity analysis*

   - Application: characterization of singularities in the subsurface and especially in the reflection operator.
   - Principle: a singularity, which is an irregular point, implies scale dependency. The wavelet transformation characterizes the scale dependency.
   - Status: successfully applied for the analysis of well-logs (Herrmann, 1997); scale dependency of reflection operator and hence of seismic reflection data is subject of current research.
   - Remark: chapter 5 deals with this subject.

5. *Image analysis*

   - Application: interpretation of migrated data for the automatic extraction of geological features of interest for the evaluation of potential reservoirs.
   - Principle: visual system can be mimicked by a proper set of operators applied to a multiscale representation of a data set (ter Haar Romeny, 1994).
   - Status: in the initial stages of development.
   - Remark: chapter 6 deals with this subject.

Figure 1.3 summarizes the above listing. Due to the fact that the third, the fourth and fifth item form the heart of seismic imaging and characterization, I have concentrated on these items. Three separate chapters are dedicated to them. The other items are only briefly reported on in chapter 2, where I discuss the basic properties of the wavelet transformation.

1.4 Outline of the thesis and its novel aspects

In chapter 2, the wavelet transformation is introduced within the realm of general transforms and representations. At the end of chapter 2 the main application areas of the wavelet transformation as introduced in the previous section are worked out in more detail. Up to section 2.6, the content of chapter 2 is a rearrangement of subjects from different literature sources. The discussion of the relevance of the

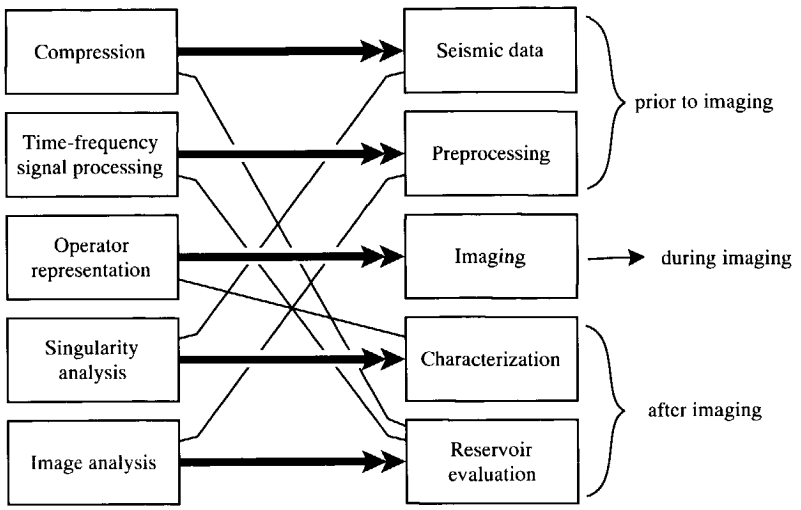


Fig. 1.3 *Successful application areas of the wavelet transformation in other fields of study are shown in the left-hand side column. The right-hand side column shows the main steps of a seismic data processing sequence. The lines between the left and the right column indicate which application area of the wavelet transformation is potentially interesting for the use in seismic data processing. The fat lines denote the main opportunities.*

wavelet transformation for the seismic method in section 2.6 has not been presented in this form before.

In chapter 3, the one-way forward model for a surface seismic experiment is presented. Here, I also pay attention to the representation of the propagation operators via a modal decomposition and to the spectral properties of the Helmholtz operator. The derivation of the forward model reshuffles publicly available literature. The concise and systematic discussion of the spectral properties of the Helmholtz operator is a novelty for the geophysical community. The subsequent derivation of the propagation operator via a modal decomposition is in its basic form not completely new, but its usefulness for geophysical imaging seems to have been overlooked.

In chapter 4, the role of the wavelet transformation in improving the time consuming process of migration is discussed. Whereas the propagation operators themselves cannot be represented particularly efficient in the wavelet transform domain, the wavelet transformation allows an efficient subdivision of the migration process enabling a coarse-to-fine reconstruction of the structural subsurface picture. The method is illustrated with synthetic and real data examples. The advantages

and disadvantages are pointed out. The results of chapter 4 have not been presented elsewhere.

Chapter 5 deals with the effect of singularities (a step function is a singularity as well) on waves reflecting at those singularities. It shows how the wavelet transformation can be utilized to analyze the effects in a deterministic way. In chapter 5 the so-called wavelet transform modulus maxima analysis is introduced first. It is a well-developed tool to determine the strength of a singularity (Mallat and Hwang, 1992). A complete characterization can be obtained with an additional analysis of the phase change induced by the singularity. In order to determine how singularities in the medium are transferred to the (angle-dependent) reflectivity and, hence, to the interacting wave field, analytical derivations and numerical experiments have been carried out. The size of the distortion with respect to the background velocity turns out to be an important factor. The analysis is concluded with a proposal for a singularity driven inversion scheme. The systematic use of the phase information to characterize a singularity and the reflection at a singularity is regarded to be new.

Chapter 6 discusses the use of the wavelet transformation in image analysis. Here, the image is a 3-D migrated data set from the coastal area of the Gulf coast, containing lots of "hidden" faults and channels. The nature of migrated data in general is such that strong reflectors are highlighted. Faults, channels, point bars, etc, can be found indirectly via inconsistencies in the strong reflections. Generally, these features are not strong reflectors themselves. Hence, the nature of migrated data is such that certain aspects of great interest for the evaluation of potential reservoirs are pushed into the background. The application of the 3-D wavelet transformation turns out to be a valuable tool in extracting the "hidden" information. The simplicity of the presented method together with its success is a novel aspect.

The chapters 2 up to and including 6 are finalized with a summarizing section, where the main aspects discussed in the chapter under consideration are put together.

A number of the more detailed mathematical en notational parts, which are not completely new, but which are otherwise essential enough to be incorporated, have been passed to the appendices. In appendix A, the matrix notation for transformations and representations is dealt with. It supports the content of chapter 2 and chapter 4. In appendix B, I deal with homogeneous distributions and their properties. It sustains the line of reasoning presented in chapter 5.

1.5 Notational conventions

To specify the position in an arbitrary configuration I employ the 3-D coordinate vector $\boldsymbol{x} = (x_1, x_2, x_3)$, which is defined with respect to a right-handed orthogonal Cartesian coordinate system. The mutually perpendicular base vectors in the coordinate system are given by $\boldsymbol{i}_1 = (1,0,0)$, $\boldsymbol{i}_2 = (0,1,0)$, and $\boldsymbol{i}_3 = (0,0,1)$. The

$\boldsymbol{i}_3$-axis is referred to as the direction of preference. In geophysical applications the vector $\boldsymbol{i}_3$ generally points vertically downwards. Roman subscripts for the position coordinates can take the values $1, 2, 3$, i.e. x_k denotes x_1, x_2 or x_3. Greek subscripts are reserved for the coordinates perpendicular to the $\boldsymbol{i}_3$-direction; they are referred to as the lateral or horizontal coordinates. Hence, x_μ denotes either x_1 or x_2. The lateral coordinates are also denoted by $\boldsymbol{x}_L = (x_1, x_2)$. Partial differentiation with respect to one of the spatial coordinates x_k is denoted by ∂_k. The time coordinate is denoted by t. Partial differentiation with respect to time is denoted by ∂_t. The parameter σ is frequently encountered. It is referred to as the scale parameter. Partial differentiation with respect to the scale parameter is denoted by ∂_σ. With respect to repeated subscripts Einstein's summation convention applies, i.e.

$$\begin{aligned} x_k y_k &= x_1 y_1 + x_2 y_2 + x_3 y_3 \quad \text{for Roman subscripts, except } t, \\ x_\mu y_\mu &= x_1 y_1 + x_2 y_2 \quad \text{for Greek subscripts, except } \sigma. \end{aligned}$$

Calligraphic uppercase characters denote operators. Bold calligraphic uppercase characters denote operator matrices. Bold sans serif uppercase characters denote discrete matrices. Bold sans serif lowercase characters denote discrete vectors. The symbols tilde $\tilde{}$, hat $\hat{}$, and check $\check{}$ are reserved for representations in the Fourier domain, in the Gabor domain and in the wavelet domain, respectively. Note, however that continuous representations of wave fields in the temporal Fourier domain are denoted by uppercase characters.

Chapter 2

Representations, analysis techniques and operators

2.1 Introduction

A mathematical model of a physical process can be written in a conceptual form as

$$\mathcal{A}u = f, \tag{2.1}$$

where $\mathcal{A}$ is an operator, whose action on the field u yields the field f. A complete understanding of the action of an operator on an arbitrary field in the domain of the operator can be acquired by solving the eigenvalue problem of $\mathcal{A}$, which reads

$$\mathcal{A}\phi_\lambda = \lambda\phi_\lambda, \tag{2.2}$$

where λ is the eigenvalue and ϕ_λ the corresponding eigenvector or eigenfunction. The solution of the eigenvalue problem provides the elementary components, into which every field f or u can be uniquely decomposed; this decomposition is called the eigenvalue or modal decomposition. The eigenvalue decomposition gives the idealized operator dependent representation of both the operator and the field. There are two major reasons to consider other decompositions and representations as well. Firstly, in the case that the operator $\mathcal{A}$ is just one part of a larger physical process, the eigenvalue decomposition of $\mathcal{A}$ is not necessarily a useful decomposition for another part of the process[1], or for effectively decomposing the field quantities f or u. Secondly, a modal decomposition draws heavily on the computational abilities. There are other decompositions that can be computed more efficiently. This chapter

[1] In terms of the conceptual model of a surface seismic experiment of Figure 1.2, this aspect can be easily illustrated. The eigenvalue decomposition of the propagation operators $\mathcal{W}^{\downarrow}$ is generally not the ideal decomposition for the reflection operator $\mathcal{R}^{+}$.

concisely introduces various representations and transformations. In chapters 4, 5, and 6 the representations and transformations will be applied to the different elements of the conceptual model of the surface seismic reflection method, which has been introduced in the previous chapter and which will be worked out in considerable detail in the next chapter.

Before I proceed, let me elucidate at this point two important notions: *representation* and *transformation*. Here, I adopt the definitions and habits of quantum mechanics (Messiah, 1958). A *representation* of a field f in the variable $a \in \mathbb{R}^n$ is a projection of this field on a particular set of functions or generalized functions[2] ψ_a of the variable a, which can be expressed symbolically by the inner product of f with the projection functions, according to[3]

$$f(a) = \langle f, \psi_a \rangle . \tag{2.3}$$

Similarly, the representation of a linear operator $\mathcal{A}$ in the variable a can be expressed by

$$A(a, a') = \langle \mathcal{A}\psi_{a'}, \psi_a \rangle . \tag{2.4}$$

What is the advantage of choosing a particular representation? David Marr (1982) in his book *Vision* put it this way:

> *"A representation is a formal system for making explicit certain entities or types of information, together with a specification of how the system does this."*

Of course, choosing one representation often means that certain aspects in the data or in the operator are highlighted at the expense of other aspects:

> *"... Thus, there is a trade-off; any particular representation makes certain information explicit at the expense of information that is pushed into the background and may be quite hard to recover."*(Marr, 1982).

Translating this to the problem at hand, it means that some representations might be more appropriate for efficient computations (in the case of an emphasis on the propagation operator), while others are designed to filter unwanted aspects in the data (emphasis on data), and again others are effectively representing the complexity in the reflectivity (emphasis on reflectivity).

The other term that requires a further elucidation is the concept transformation. A *transformation* is a rule prescribing how to go from one representation of a field or operator to another representation of the same field or operator. Figure 2.1

[2]For the concept of generalized functions or distributions the reader is referred to Zemanian (1965) or to section 2.4.

[3]The abstract equations (2.3) and (2.4) will be elucidated in the course of this chapter.

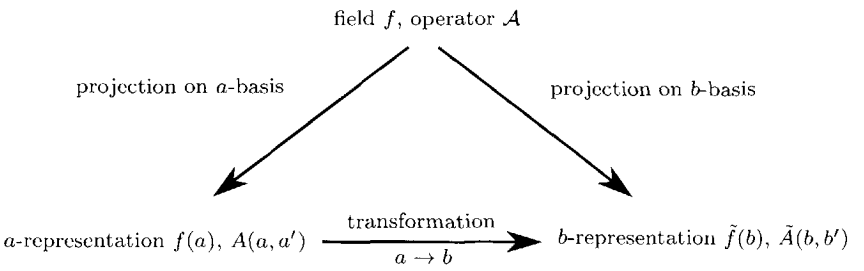


Fig. 2.1 *The representation of a field f and an operator $\mathcal{A}$ consists of a projection on a particular set of functions. A transformation links two representations. The tilde in the b-representation makes the difference with the a-representation explicit.*

schematically shows the relation between representations and transformations. Unlike the situation in quantum mechanics, in seismology a wave field is almost always dealt with as a function of time and spatial coordinates, i.e. as a representation. Hence, the original representation is often fixed by the temporal and spatial filters, which the wave fields are projected on.

A large number of representations and associated transformations are available. Some ordering is required to be able to make a sound choice. Here, I make a discrimination between representations associated with self-adjoint operators and representations associated with so-called coherent states. Self-adjoint operators give rise to an orthogonal basis of eigenfunctions. Every function in an appropriate space can be uniquely decomposed in the orthogonal basis. Well-known examples are Legendre polynomials, Bessel functions, or complex exponentials, but also δ-distributions, Hermite polynomials or spherical harmonics (Churchill and Brown, 1978; Messiah, 1958). As I am looking for rather general representations, which nevertheless have attractive properties in various applications, the number of possibilities is limited. I will consider representations induced by complex exponentials, i.e. the Fourier representation, and the representation related to the δ-distribution, i.e. the spatial or temporal representation. They are associated with the derivative operator and the space or time operator, respectively. I will also study an intermediate representation, related to a combination of the derivative operator and the time operator, i.e. the Mellin or log-modulation representation.

While a decomposition in eigenfunctions has proven undoubtedly its usefulness, it is in other instances a limitation. A field which is not exactly an eigenfunction should be described by a linear combination of the eigenfunctions. Especially, in classical physics or classical limits of quantum physics, one is not bounded to work with eigenfunctions. More than that, the eigenfunctions are often physically unre-

alistic. A Fourier component or a δ-distribution, albeit very powerful mathematical objects, will never occur in real life. These limitations, though not always motivated in the same way, caused both quantum physicists and engineers to introduce a continuous set of square-integrable, i.e. measurable, states. Quantum physicists called them coherent states (see Primer, Klauder and Skagerstam, 1985). Coherent states are generated by applying unitary operators to a fixed ground state. A well-known example of a coherent state has been introduced in the signal analysis community by Dennis Gabor (1946)[1]:

$$g_{ab}(t) = e^{jat} g(t - b),$$

where $g(t)$ is the normalized ground state. The coherent states $g_{ab}(t)$ are generated by applying unitary translation and modulation operators to the function $g(t)$. If the function $g(t)$ is concentrated around $t = 0$ and its Fourier transform $\tilde{g}(\omega)$ around $\omega = 0$, then it can be easily seen that $g_{ab}(t)$ is concentrated around $(t, \omega) = (b, a)$. Via the parameters a and b the functions $g_{ab}(t)$ can be used to carry out a local Fourier analysis, whence the often employed term windowed Fourier transform. Are the functions $g_{ab}(t)$ with $(a, b) \in \mathbb{R}^2$ completely describing any function in a certain functional space? Are any restrictions applicable to the ground state $g(t)$? Answers to these questions have been given in the past decades and will be discussed in section 2.3.

Another set of coherent states are generated by translating and dilating a particular square-integrable ground state $\psi(t)$, i.e.

$$\psi_{\sigma b}(t) = \frac{1}{\sqrt{|\sigma|}} \psi\left(\frac{t - b}{\sigma}\right),$$

with $\sigma \neq 0$. The function $\psi(t)$ has to obey the 'admissibility' condition:

$$\int \psi(t)\, \mathrm{d}t = 0.$$

The functions $\psi_{\sigma b}(t)$ are called affine coherent states or wavelets. They have been introduced by Aslaksen and Klauder (1968, 1969). The geophysicist Morlet (Morlet et al., 1982) introduced wavelets as a useful tool to carry out a joint time-frequency analysis of dispersed seismic reflection and transmission data. The importance of wavelets stems firstly from the fact that they are "blind" to trends in a signal, which is directly due to the admissibility condition. Secondly, wavelets act through the scaling parameter σ as a mathematical zoom. By enlarging σ the analyzing wavelet $\psi(t)$ is stretched, whereas a smaller σ causes a highly compact version of the wavelet, which allows one to zoom in on small details at the location given by the translation parameter b. Hence, the parameter σ enables a local analysis at multiple scales or

[1]The term coherent state was only introduced in the sixties.

resolutions. The zooming property in addition to its blindness to trends has resulted in a number of interesting applications such as regularity analysis and singularity detection (Holschneider and Tchamitchian, 1990; Mallat and Hwang, 1992; Bacry et al., 1993; Herrmann, 1997). The importance of singularities in physics has been beautifully formulated by Maxwell (1882, p. 443):

> *"For example, the rock loosed by frost and balanced on a singular point of the mountain-side, the little spark which kindles the great forest, the little word which sets the world a fighting, the little scruple which prevents a man from doing his will, the little spore which blights all the potatoes, the little gemmule which makes us philosophers or idiots. Every existence above a certain rank has its singular points: the higher the rank, the more of them. At these points, influences whose physical magnitude is too small to be taken account of by a finite being, may produce results of the greatest importance. All great results produced by human endeavour depend on taking advantage of these singular states when they occur."*

Realizing that reflection can be seen as the local interaction of a wave with a singular medium, the association with wavelets is easily made and will be worked out in the remaining part of the thesis.

In the fall of 1986, Mallat and Meyer realized that a representation in the wavelet domain also expresses the difference between observations made at two different scales. Their theory of a multiresolution approximation made the link between wavelets and measurements more manifest. Without loosing too much of the properties of the *continuous* wavelet transform, it turns out that, within the theory of a multiresolution approximation, a *discrete* wavelet transform can be carried out, at least as fast as the fast Fourier transform. The discrete wavelet transform is obtained by choosing particular discrete translation values for b and discrete dilation values for σ. The advent of the multiresolution approximation scheme by Meyer (1986) and by Mallat (1989b), to be discussed in section 2.5, and the advent of orthogonal wavelets of compact support (Daubechies, 1988) were major factors in increasing the popularity of the wavelet representation.

Of course, the discussion of possible representations in a thesis, or even in a complete textbook on representations, cannot be all encompassing. The reader can always point out properties, applications, or related subjects that are not discussed. The exposition is dictated by a personal historical development placed in the aforementioned framework and by the applications in subsequent chapters. Let me discuss some of the limitations. The body of this chapter is concerned with signals or functions depending on one variable only, and the operators discussed are set to work on one-dimensional signals. More dimensional versions are not necessarily straightforwardly obtained. In subsequent chapters, I will introduce more dimensional versions as soon as I will utilize them. Another restriction is related to the

choice to work with linear representations only. The fact that I am considering the linearized wave equation, and the fact that it is appealing to consider the decomposing building blocks as special wave field functions, which can be manipulated in the transformed domain, made me not consider non-linear representations, like the Wigner-Ville distribution (Cohen, 1993)[5]. I will also not consider a completely signal dependent transformation like the Karhunen-Loève transformation (Vetterli and Kovačević, 1995). It optimally decorrelates a signal, but it is not related to any operator. The reader will also not find an extensive historical overview of the field of wavelets. I would like to refer the reader to the books of Daubechies (1992) or Meyer (1993), and to last year's *Special issue on wavelets* of the IEEE (see, for example, Daubechies, 1996).

In the sequel, the concepts that have been hinted at in this introduction will be worked out. I will start with the framework for representations and transformations provided by self-adjoint operators in section 2.2. Here, the fundamental variables time t, temporal frequency ω and the logarithmic modulation c and their associated eigenfunctions will be introduced[6]. In section 2.3 coherent states will be introduced as an alternative for the representations related to self-adjoint operators. The Gabor representation and the wavelet representation will be presented. A number of properties of the two representations will be discussed. In section 2.4 the properties of the wavelet transform will be discussed. Emphasis will be put on the concept scale. Subsequently, in section 2.5 discrete versions of the introduced representations will be discussed. Here, the multiresolution framework and the discrete wavelet transform (as opposed to the continuous wavelet transform) will be brought in. In section 2.6 the main applications of the wavelet transform will be discussed in relation to the seismic method. Section 2.7 summarizes the main aspects of this chapter.

2.2 Representations and transformations related to self-adjoint operators

The common use in quantum physics to associate physical quantities with operators has been adopted in the signal analysis community, especially since Gabor (1946) wrote his masterpiece on joint time-frequency analysis[7]. Gabor introduced operators for time and frequency similar to the operators for position and momentum in quantum physics. Although it is not strictly necessary to draw on the extensive

[5]Steeghs (1997) deals with a number of non-linear distributions and their applications in geophysical processing.

[6]The derived properties apply equally well, mutatis mutandis, to the variables position, spatial frequency and spatial logarithmic modulation.

[7]It is important to realize that quantum physics, here or in the rest of this chapter, is not used to clarify aspects in signal analysis. The tools developed in quantum physics are translated to a time-frequency framework instead.

operator theory, it provides a unifying framework of which the variable, the basis and the properties of the basis form integral parts. Moreover, the results obtained in quantum physics for the conjugate pair position-momentum based upon the operator theory can be translated to the conjugate pair time-frequency. The translation has been clearly reported on, amongst others, by Cohen (1993, 1995), and by Baraniuk and Jones (1995). The operator framework will be given for general operators in section 2.2.1. In section 2.2.2 I will specify the exposition for operators relevant for a time, frequency and mixed time-frequency analysis.

2.2.1 General theory

In this subsection the following notions will be introduced: Hilbert space, linear operator, adjoint operator, self-adjoint operator, eigenfunction, unitary operator, and invariance and covariance properties. The exposition will be necessarily brief. For more extensive introductions the reader is referred to Messiah (1958), for a broad overview, or to chapter 3, for a more extensive treatment of some specific subjects.

Hilbert space —A mathematical description of a physical process requires a careful analysis of the states the observable field can possibly attain. The possible states depend on the operators ruling the physical process. They are mathematically described by defining an appropriate Hilbert space. A Hilbert space is a linear space[8]; it is equipped with an inner product denoted by $\langle\,,\,\rangle$, which has to satisfy the following properties:

(a) the inner product of ϕ and ψ is the complex conjugate of the inner product of ψ and ϕ, i.e.

$$\langle \phi, \psi \rangle \;=\; \langle \psi, \phi \rangle^{*}; \tag{2.5}$$

(b) the inner product is linear with respect to the first argument

$$\langle \lambda_1 \phi_1 + \lambda_2 \phi_2, \psi \rangle \;=\; \lambda_1 \langle \phi_1, \psi \rangle \;+\; \lambda_2 \langle \phi_2, \psi \rangle\,; \tag{2.6}$$

(c) the norm of a function ψ is a real non-negative number defined by

$$\|\psi\| \triangleq \sqrt{\langle \psi, \psi \rangle} \geq 0, \tag{2.7}$$

where $\langle \psi, \psi \rangle = 0$, if and only if $\psi = 0$.

In order that a space of functions is a Hilbert space, it is required that $\langle \psi, \psi \rangle < \infty$ for any ψ in the Hilbert space. Finally, a Hilbert space is required to be complete

[8] A vector space is linear in the case that a linear combination of two vectors or functions ψ_1 and ψ_2 in the vector space, is in the vector space as well. In terms of wave fields, this means that the superposition principle applies, i.e. that the linearized wave equation is assumed to be valid.

and separable. Completeness in this context means that every so-called Cauchy sequence[9] of elements in the Hilbert space converges towards an element in the same Hilbert space. Separability means that any element in the Hilbert space can be approached infinitely close by a Cauchy sequence of elements in that space.

Operators —An operator $\mathcal{A}$ assigns to a function ψ in the domain of $\mathcal{A}$ another function ϕ, according to

$$\phi = \mathcal{A}\psi, \tag{2.8}$$

where ψ and ϕ do not necessarily belong to one and the same Hilbert space. If the action of the operator $\mathcal{A}$ is linear in ψ, it is called a linear operator. The inverse of an operator (if it exists) is the operator $\mathcal{B} = \mathcal{A}^{-1}$ that undoes the action of the operator $\mathcal{A}$ on ψ, according to

$$\psi = \mathcal{A}^{-1}\phi. \tag{2.9}$$

Consequently, $\mathcal{B}\mathcal{A} = \mathcal{A}^{-1}\mathcal{A} = \mathcal{I}$, where $\mathcal{I}$ is the identity operator. The adjoint of $\mathcal{A}$ is denoted by $\mathcal{A}^\dagger$. It is the operator for which the equality

$$\langle \mathcal{A}\psi, \phi \rangle \;=\; \langle \psi, \mathcal{A}^\dagger \phi \rangle \tag{2.10}$$

holds true.

Self-adjoint operators —For a self-adjoint operator the adjoint operator equals the operator itself, according to $\mathcal{A}^\dagger = \mathcal{A}$. Consequently, for a self-adjoint operator, the following equality holds:

$$\langle \mathcal{A}\psi, \phi \rangle \;=\; \langle \psi, \mathcal{A}\phi \rangle \,. \tag{2.11}$$

The proper definition of a self-adjoint operator goes along the subtle lines sketched in section 3.3.1 on page 81. Self-adjoint operators form an important subclass of the set of linear operators. They provide insight in the properties of a physical system amongst others through a study of the eigenvalue problem

$$\mathcal{A}\phi_a = a\phi_a, \tag{2.12}$$

where a is the eigenvalue and ϕ_a is the corresponding eigenfunction. The advantage of a self-adjoint operator becomes manifest, if one realizes that all the eigenvalues of a self-adjoint operator are real (in quantum physics it is postulated that they are

[9]The sequence v_n is said to be a Cauchy sequence if, given $\epsilon > 0$, there exists an N such that $\|v_n - v_m\| < \epsilon$ for every n and m greater than N (see, e.g. Higgins, 1977; Gohberg and Goldberg, 1981).

related to the observables, see section 3.3.1) and that the corresponding eigenfunctions form an orthonormal and complete set, hence an orthonormal basis[10]. The orthonormality condition takes the mathematical form

$$\langle \phi_a, \phi_{a'} \rangle = \delta(a - a'), \tag{2.13}$$

and the requirement for completeness is given by

$$\langle \phi_a(t), \phi_a(t') \rangle = \delta(t - t'), \tag{2.14}$$

where on the left-hand side an integration over a in the spectrum $\sigma(\mathcal{A})$ of $\mathcal{A}$ is carried out[11]. In relation (2.14) I have assumed an explicit dependence on a variable t which can be in $\mathbb{R}^n$, but in this chapter t is assumed to be a 1-D variable. The completeness relation means that any function in the Hilbert space can be expanded in a linear combinations of (generalized) eigenfunctions. Another form of the completeness relation is given by

$$v(t) = \int v(a)\, \phi_a(t)\, \mathrm{d}a, \tag{2.15}$$

where $v(a)$ is the representation of a field v in the basis $\{\phi_a\}$, according to

$$v(a) = \langle v, \phi_a \rangle . \tag{2.16}$$

Equation (2.15) is called the closure relation. The integral in equation (2.15) is taken over the values a in the spectrum $\sigma(\mathcal{A})$ of $\mathcal{A}$. The closure relation expresses that the projection of a function on a complete set of eigenfunctions acts as an identity operator. Note that equations (2.13), (2.14), and (2.15) should be understood in the sense of generalized functions.

The representation of an operator $\mathcal{K}$ in the basis $\{\phi_a\}$ is given by

$$\mathrm{K}(a, a') = \langle \mathcal{K}\phi_{a'}, \phi_a \rangle . \tag{2.17}$$

$\mathrm{K}(a, a')$ is called the kernel of the operator $\mathcal{K}$.

[10] Here, it is tacitly understood that the generalized eigenfunctions, those belonging to the absolute continuous or singular continuous part of the spectrum, are included as well. Strictly speaking, the generalized eigenfunctions do not belong to the domain of the operator. See chapter 3 for a more extensive treatment of this issue.

[11] The spectrum $\sigma(\mathcal{A})$ of $\mathcal{A}$ is the set of all eigenvalues of $\mathcal{A}$. Note that in equations (2.13), (2.14) and (2.15) the spectrum is assumed to be non-degenerate and continuous. In the case of a degenerate spectrum or in the presence of other types of spectra as well, the expressions for orthonormality and completeness get slightly more involved, but not fundamentally different (Messiah, 1958). The physical interpretation of the various types of spectra, however, can be quite different.

Unitary operators —Another important class of operators is the unitary operator. For a unitary operator $\mathcal{U}$ the adjoint operator equals the inverse, i.e.

$$\mathcal{U}^{\dagger} = \mathcal{U}^{-1}. \tag{2.18}$$

A unitary operator preserves the inner product, according to

$$\langle \mathcal{U}\psi, \mathcal{U}\phi \rangle = \langle \psi, \mathcal{U}^{\dagger}\mathcal{U}\phi \rangle = \langle \psi, \phi \rangle . \tag{2.19}$$

Consequently, a unitary operator is norm-preserving, that is

$$\langle \mathcal{U}\psi, \mathcal{U}\psi \rangle = \langle \psi, \psi \rangle . \tag{2.20}$$

In this chapter, two types of unitary operators will be encountered[12]. One is related to the transformation between two representations. It is the unitary operator which straddles the representation related to different sets of eigenfunctions. I will refer to this type of unitary operator as *transformation unitary operator*. The other type of unitary operator is related to the representation of a group. Operations like translation, modulation, scaling and rotation, or any mixture of these operations belong to the latter class. I will refer to this type of unitary operator as *parameterized unitary operator*.

Transformation unitary operators –Suppose two different self-adjoint operators $\mathcal{A}$ and $\mathcal{B}$ in a particular Hilbert space are defined with two different complete sets of eigenfunctions denoted by $\{\phi_a\}$ and $\{\psi_b\}$, respectively. Both the spectral values a of $\mathcal{A}$ and b of $\mathcal{B}$ are assumed to be absolutely continuous for ease of notation (see chapter 3, Messiah (1958) or Reed and Simon (1972), for a complete description of the spectrum). Due to the completeness of both sets any function of the set $\{\phi_a\}$ can be expressed as a linear combination of functions from the set $\{\psi_b\}$ and vice versa, cf. equation (2.15),

$$\phi_a = \int \langle \phi_a, \psi_b \rangle \, \psi_b \, \mathrm{d}b \qquad \psi_b = \int \langle \psi_b, \phi_a \rangle \, \phi_a \, \mathrm{d}a. \tag{2.21}$$

The coefficients $\langle \phi_a, \psi_b \rangle$ or $\langle \psi_b, \phi_a \rangle$ can be regarded as the coefficients of an integral kernel[13]. It can be easily shown that $\mathrm{U}(a,b) = \langle \phi_a, \psi_b \rangle^*$ is a unitary transformation kernel relating the representation of a function v in the sets $\{\phi_a\}$ and $\{\psi_b\}$, according to

$$v(a) = \int \mathrm{U}(a,b) v(b) \, \mathrm{d}b \qquad v(b) = \int \mathrm{U}^{-1}(a,b) v(a) \, \mathrm{d}a, \tag{2.22}$$

[12]Baraniuk and Jones (1995) discriminate three types.

[13]In the case of two complete discrete spectra $\langle \phi_a, \psi_b \rangle$ and $\langle \psi_b, \phi_a \rangle$ are matrix kernels.

where $\mathrm{U}^{-1}(a,b)$ is the kernel of the inverse operator, given by $\mathrm{U}^{-1}(a,b) = \langle \phi_a, \psi_b \rangle$. Similarly, the representation of an operator $\mathcal{K}$ in the basis $\{\psi_b\}$, given by $\mathrm{K}(b,b') = \langle \mathcal{K}\psi_{b'}, \psi_b \rangle$, can be transformed to a representation in the basis $\{\phi_a\}$ by

$$\mathrm{K}(a,a') = \int_b \int_{b'} \langle \phi_a, \psi_b \rangle^* \mathrm{K}(b,b') \langle \phi_{a'}, \psi_{b'} \rangle \, \mathrm{d}b' \, \mathrm{d}b \tag{2.23}$$

or

$$\mathrm{K}(a,a') = \int_b \int_{b'} \mathrm{U}(a,b)\mathrm{K}(b,b')\mathrm{U}^{-1}(a',b') \, \mathrm{d}b' \, \mathrm{d}b. \tag{2.24}$$

Equations (2.22) and (2.24) make the symbolic transformations shown in Figure 2.1 explicit.

In the case that a one-to-one correspondence exists between eigenfunctions of the two sets, the matrix or integral kernel U can be expressed as a transformation unitary operator $\mathcal{U}$. Hence, equation (2.24) can be casted in operator form

$$\bar{\mathcal{K}} = \mathcal{U}\mathcal{K}\mathcal{U}^{-1}. \tag{2.25}$$

The operators $\bar{\mathcal{K}}$ and $\mathcal{K}$ are called unitarily equivalent (Messiah, 1958; Daubechies et al., 1986; Baraniuk and Jones, 1995). Unitarily equivalent self-adjoint operators share the same spectrum; however, the corresponding eigenfunctions are generally different; they are related by the unitary operator $\mathcal{U}$. An eigenfunction ϕ of $\mathcal{K}$ transforms to the eigenfunction $\bar{\phi}$ of $\bar{\mathcal{K}}$ according to

$$\bar{\phi} = \mathcal{U}\phi. \tag{2.26}$$

The concept of unitarily equivalent operators has been used by Baraniuk and Jones (1995) to derive, in an elegant way, from existing orthogonal decompositions, like the Fourier transform or the wavelet transform, other orthogonal decompositions better adapted to the signal to be analyzed.

Parameterized unitary operators —Another type of unitary operators can be associated with groups. The concept group[14] formalizes the symmetry aspects in a

[14] **Definition 2.1:** Group (Messiah,1961)
A set $\mathfrak{G}$ *of operations* $a, b, c, \ldots$ *forms a group if*

(i) the product *of any two of them also belongs to the set:*
two operations a and b carried out consecutively can also be applied in one step according to the rule $c = ab$*;*

(ii) one of the elements, I, is the unit element*: application of the unit element does not change the system;*

(iii) each of the elements has an inverse*: each of the operations does not destroy information, i.e., it can be undone;*

(iv) the operation is associative*: in a series of three operations the operations can be done in pairs.*

Note that if the inverse cannot be defined, a semi-group is dealt with.

given system. Following Duffey (1992, page 38): "Every operation that transforms a given system into an equivalent system is called a symmetry operation. By equivalent system, we mean one that is indistinguishable from the original one with respect to the properties under study. All the symmetry operations for a system considered in a certain way form a group."

Example 2.1 *The set of translations or shifts over an axis forms a continuous group with an infinite number of elements. It is defined by*

$$(\mathcal{S}_a f)(t) = f(t-a). \tag{2.27}$$

It can be easily verified that the group requirements for the translation operator are fulfilled.

The shift operator $\mathcal{S}_a$ is a unitary operator and it constitutes a representation of the translation group. According to Stone's theorem (Reed and Simon, 1972, Theorem VIII.8) one can associate to a parameterized unitary operator, like the translation operator or the rotation operator, a self-adjoint operator $\mathcal{A}$ such that

$$\mathcal{U}_a = \mathrm{e}^{ja\mathcal{A}}. \tag{2.28}$$

In the case of a bounded self-adjoint operator $\mathcal{A}$, equation (2.28) can be evaluated with a Taylor series expansion. For an unbounded operator functional calculus is required. The self-adjoint operator $\mathcal{A}$ is called the infinitesimal generator of the unitary operator $\mathcal{U}_a$. In the previous example the infinitesimal generator is the self-adjoint operator $\mathcal{A} = \frac{1}{j}\partial_t$.

Invariance and covariance properties —In order to fully explore a particular representation, one has to be aware of its properties with respect to other representations. The power spectrum of a signal is not sensitive to temporal shifts of the signal, however, it is sensitive to modulations of the original signal. The search for the properties of a system can be carried out by looking at commutation relations between the parameterized unitary operators and self-adjoint operators representing a physical system (Messiah, 1961, Ch. XV, Section III). Two operators $\mathcal{A}$ and $\mathcal{B}$ commute if their commutator $[\mathcal{A}\,,\,\mathcal{B}]$ is zero, according to

$$[\mathcal{A}\,,\,\mathcal{B}] \triangleq \mathcal{A}\mathcal{B} - \mathcal{B}\mathcal{A} = 0. \tag{2.29}$$

In the case that $\mathcal{A}$ is a self-adjoint operator, i.e. an operator representing a certain physical variable a, and $\mathcal{B}$ is a parameterized unitary operator $\mathcal{U}_b$, equation (2.29) reads

$$[\mathcal{A}\,,\,\mathcal{U}_b] = 0, \tag{2.30}$$

and consequently, the variable a is invariant under the group of transformations associated with $\mathcal{U}_b$.

Example 2.2 *Consider again the shift operator* $(\mathcal{S}_a f)(t) = f(t-a)$. *It is well-known that the frequency content of a signal is invariant under temporal translations, except for a phase factor. The frequency content of a signal can be obtained by doing a decomposition in complex exponentials. The complex exponentials are the generalized eigenfunctions of the self-adjoint operator* $\mathcal{F} = \frac{1}{j}\partial_t$. *Consequently, one might expect that* $\mathcal{S}_a$ *and* $\mathcal{F}$ *commute, i.e.*

$$[\mathcal{F}, \mathcal{S}_a] = 0,$$

which can be easily verified to be true[15].

Besides the invariance properties, it is often interesting to know how a representation in a specific domain behaves under (unitary) changes of the representation in another domain. If the changes carried out in one domain can be carried out equally well in the other domain, the associated properties are called covariant (Baraniuk and Jones, 1995). In the discussion of the different representations attention will be paid to both invariance and covariance properties. More particularly, I will look for invariance and covariance properties under the action of a shift operator, a dilation operator, and a modulation operator.

2.2.2 Time, frequency, and logarithmic modulation

In the previous section the apparatus has been developed which enables an investigation of the physical variables time t, frequency ω and logarithmic modulation[16] parameter c in terms of self-adjoint operators and associated unitary operators[17]. Following Cohen (1993, 1995), and Baraniuk and Jones (1995), the self-adjoint representations of time, frequency and log modulation are in the time domain[18] given by

$$\begin{aligned} &\text{Time operator} && (\mathcal{T}f)(t) = tf(t) \\ &\text{Frequency operator} && (\mathcal{F}f)(t) = \frac{1}{j}\partial_t f(t) \\ &\text{Log modulation} && (\mathcal{C}f)(t) = \left(\frac{\mathcal{T}\mathcal{F} + \mathcal{F}\mathcal{T}}{2} f\right)(t). \end{aligned} \tag{2.31}$$

Although the terminology might not be a priori clear, the solution of the (generalized) eigenvalue problems of the three self-adjoint operators will show that the

[15] Knowing that the infinitesimal generator of $\mathcal{S}_a$ is $\mathcal{F}$, it is a trivial result.

[16] In the sequel abbreviated as log modulation.

[17] In the introduction I explained that the results can be used, mutatis mutandis, for the corresponding variables in space as well.

[18] A representation of an operator in the time domain means that the operator is set to work on a function projected on the temporal basis.

terminology makes sense. The eigenvalue problems (cf. equation 2.12) are

$$\begin{array}{ll} \text{Time} & \mathcal{T}\phi_{t'} = t'\phi_{t'} \\ \text{Frequency} & \mathcal{F}\phi_{\omega} = \omega\phi_{\omega} \\ \text{Log modulation} & \mathcal{C}\phi_{c} = c\phi_{c}, \end{array} \tag{2.32}$$

and the resulting generalized eigenfunctions read in the temporal domain

$$\begin{array}{ll} \text{Time} & \phi_{t'}(t) = \delta(t - t') \\ \text{Frequency} & \phi_{\omega}(t) = \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{j\omega t} \\ \text{Log modulation} & \phi_{c}(t) = \frac{1}{\sqrt{2\pi}}\, t^{jc-1/2} \quad t \geq 0. \end{array} \tag{2.33}$$

The representation of a function in the domain associated with these eigenfunctions, reads

$$\begin{array}{ll} \text{Time} & f(t') = \langle f(t), \delta(t - t') \rangle \\ \text{Frequency} & \tilde{f}(\omega) = \langle f(t), \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{j\omega t} \rangle \\ \text{Log modulation} & \check{f}(c) = \langle f(t), \frac{1}{\sqrt{2\pi}}\, t^{jc-1/2} \rangle, \end{array} \tag{2.34}$$

which are identified as the time domain representation, the frequency domain representation, and the log-modulation domain representation, respectively. The frequency domain representation is better known as the Fourier domain representation, associated with the Fourier transform. The log-modulation domain representation is also referred to as the Mellin domain representation, because $\check{f}(c)$ can be recognized as the Mellin transform of $f(t)$ with the complex argument $-jc - 1/2$ (Cohen, 1995; Dautray and Lions, 1990). The term log modulation becomes clear by rewriting the eigenfunction $\phi_c(t)$ as

$$\phi_{c}(t) = \frac{1}{\sqrt{2\pi}} \frac{\mathrm{e}^{jc \log t}}{\sqrt{t}}, \tag{2.35}$$

which is recognized as a complex exponential with a logarithmic t-axis (see also Figure 2.2).

The eigenfunctions of equation (2.33) are given in the time domain. Upon comparison of equation (2.34) with the first of equations (2.22) one directly recognizes the unitary transformation kernel from time to time, time to frequency, and time to Mellin, respectively. By explicitly defining the unitary transformation kernels[19]

$$\begin{aligned} \mathrm{U}_{\mathcal{T}}(t', t) &= \delta(t - t') \\ \mathrm{U}_{\mathcal{F}}(\omega, t) &= \frac{1}{\sqrt{2\pi}}\, \mathrm{e}^{j\omega t} \\ \mathrm{U}_{\mathcal{C}}(c, t) &= \frac{1}{\sqrt{2\pi}}\, t^{jc-1/2}. \end{aligned} \tag{2.36}$$

[19] These are of course equal to the temporal domain eigenfunctions given in equation (2.33).

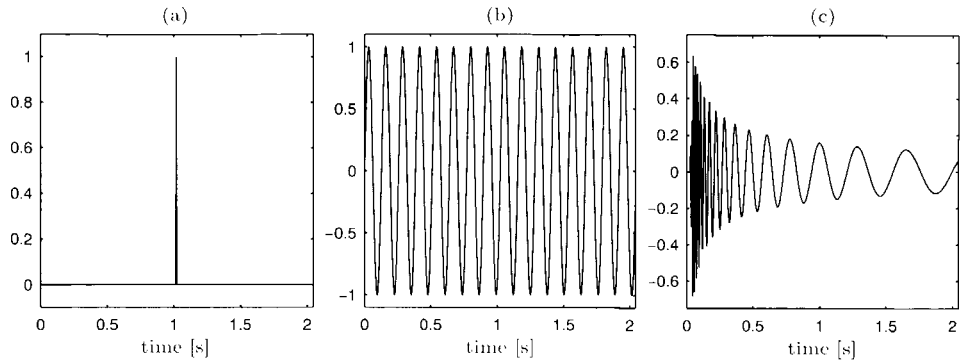


Fig. 2.2 *The eigenfunctions of belonging to the time, frequency and Mellin representation. (a) δ-distribution. (b) Real part of a complex exponential. (c) Real part of a Mellin eigenfunction. Note the logarithmic course of the modulation frequency.*

the correspondence with equations (2.22) is even more apparent:

$$\begin{aligned} f(t') &= (\mathcal{U}_{\mathcal{T}} f)(t') \\ \tilde{f}(\omega) &= (\mathcal{U}_{\mathcal{F}} f)(\omega) \\ \breve{f}(c) &= (\mathcal{U}_{\mathcal{C}} f)(c), \end{aligned} \tag{2.37}$$

where $\mathcal{U}_{\mathcal{T}}$, $\mathcal{U}_{\mathcal{F}}$ and $\mathcal{U}_{\mathcal{C}}$ are the transformation unitary operators corresponding to the kernels $\mathrm{U}_{\mathcal{T}}(t',t)$, $\mathrm{U}_{\mathcal{F}}(\omega,t)$ and $\mathrm{U}_{\mathcal{C}}(c,t)$ respectively. In a similar way, following equation (2.24) the representation of an operator $\mathcal{K}$ in the time domain, $\mathrm{K}(t,t')$ can be transformed according to

$$\begin{aligned} \mathrm{K}(t,t') &= \mathcal{U}_{\mathcal{T}} \mathrm{K}\, \mathcal{U}_{\mathcal{T}}^{-1} \\ \tilde{\mathrm{K}}(\omega,\omega') &= \mathcal{U}_{\mathcal{F}} \mathrm{K}\, \mathcal{U}_{\mathcal{F}}^{-1} \\ \breve{\mathrm{K}}(c,c') &= \mathcal{U}_{\mathcal{C}} \mathrm{K}\, \mathcal{U}_{\mathcal{C}}^{-1}. \end{aligned} \tag{2.38}$$

Let me discuss here some obvious features of the representations associated with these unitary transformations. Figure 2.2 shows (symbolically) a δ-distribution, a complex exponential and a Mellin eigenfunction. The time and frequency eigenfunctions are extremes, in the sense that the former has an infinitely small temporal influence area and the latter has an infinitely long temporal influence area and the other way around in the frequency representation. Another way to interpret the eigenfunctions is by introducing a time-frequency plane (reminiscent of the phase-space in classical or quantum mechanics) and by inspecting how the eigenfunctions cover it. Figure 2.3 shows the time-frequency distribution of single eigenfunctions. Note that the Mellin eigenfunction has a symmetry axis. This specific symmetry was to be expected on the basis of the symmetric structure of the associated self-adjoint

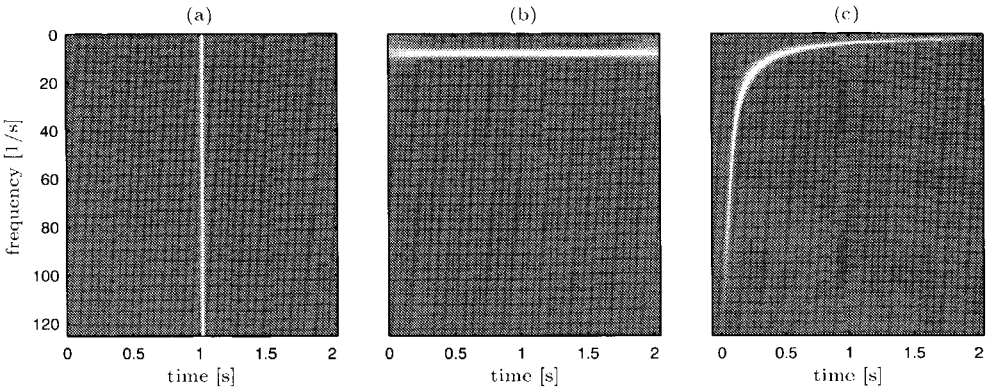


Fig. 2.3 *The coverage in the time-frequency plane of the time (left), frequency (middle) and log-modulation or Mellin eigenfunctions (right). The symmetry of the self-adjoint operator associated with the latter is also clear from the symmetry in the time-frequency plane.*

operator $\mathcal{C}$ in equation (2.31). More properties of the representations are discussed below.

Invariance and covariance properties —Using the apparatus developed in section 2.2.1, the invariance and covariance properties of the transformations can be inspected. Invariance properties of a physical variable become mathematically clear if the commutator of the governing self-adjoint operator and a parameterized unitary operator equals zero. A number of straightforward parameterized unitary operators can be made, following equation (2.28):

Frequency shift/modulation operator	$\mathcal{M}_a = e^{ja\mathcal{T}}$	$(\mathcal{M}_a f)(t) = e^{jat} f(t)$
Time shift operator	$\mathcal{S}_b = e^{-jb\mathcal{F}}$	$(\mathcal{S}_b f)(t) = f(t-b)$
Dilation operator	$\mathcal{E}_\sigma = e^{j \log\sigma \mathcal{C}}$	$(\mathcal{E}_\sigma f)(t) = \frac{1}{\sqrt{\sigma}} f\left(\frac{t}{\sigma}\right).$

The fact that the log-modulation operator $\mathcal{C}$ is the infinitesimal generator of the dilation operator[20], is the most important reason for introducing it. The invariance properties are now readily obtained, they are given by

$$[\mathcal{M}_a\,,\,\mathcal{T}] = 0 \qquad [\mathcal{S}_b\,,\,\mathcal{F}] = 0 \qquad [\mathcal{E}_\sigma\,,\,\mathcal{C}] = 0, \tag{2.39}$$

which means that, up to a phase factor, the temporal representation of a function is invariant under modulations, the spectrum of a function is invariant under translations, and the Mellin representation of a signal is invariant under dilations or scale

[20]Note that the $\mathcal{E}$ of expansion is used as a symbol for the dilation operator.

changes. More mathematically, this denotes (Baraniuk and Jones, 1995)

$$|(\mathcal{U}_T \mathcal{M}_a f)(t')| = |(\mathcal{U}_T f)(t')| \tag{2.40}$$

$$|(\mathcal{U}_\mathcal{F} \mathcal{S}_b f)(\omega)| = |(\mathcal{U}_\mathcal{F} f)(\omega)| \tag{2.41}$$

$$|(\mathcal{U}_C \mathcal{E}_\sigma f)(c)| = |(\mathcal{U}_C f)(c)|. \tag{2.42}$$

The invariance properties of the time and frequency representation are well-known. The invariance of the Mellin representation under scale changes of the original signal is less well-known. Altes (1978) uses it in combination with the Fourier transform to analyze mammalian hearing. Cohen (1993, 1995) pays a lot of attention to it. He calls the Mellin representation the scale representation. In view of its invariance under scale changes or dilations, Cohen chose an awkward terminology. On the other hand, just because of their invariance under scaling the Mellin eigenfunctions are useful objects in the analysis of self-similar phenomena, like (multi)fractals; a conjecture worked out in more detail by Herrmann (1997). In chapter 5, the Mellin transform will be used again while analyzing the wavelet transform of (scale invariant) homogeneous distributions.

Covariance properties follow from other commutation relations. The commutators of the frequency operator and the frequency shift, and of the time operator and the time shift, read

$$[\mathcal{S}_b\,,\,\mathcal{T}] = b\mathcal{S}_b \quad \text{and} \quad [\mathcal{M}_a\,,\,\mathcal{F}] = a\mathcal{M}_a, \tag{2.43}$$

which leads to the following covariance properties

$$(\mathcal{U}_T \mathcal{S}_b f)(t') = (\mathcal{S}_b \mathcal{U}_T f)(t') = (\mathcal{U}_T f)(t' - b) \tag{2.44}$$

$$(\mathcal{U}_\mathcal{F} \mathcal{M}_a f)(\omega) = (\mathcal{M}_a \mathcal{U}_\mathcal{F} f)(\omega) = (\mathcal{U}_\mathcal{F} f)(\omega - a). \tag{2.45}$$

The last two relations make again clear that $f(t')$ measures temporal content and $\tilde{f}(\omega)$ measures spectral content.

Cohen (1995) and Baraniuk and Jones (1995) discuss more properties of the different representations on the basis of the operator approach. Moreover, Baraniuk and Jones introduce in a consistent way unitarily equivalent self-adjoint operators $\bar{\mathcal{A}}$ (cf. equation 2.25) and a number of derived unitarily equivalent operators: unitarily equivalent parameterized unitary operators $\exp(ja\bar{\mathcal{A}})$, and unitarily equivalent transformation unitary operators $\bar{\mathrm{U}}_{\bar{\mathcal{A}}}$. This approach opens up a whole class of representations, which covers the time-frequency plane as introduced in Figure 2.3 in a different adaptive way. However, the coverage of the time-frequency plane is still carried out with eigenfunctions of infinite length in this plane. Is it possible to work with more localized decomposition functions? This question will be addressed in the next section.

2.3 Representations and transformations related to coherent states

Gabor (1946) dealt with the question on more localized decomposition functions. He worked in the field of communication theory and acoustics. He was aware of the attractive properties of the frequency domain representation Fourier had provided. At the same time Gabor realized, however, that Fourier integrals, which consider intervals "sub specie aeternatis", are far from our every day point of view. He realized that measuring instruments, but also the human ear, sense information in both time and frequency together. A musical tone, for example, can be discriminated by the human ear on the basis of the moment it is played but at the same time two tones can be discriminated on the basis of their pitch. A musical score is based on this idea. Different tones can be played at different times. It was also realized in the time of Gabor that there are certain limitations on the simultaneous discrimination of the time and frequency of a tone. A tone of a certain pitch requires a certain minimum amount of playing time in order to be recognized as that specific tone. Although it could have been a limitation of the human ear, Gabor showed[21] that the limitations on a simultaneous time-frequency analysis are fundamental. He derived the time-frequency equivalent of Heisenberg's uncertainty relation by stating that a signal cannot be characterized infinitely accurate by its time and frequency localization together. Given this fundamental uncertainty, Gabor derived the set of functions for which the smallest uncertainty is obtained. These functions are translated and modulated Gaussians, according to

$$g_{ab}(t) = (\mathcal{M}_a \mathcal{S}_b g)(t) = e^{jat} g(t-b), \tag{2.46}$$

where $g(t)$ is a normalized Gaussian function. For all values of a and b the functions $g_{ab}(t)$ occupy a minimum area with equal shape and size in the time-frequency plane[22]. Gabor intuitively understood that a set of functions $\{g_{ab}(t)\}$ with a, b continuously varying, forms a redundant family. By choosing $a = ma_0$ and $b = nb_0$, he assumed that he could completely cover the time-frequency, i.e. that each function f could be uniquely defined by the expansion coefficients on the discrete lattice[23]

$$\langle f, g_{ma_0, nb_0} \rangle . \tag{2.47}$$

Among the applications for his newly defined functions, Gabor especially focussed on the accurate and efficient representation of the receiver characteristics of the ear.

[21] At this point Gabor (1946) gives credits to Stewart (1931) who after a suggestion of the theoretical physicist A. Landeé brought the possibility of an acoustical version of Heisenberg's uncertainty relation to the awareness of the acousticians.

[22] The notion minimum area should be interpreted in the sense of an energy concentration. This notion will be elucidated on page 33.

[23] Bastiaans (1980) showed that the lattice Gabor chose, does not lead to a stable reconstruction.

With the benefit of the hindsight, it is easy to say that the constant-Q shape of the aural filter can better be represented by constant-Q filters, like wavelets, than by Gabor functions, which do not have a constant-Q behavior.

The long introduction on the work of Gabor serves two goals: (1) it honors a great paper; (2) it essentially introduces all the ingredients for a joint time-frequency analysis. The ingredients are completeness, reconstruction and physical relevance of particular divisions of the time-frequency plane. These ingredients can best be dealt with in the framework of coherent states. Although coherent states have their origin in quantum physics (Glauber, 1963), the applicability of the associated theory is elucidating enough to introduce them. The general theory of coherent states will be dealt with in section 2.3.1. Section 2.3.2 deals with translated and modulated Gaussians, giving rise to the Gabor representation. In section 2.3.3 affine coherent states will be dealt with, leading to the wavelet representation.

2.3.1 General theory

In their book, Klauder and Skagerstam (1985) give an extensive account of the developments related to coherent states. Sets of coherent states share the following properties:

1. *Continuity* . The state or vector g_a is a continuous function of the (possibly more dimensional) label a. Continuity denotes that when $a - a' \to 0$, it is required that $g_a \to g_{a'}$. This requirement typically excludes all basis sets of eigenfunctions related to self-adjoint operators[24].

2. *Completeness (Resolution of identity).* There exists an integration measure δa such that

$$f = \int \langle f, g_a \rangle \, g_a \, \delta a \tag{2.48}$$

for any f in an appropriately chosen Hilbert space. Equation (2.48) has to be understood in the sense of distributions. Though (2.48) seems to be equivalent to the closure relation (2.15), which shows a reconstruction from an expansion in a complete set of orthogonal eigenfunctions, the lack of orthogonality of the coherent states makes expansion (2.48) notably different.

Let me briefly list a number of properties that readily follow from the aforementioned requirements and that can be directly applied to the continuous Gabor expansions and the continuous wavelet expansions that follow.

[24]In the case of a discrete spectrum the labels are not continuous. In the case of a continuous label the eigenfunctions are still not coherent states because they will always be orthogonal, and hence g_a will not approach $g_{a'}$ infinitely close.

1. Coherent states can always assumed to be normalized, i.e.

$$\langle g_a, g_a \rangle = 1, \tag{2.49}$$

which is not true for generalized eigenfunctions. This is a physically appealing property. One can thus associate states with physical phenomena.

2. Just as $f(a) = \langle f, g_a \rangle$ is the functional representation of the measurable function f in the coherent states, an operator $\mathcal{K}$ can be given a functional representation according to

$$\mathrm{K}(a, a') = \langle \mathcal{K} g_{a'}, g_a \rangle , \tag{2.50}$$

which is again reminiscent of the representation of an operator in an orthogonal basis of eigenfunctions (see equation 2.17).

3. Due to the resolution of identity (2.48), the representation of a function f in the coherent state set g_a can be written as

$$\langle f, g_a \rangle = \int \langle f, g_{a'} \rangle \, \mathrm{R}(a', a) \delta a', \tag{2.51}$$

where $\mathrm{R}(a', a) \triangleq \langle g_{a'}, g_a \rangle$ is the so-called reproducing kernel. In the case of an orthonormal basis, equation (2.51) would represent a trivial identity placing no extra restrictions on the representations in the transformed domain. However, due to the continuity of g_a, the inner product $\langle g_{a'}, g_a \rangle$ is generally non-zero in a neighborhood of a' around a. Consequently, the representation $f(a') = \langle f, g_{a'} \rangle$ cannot be chosen arbitrarily; it has to be chosen 'admissible'. It means that the space of possible functions $f(a)$ is in general larger than the space of projections $\langle f, g_a \rangle$.

4. A set of coherent states is highly redundant. Any single coherent state can be expressed as a linear combination of the other coherent states, according to

$$g_a = \int \langle g_a, g_{a'} \rangle \, g_{a'} \delta a'. \tag{2.52}$$

The redundancy can be utilized to look for suitable subsets which still represent the whole space. In section 2.5 discrete subsets will be discussed.

How can coherent states be constructed? They can be obtained by applying parameterized unitary operators $\mathcal{U}$ of an irreducible representation of a Lie group to a ground state or fiducial vector (Perelomov, 1972)[25]. I will limit the exposition to

[25] Roughly speaking, a representation is irreducible, if for any $g \neq 0$, there does not exist nontrivial f orthogonal to all the $\mathcal{U}g$ (Messiah, 1961; Daubechies, 1992).

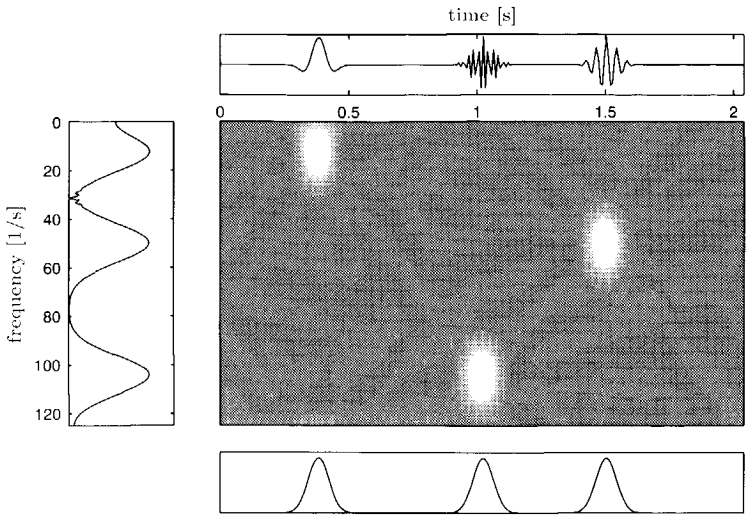


Fig. 2.4 *Time-frequency distribution of three Weyl-Heisenberg coherent states or Gabor functions. The top picture shows the three individual states (only the real parts) in the time domain. The left picture shows their power spectrum. The middle picture shows their position and shape in the time-frequency plane. The bottom picture shows the envelope of the three states. Note that only positive frequencies are considered in this example.*

two types of coherent states. One is generated by applying translations and modulations to a ground state, according to the ideas of Gabor (1946), but now placed in the new coherent-state framework. Within this framework they will be called Weyl-Heisenberg coherent states. The other one is generated by applying translations and dilations to a ground state, yielding wavelets. Within the coherent-state framework, wavelets are called affine coherent states. For details on the group theoretical background of both types of coherent states, the reader is especially referred to Heil and Walnut (1989).

2.3.2 Weyl-Heisenberg coherent states

The Weyl-Heisenberg coherent states are the quantum mechanical equivalents of Gabor's proposal for joint time-frequency analysis functions. They are created by translations and modulations of a ground state, according to

$$\begin{aligned} g_{ab}(t) &\triangleq (\mathcal{G}_{ab}g)(t) \\ &= e^{jat}g(t-b), \quad (a,b) \in \mathbb{R}^2, \end{aligned} \tag{2.53}$$

where the parameterized unitary operator $\mathcal{G}_{ab}$ is defined by

$$\mathcal{G}_{ab} \triangleq \mathcal{M}_a \mathcal{S}_b, \tag{2.54}$$

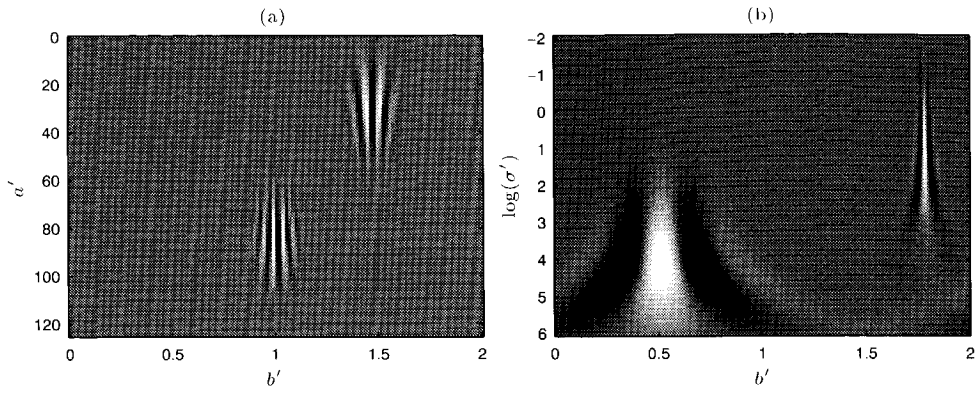


Fig. 2.5 *(a) Reproducing kernel* $\mathrm{R}_g(a', b'; a, b)$ *for the Gabor representation for two values of* a *and* b*. (b) Reproducing kernel* $\mathrm{R}_w(\sigma', b'; \sigma, b)$ *for two values of* σ *and* b *for the wavelet representation. Both kernels are continuous in their parameters.*

which is very closely related to the Weyl operator of quantum mechanics (Klauder and Skagerstam, 1985; Grossmann et al., 1986). It is the representation of the Weyl-Heisenberg group of phase-space translations with the following group multiplication rule

$$\mathcal{G}_{ab}\mathcal{G}_{a'b'} = C\mathcal{G}_{a+a',b+b'}, \tag{2.55}$$

where C is a phase factor depending on a, a', b, and b'. Following section 2.3.1, the ground state $g(t)$ is an arbitrary square-integrable function, which can be chosen to one's advantage. Figure 2.4 shows examples of Weyl-Heisenberg coherent states for a number of values a, b (here with $a > 0$) and for the most commonly used ground state both in signal analysis and quantum mechanics, the Gaussian function. Figure 2.4 shows the location of the coherent states in the time-frequency plane, which has been introduced in Figure 2.3, as well. All coherent states $g_{ab}(t)$ occupy areas of equal size and form, which is an important characteristic of these states in comparison with affine coherent states (cf. section 2.3.3).

The requirements given in the beginning of section 2.3.1 for $g_{ab}(t)$ to be coherent states, can be checked. The continuity in the labels a, b is easily seen by looking, for example, at the reproducing kernel $\mathrm{R}_g(a', b'; a, b)$, cf. equation (2.51), for a fixed value of a, b. Figure 2.5a shows two Gabor reproducing kernels for two values of a and b. Similarly, the resolution of identity can be checked. The decomposition of an arbitrary function into Gabor functions reads

$$\mathfrak{G}\{f, g\}(a, b) \triangleq \langle f, g_{ab} \rangle = \int f(t) g^*_{ab}(t)\, \mathrm{d}t. \tag{2.56}$$

This decomposition goes under the name Gabor transform. It is also referred to as the short-time or windowed Fourier transform. If there cannot be any confusion with

respect to the decomposition function, $\mathfrak{G}\{f,g\}(a,b)$ will also be denoted by $\hat{f}(a,b)$ and it will be called the Gabor representation of f. According to the resolution of identity, the reconstruction of f from its Gabor representation reads

$$f(t) = \int \hat{f}(a,b) g_{ab}(t)\, da\, db, \tag{2.57}$$

where $da\, db$ is the invariant integration measure δa of equation (2.48).

Properties of the Gabor representation —Let me discuss three important properties of the Gabor representation (for other properties the reader is referred to Vetterli and Kovačević (1995)):

- *Covariance under translations*
  The Gabor representation is covariant under translations in the time domain, according to

$$\begin{aligned} \langle \mathcal{S}_\xi f, \mathcal{M}_a \mathcal{S}_b g \rangle &= \langle f, \mathcal{S}_{-\xi} \mathcal{M}_a \mathcal{S}_b g \rangle \\ &= e^{ja\xi} \langle f, \mathcal{M}_a \mathcal{S}_{b-\xi} g \rangle \\ &= e^{ja\xi} \hat{f}(a, b-\xi). \end{aligned} \tag{2.58}$$

- *Covariance under modulations*
  The Gabor representation is covariant under modulations (or translations in the frequency domain), according to[26]

$$\begin{aligned} \langle \mathcal{M}_\xi f, \mathcal{M}_a \mathcal{S}_b g \rangle &= \langle f, \mathcal{M}_{-\xi} \mathcal{M}_a \mathcal{S}_b g \rangle \\ &= \langle f, \mathcal{M}_{a-\xi} \mathcal{S}_b g \rangle \\ &= \hat{f}(a-\xi, b). \end{aligned} \tag{2.59}$$

- *Localization*
  In order to better understand how the Gabor representation is subdividing the time-frequency plane, let me discuss in more detail the localization properties. Suppose $g(t)$ is concentrated around $t = 0$ and its Fourier transform $\tilde{g}(\omega)$ around $\omega = 0$, which can be expressed by

$$\bar{t} \triangleq \int t |g(t)|^2 \, dt = 0 \tag{2.60}$$

$$\bar{\omega} \triangleq \int \omega |\tilde{g}(\omega)|^2 \, d\omega = 0. \tag{2.61}$$

[26] By including an appropriate phase factor in the Gabor decomposition functions, the covariance properties for time and frequency translations can be symmetrized (Grossmann et al., 1986). In order to make it a correct representation of the Weyl-Heisenberg group, the phase factor is required.

The time width Δ_t and frequency width Δ_ω are defined by

$$\Delta_t^2 \triangleq \int (t - \bar{t})^2 |g(t)|^2 \, \mathrm{d}t \tag{2.62}$$

$$\Delta_\omega^2 \triangleq \int (\omega - \bar{\omega})^2 |\tilde{g}(\omega)|^2 \, \mathrm{d}\omega, \tag{2.63}$$

and they take some finite value. The quantities Δ_t and Δ_ω measure the energy concentration in the time and frequency domain. In equations (2.60)-(2.63) the function $g(t)$ is assumed to be normalized. By using a version of the proof of Schwartz' inequality due to Weyl (1931), Gabor (1946) was able to derive the time-frequency equivalent of Heisenberg's uncertainty relation:

$$\Delta_t \Delta_\omega \geq \frac{1}{2}. \tag{2.64}$$

This inequality says that it is not possible to simultaneously measure the time and frequency of a physical event with arbitrary precision. Equality is reached for a Gaussian function

$$g(t) = \left(\frac{1}{\gamma^2 \pi}\right)^{1/4} e^{-t^2/(2\gamma^2)} \tag{2.65}$$

and its Fourier transform

$$\tilde{g}(\omega) = \left(\frac{\gamma^2}{\pi}\right)^{1/4} e^{-\gamma^2 \omega^2/2}. \tag{2.66}$$

With this choice, the width in time and frequency become

$$\Delta_t = \frac{\gamma}{\sqrt{2}} \quad \text{and} \quad \Delta_\omega = \frac{1}{\sqrt{2}\,\gamma}, \tag{2.67}$$

respectively. The temporal and frequency localization of the coherent states are not affected by the modulations and translations. Hence, the shape and the width of the window functions do not change, but the shape of the actual analyzing functions does change (see Figure 2.4). Finally, note that in the limiting situation of $\gamma \to 0$, the Gabor representation reduces to the temporal representation, while for $\gamma \to \infty$ the Gabor representation goes to the Fourier representation.

2.3.3 Affine coherent states

The term affine coherent state has been introduced by Aslaksen and Klauder (1968, 1969). Affine coherent states are formed by applying parameterized unitary operators associated with the affine group to a ground state, i.e. by applying translations

and dilations to a ground state $\psi(t)$, according to

$$\begin{aligned}\psi_{\sigma b}(t) &\triangleq (\mathcal{A}_{\sigma b}\psi)(t) \\ &= \frac{1}{\sqrt{|\sigma|}}\psi\left(\frac{t-b}{\sigma}\right) \quad \sigma \in \mathbb{R}/\{0\}, b \in \mathbb{R}.\end{aligned} \tag{2.68}$$

where the parameterized unitary operator $\mathcal{A}_{\sigma b}$ is defined by $\mathcal{A}_{\sigma b} \triangleq \mathcal{S}_b\mathcal{E}_\sigma$. It is the representation of the affine or '$ax + b$'-group which fulfills the following group multiplication rule

$$\mathcal{A}_{\sigma b}\mathcal{A}_{\sigma' b'} = \mathcal{A}_{\sigma\sigma', b+\sigma b'}. \tag{2.69}$$

In contrast with the ground state $g(t)$ for the Weyl-Heisenberg coherent states, the ground state $\psi(t)$ for the affine coherent state can not be chosen arbitrarily square-integrable. It has to be chosen such that

$$0 < C_\psi \triangleq \int |\tilde{\psi}(\omega)|^2 \frac{\mathrm{d}\omega}{\omega} < \infty. \tag{2.70}$$

For sufficiently fast decaying functions $\psi(t)$ this condition reduces to

$$\int \psi(t)\,\mathrm{d}t = 0, \tag{2.71}$$

which is referred to as the admissibility condition. The fast decay in the time domain means regularity in the frequency domain. Regularity is required for equation (2.71) to be a sufficient condition. Albeit a limitation, equation (2.71) still allows for an amazing amount of affine coherent states. Various possibilities will be discussed in the subsequent section. From here onwards, affine coherent states defined according to equations (2.68) and (2.71) will be called wavelets. Figure 2.6 shows a number of wavelets for different values of σ and b. The ground state is here the second derivative of a Gaussian, the Mexican hat. Figure 2.6 shows their positions and shapes in the time-frequency plane as well. It is clear that, unlike the Gabor functions, wavelets are not of constant width. The width varies as a function of σ, while the shape remains fixed.

The requirements given in the beginning of section 2.3.1 for the wavelets $\psi_{\sigma b}(t)$ to be coherent states can be easily checked. The continuity in the labels can be intuitively understood by looking at the reproducing kernel $\mathrm{R}_w(\sigma', b'; \sigma, b)$, defined by equation (2.51). Figure 2.5b shows it for two values of the pair (σ, b). The proof of the resolution of identity can be found, for example, in Daubechies (1992). The decomposition of an arbitrary function into wavelets can now be uniquely written as

$$\mathfrak{W}\{f, \psi\}(\sigma, b) \triangleq \langle f, \psi_{\sigma b}\rangle = \int f(t)\psi^*_{\sigma b}(t)\,\mathrm{d}t. \tag{2.72}$$

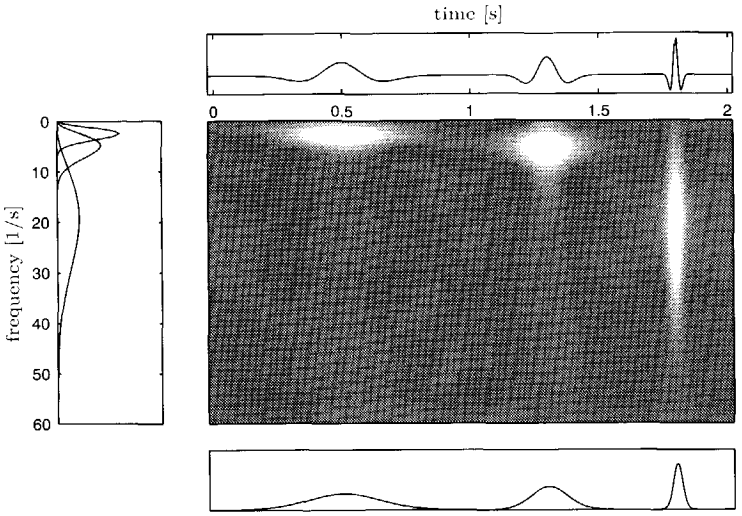


Fig. 2.6 *Affine coherent states or wavelets for three different values of σ and b. The top picture shows the three wavelets in the temporal domain. The left picture shows the power spectrum of the three wavelets. The middle picture shows the time-frequency distribution. The bottom picture contains the envelope of the three wavelets. Note that the shape of the wavelets remains fixed.*

This decomposition goes under the name wavelet transform or continuous wavelet transform. In the case that there is no confusion with respect to the choice of the wavelet ψ, I will represent $\mathfrak{W}\{f,\psi\}(\sigma,b)$ by $\check{f}(\sigma,b)$. In the remainder of the thesis, $\check{f}(\sigma,b)$ will be called the wavelet representation. According to the resolution of identity, the reconstruction of f from its wavelet representation reads

$$f = C_\psi^{-1} \int \check{f}(\sigma,b)\psi_{\sigma b}(t)\frac{\mathrm{d}\sigma\,\mathrm{d}b}{\sigma^2}, \tag{2.73}$$

where $\frac{\mathrm{d}\sigma\,\mathrm{d}b}{\sigma^2}$ is the integration measure δa of equation (2.48).

Remark 2.1 *The decomposition and subsequent composition do not strictly require L^2-normalized wavelets as given in equation (2.68). Kaiser (1994) shows that similar decompositions and compositions can be carried out for wavelets in L^p, which are of the form*

$$\frac{1}{|\sigma|^{1/p}}\psi\left(\frac{t-b}{\sigma}\right). \tag{2.74}$$

The decomposition of equation (2.72) now takes the form

$$\mathfrak{W}\{f,\upsilon\}(\sigma,b) = \int f(t)\frac{1}{|\sigma|^{1/p}}\upsilon^*\left(\frac{t-b}{\sigma}\right)\mathrm{d}t. \tag{2.75}$$

In the next section and in later chapters, this freedom will be used, to work for example with an L^1-normalization according to the choice $p = 1$.

Properties of the continuous wavelet representation —Let me discuss some basic properties of wavelets and the wavelet representation.

- *Covariance under translations*
  The wavelet representation $\check{f}(\sigma, b)$ of the function f is covariant under translations, according to

$$\begin{aligned} \langle \mathcal{S}_\xi f, \mathcal{S}_b \mathcal{E}_\sigma \psi \rangle &= \langle f, \mathcal{S}_{-\xi} \mathcal{S}_b \mathcal{E}_\sigma \psi \rangle \\ &= \check{f}(\sigma, b - \xi). \end{aligned} \tag{2.76}$$

  Note that this important property disappears for the discrete wavelet transform which will be discussed in section 2.5.

- *Covariance under scaling*
  The wavelet representation $\check{f}(\sigma, b)$ of f is covariant under scaling. This becomes clear by looking at the wavelet representation of a scaled function, that is

$$\begin{aligned} \langle \mathcal{E}_\xi f, \mathcal{S}_b \mathcal{E}_\sigma \psi \rangle &= \langle f, \mathcal{E}_{-\xi} \mathcal{S}_b \mathcal{E}_\sigma \psi \rangle \\ &= \langle f, \mathcal{S}_{b/\xi} \mathcal{E}_{-\xi} \mathcal{E}_\sigma \psi \rangle \\ &= \langle f, \mathcal{S}_{b/\xi} \mathcal{E}_{\sigma - \xi} \psi \rangle \\ &= \check{f}\left(\frac{\sigma}{\xi}, \frac{b}{\xi} \right). \end{aligned} \tag{2.77}$$

  Note that this behavior is opposite to that of the Mellin representation under scaling (see section 2.2.2). The Mellin representation is invariant under scale changes of the original function. The derived property helps in understanding the integration measure $\frac{d\sigma\, db}{\sigma^2}$ in the reconstruction of equation (2.73). Consider the contribution of f in the region $\sigma \in (\sigma_0, \sigma_0 + \epsilon)$ and $b \in (b_0, b_0 + \epsilon)$, see Figure 2.7. Under dilation of the function f, according to $f \mapsto \mathcal{E}_\xi f$, the same contribution goes to the region $\sigma \in (\sigma_0/\xi, \sigma_0/\xi + \epsilon/\xi)$ and $b \in (b_0/\xi, b_0/\xi + \epsilon/\xi)$. Consequently, the integration measure $\frac{d\sigma\, db}{\sigma^2}$ is such that elementary squares contribute equal energy to the inverse continuous wavelet transform (Vetterli and Kovačević, 1995).

- *Localization*
  A better understanding of the position of the scale parameter σ with respect to the frequency parameter ω can be obtained by looking at the time and frequency localization of the function $\psi_{\sigma b}(t)$. If ψ is localized around $t = 0$ and $\tilde{\psi}(\omega)$ around $\omega = \omega_0$, and if their widths in the time and frequency domain

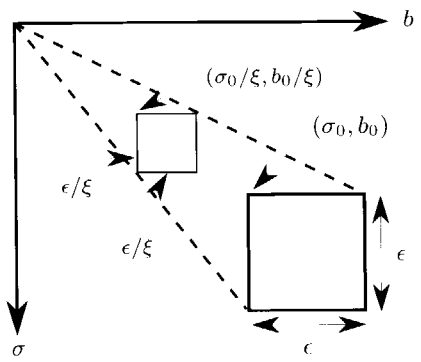


Fig. 2.7 *The contribution of a function f to its wavelet representation $\check{f}(\sigma, b)$ around the square at (σ_0, b_0) is equal to the contribution of $\mathcal{E}_\xi f$, with $\xi > 0$ around the square at $(\sigma_0/\xi, b_0/\xi)$ (Vetterli and Kovačević, 1995).*

are given by Δ_{t_0} and Δ_{ω_0}, respectively (see equations 2.62 and 2.63), then $\psi(t)$ is effectively supported in the time-frequency range

$$t \in \left(-\frac{\Delta_{t_0}}{2}, \frac{\Delta_{t_0}}{2}\right) \qquad \omega \in \left(\omega_0 - \frac{\Delta_{\omega_0}}{2}, \omega_0 + \frac{\Delta_{\omega_0}}{2}\right).$$

What happens with the time-frequency localization if the basic wavelet is translated and dilated? Due to the straightforward translation covariance of wavelets given in equation (2.76), translations can be easily dealt with. According to equation (2.77) dilations influence both the time and frequency location. For $\psi_{\sigma b}(t)$ the time and frequency localization become

$$t \in \left(\sigma b - \frac{\sigma\Delta_{t_0}}{2}, \sigma b + \frac{\sigma\Delta_{t_0}}{2}\right) \quad \omega \in \left(\frac{\omega_0}{\sigma} - \frac{1}{2}\frac{\Delta_{\omega_0}}{\sigma}, \frac{\omega_0}{\sigma} + \frac{1}{2}\frac{\Delta_{\omega_0}}{\sigma}\right), \tag{2.78}$$

respectively. Expressions (2.78) make clear that for increasing scale σ the temporal localization decreases and the spectral localization increases. Given the time-frequency localization properties it is interesting to see which part of the wavelet representation $\check{f}(\sigma, b)$ is influenced by the function $f(t)$ for $t = t_1$? The cone of influence of the point t_1 is the set of points (σ, b) that satisfies

$$\{(\sigma, b) \in \mathbb{R}^2 \mid |t_1 - b| < \frac{1}{2}\sigma\Delta_{t_0}\}. \tag{2.79}$$

It is the set of points (σ, b) for which $\check{f}(\sigma, b)$ is influenced by the value of $f(t)$ at $t = t_1$ (Mallat and Hwang, 1992). The cone of influence turns out to be an important quantity in the analysis of singularities which will be discussed briefly in section (2.4) and in chapter 5 in full detail.

Another aspect that directly follows from expression (2.78) is the constant-Q property. It expresses that the ratio of bandwidth and central frequency is unaltered by translations or dilations, according to

$$\frac{1}{Q} \triangleq \frac{\Delta_\omega}{\omega} = \frac{\Delta_{\omega_0}/\sigma}{\omega_0/\sigma} = \frac{\Delta_{\omega_0}}{\omega_0}. \tag{2.80}$$

This property is just another reason to consider the wavelet representation for the analysis of seismic data, because the constant-Q behavior is experimentally observed in seismic reflection data (see, e.g. Engelhard, 1996; Kjartansson, 1979, and references therein).

2.4 Understanding the continuous wavelet transform

In order to better understand the continuous wavelet transform and to give a sound motivation for forthcoming applications, let me investigate a general physical situation where wavelets come into play. The general physical situation deals with measurements of a physical variable and it is directly related to the theory of distributions introduced by Schwartz (1951, 1952). Consider a measurement of a physical variable obtained with a measuring device with a non-vanishing support. A question concerning the dependence of the physical variable on the location of the measurement or its dependence on the size of measuring device will naturally lead to the wavelet transform of the physical variable. In section 2.4.1 the link between distributions and the wavelet transform will be explored. In section 2.4.2 a number of properties will be discussed which determine the ability of wavelets to analyze local points of rapid variations. These properties turn out to be their degree of differentiability and their number of vanishing moments.

2.4.1 Wavelets and measurements

Normally, a physical variable is represented by a function $f(t)$. This notation suggests that the physical variable is known for any t. However, due to the non-vanishing support of a measuring device, a quantity can never be known for any t. Its average behavior at and around t is known. This observation, among other observations, made Schwartz (1951, 1952) decide to introduce the concept functional instead of a function. In addition, Schwartz introduced the concept test function to represent a measuring device. The measuring process is given by the action of the functional on the test function; it is called the testing operation. Hence, the functional gets a meaning only if it is tested against the test function.

Example 2.3 *Consider as a physical variable the temperature in a room. Measuring the temperature at a certain location implies that a thermometer, which has*

a non-vanishing size, has to be placed in the room. The thermometer carries out a local averaging process over its size. It can consequently be seen as a test function. The temperature in the room can be seen as the functional, and the measured temperature can be interpreted as the outcome of the test operation.

Following Zemanian (1965), I denote the functional by f. The complex number that the functional f assigns to a particular test function ϕ is designated (f, ϕ). Here, linear and continuous functionals are considered[27], which are referred to as distributions or generalized functions (the latter term has been coined by Gel'fand and Shilow (1960)). Two types of distributions are generally discriminated: regular and singular distributions. Regular distributions are locally integrable. For regular distributions the testing operation is explicitly given by an integral according to

$$(f, \phi) = \int f(t)\phi(t)\,\mathrm{d}t. \tag{2.81}$$

Singular distributions are not locally integrable. For singular distributions the notation on the right-hand side of equation (2.81) can be utilized, but in that case it is a symbolic notation. A well-known example of a singular distribution is the δ-distribution. The action of the δ-distribution on a test function is often expressed by

$$\phi(0) = \int \delta(t)\phi(t)\,\mathrm{d}t, \tag{2.82}$$

where the right-hand side is a symbolic notation: expression (2.82) has a meaning via its left-hand side and not via its right-hand side. Other singular distributions which can be given a meaning in the sense of distributions are (multi)fractals. Herrmann (1997) pays extensive attention to the proper distributional interpretation of multifractals.

The test function $\phi(t)$ is assumed to be real and sufficiently smooth[28]. Moreover, it is assumed that it has a non-vanishing and finite zeroth moment, i.e.

$$0 < \int \phi(t)\,\mathrm{d}t < \infty. \tag{2.83}$$

In other words, $\phi(t)$ has a DC-component. Since $\phi(t)$ is a representation of a measuring device, it is generally of compact support or rapidly decaying. In the wavelet literature the function $\phi(t)$ is referred to as the scaling or smoothing function.

The outcome of a measurement or a testing operation depends on the characteristics of the distribution, but also on the test function, most notably on its

[27] A *linear* functional fulfills $(f, \phi_1 + \phi_2) = (f, \phi_1) + (f, \phi_2)$. A *continuous* functional fulfills $(f, a\phi) = a(f, \phi)$.

[28] Sufficient is a relative notion: the more irregular the distribution, the more regular the test function should be.

resolution or scale and on its location in time or space. A series of measurements in the variable t at a specific scale σ can be written as (cf. equation 2.81)[29]

$$f(\sigma,t) = (f, \frac{1}{\sigma}\phi\left(\frac{t'-t}{\sigma}\right)) = \int f(t')\frac{1}{\sigma}\phi\left(\frac{t'-t}{\sigma}\right)\,\mathrm{d}t', \tag{2.84}$$

where an L^1-normalization is utilized for the scaling function. The tested distribution $f(\sigma,t)$ is an approximation of the distribution f at the scale σ. It is also called a regularized distribution. It is not the wavelet transform of f, hence the check $\check{}$ is correctly omitted. Equation (2.84) introduces the mathematical formulation of a scale or resolution dependent physical measurement. A measurement at a single time t or a single scale σ can be of great value, however one is often interested in how the measurement changes in the course of time or how it would have been if taken at a different scale σ. The next two paragraphs discuss what happens in those situations.

The changes as a function of time become apparent by looking at the nth order derivative ∂_t^n of $f(\sigma,t)$, which is allowed for $\phi(t)$ sufficiently smooth. Consider the nth order derivative of the scaling function $\phi(t)$, according to

$$\psi^{(n)}(t) = \partial_t^n \phi(t).$$

It can be easily seen that $\psi^{(n)}$ is a wavelet, by checking that its Fourier transform vanishes at $\omega = 0$, according to $\tilde{\psi}(0) = 0$. It can now be derived that the changes as function of time in $f(\sigma,t)$ are directly related to the wavelet transform of the distribution f, according to

$$\begin{aligned}\mathfrak{W}\{f,\psi^{(n)}\}(\sigma,t) &= \int f(t')\frac{1}{\sigma}\psi^{(n)}\left(\frac{t'-t}{\sigma}\right)\,\mathrm{d}t' \\ &= (-\sigma)^n\partial_t^n f(\sigma,t).\end{aligned} \tag{2.85}$$

Consequently, the nth derivative with respect to t of $f(\sigma,t)$ corresponds to the wavelet representation of the distribution f with the nth derivative of ϕ as analyzing wavelet. Figure 2.8 illustrates the relation between a distribution, a tested distribution and the wavelet representation at a particular scale. It is clear that the wavelet representation senses changes in the distribution. Because of the fact that essential information in a signal or image is generally carried by its points of rapid variation, i.e. at points where the underlying physical process or physical system is changing, the wavelet representation is of natural interest (Mallat and Hwang, 1992; Mallat and Zhong, 1992; Herrmann, 1997)[30]. It is able to delineate and characterize the points of rapid variations. The points of rapid variation are also referred to as singularities.

[29]Note that I adopt the distributional concepts of Schwartz (1951, 1952), see also Zemanian (1965), but that the notation is slightly altered in order to bring it more in line with the wavelet representation.

[30]See also the quote of Maxwell (1882) on page 15.

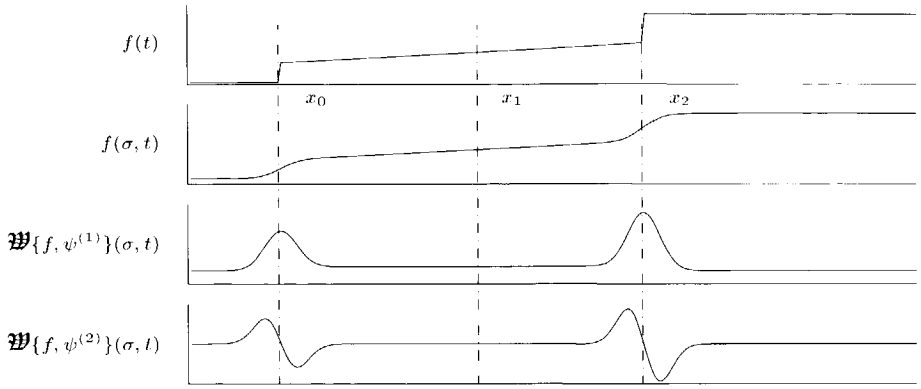


Fig. 2.8 *The top picture shows the distribution f. The measurement or testing operation yields the tested distribution $f(\sigma, t)$ at a particular scale. The test function (or scaling function in the wavelet literature) is chosen to be the Gauss function. The scaled first derivative with respect to time yields $\mathfrak{W}\{f, \psi^{(1)}\}(\sigma, t)$, which expresses how the distribution is changing as a function of time at the scale σ. The scaled second derivative $\mathfrak{W}\{f, \psi^{(2)}\}(\sigma, t)$ expresses how the distribution is changing in the scale direction, but it also expresses how $\mathfrak{W}\{f, \psi^{(1)}\}(\sigma, t)$ is changing as a function of time.*

Generally a measurement represented by (2.84) is taken at a single scale σ. Very well developed measurement devices for seeing or hearing, for example the human eye or ear, are, however, taking measurements at multiple scales (Marr, 1982; ter Haar Romeny, 1994). The reason for this is related to the fact that singularities are, besides location dependent, generally scale dependent as well (Nottale et al., 1997; Herrmann, 1997). The question arises how the measurement is exactly changing as a function of scale, i.e. what happens if we look at an infinitesimal different scale. Consider a function $\psi^H(t)$ given by $\psi^H(t) = (t\partial_t + 1)\phi(t)$. It can be easily seen that for $\phi(t)$ sufficiently smooth and sufficiently fast decaying, the function $\psi^H(t)$ is a wavelet. The wavelet transform with respect to this wavelet reads

$$\begin{aligned} \mathfrak{W}\{f, \psi^H\}(\sigma, t) &= \int f(t') \frac{1}{\sigma} \psi^H \left(\frac{t' - t}{\sigma} \right) \mathrm{d}t' \\ &= -\sigma \partial_\sigma f(\sigma, t) \\ &= -\frac{\partial f}{\partial \log \sigma}, \end{aligned} \tag{2.86}$$

which is a result due to Holschneider (1995). Equation (2.86) expresses that the logarithmic scale derivative (i.e. $\partial_{\log \sigma}$) of $f(\sigma, t)$ is equal to the wavelet transform of the distribution f at scale σ with ψ^H as analyzing wavelet. Generally, this result

holds only true for L^1-normalized scaling functions, however for a Gaussian scaling function, other normalizations yield equivalent results, which gives the Gaussian scaling function a unique position.

In addition, for a Gaussian scaling function $\phi(t)$, the wavelet $\psi^H(t)$ can also be written as

$$\psi^H(t) = -\frac{1}{2}\partial_t^2 \phi(t), \tag{2.87}$$

i.e. as a second order temporal derivative. In the wavelet literature the latter function is referred to as the Mexican hat (see, e.g. Daubechies, 1992) and in the geophysical literature it is referred to as a Ricker wavelet. Combining equations (2.85) and (2.86) I obtain the following equality

$$-\sigma\partial_\sigma f(\sigma, t) = -\frac{1}{2}\sigma^2 \partial_t^2 f(\sigma, t), \tag{2.88}$$

which expresses that changes in the (logarithmic) scale direction are equal to second order changes in the time direction. A remarkable relation that has a clear physical background, because if one substitutes $\sigma = \sqrt{4\tau}$ it turns out that

$$\partial_\tau f(\tau, t) = \partial_t^2 f(\tau, t), \tag{2.89}$$

which can be recognized as the diffusion equation. Hence, if one chooses a Gaussian smoothing function, the multiscale analysis fulfills the diffusion equation. From an image analysis point of view, the diffusion equation has a unique position (see, Koenderink (1984), ter Haar Romeny (1994), and chapter 6).

Concluding, any physical measurement should be accompanied by a notion of scale, because generally a measurement depends on the scale of the measuring device. Changes of the measurement due to changing the scale or due to changing the time naturally lead to the wavelet transform of the underlying distribution. In the case that a Gaussian smoothing function is chosen, it can be shown that the changes in the scale direction and in the time direction are closely related. The relation is expressed in the diffusion equation.

2.4.2 Decay, regularity and vanishing moments

As argued in the previous subsection points of rapid variation, the singularities, carry the important information in a signal. The strength of the singularity is estimated via the local degree of differentiability or regularity. The local degree of differentiability is given in terms of the Hölder exponent α. What requirements have to be imposed on a wavelet to be able to analyze a certain range of Hölder exponents? This question has been addressed in considerable detail by Muzy et al. (1993), Bacry et al. (1993), Staal (1995), and Herrmann (1997). Here, I will briefly discuss the main aspects. The requirements on the wavelet are expressed in terms of the regularity of the wavelet and the number of vanishing moments of the wavelet.

Consider a distribution of the form $f = g + (t - t_0)^\alpha$, $t > 0$. It consists of a relatively smooth background g and a distortion on top. The local degree of differentiability at $t = t_0$ is captured by the exponent α. For $\alpha < 1$ the distribution f gives rise to a singularity at $t = t_0$. Despite the fact that it comprises a small sample of possible singularities (namely of the algebraic type), it is generic enough to be discussed.

The ability of a wavelet to analyze local points of irregularity, i.e. the singularities, is based upon the fact that wavelets are able to "absorb" derivatives and/or integrals, provided that an appropriate wavelet is chosen. Appropriate is here defined in terms of its smoothness and number of vanishing moments. The smoothness or regularity of the wavelet $\psi(t)$ refers of course to the degree of differentiability. The wavelet $\psi(t)$ has M vanishing moments if

$$\int t^m \psi(t)\,\mathrm{d}t = 0 \quad \text{for} \quad m = 0, \ldots, M-1. \tag{2.90}$$

In the spectral domain regularity corresponds to decay behavior for $\omega \to \infty$, and the number of vanishing moments correspond to the decay behavior for $\omega \to 0$, respectively. If I introduce the operators

$$\begin{aligned} &\partial_t^m \quad m > 0 \quad \text{for an } m\text{th order differentiation,} \\ &\partial_t^m \quad m = 0 \quad \text{for the identity operator, and} \\ &\partial_t^m \quad m < 0 \quad \text{for an } m\text{th order integration,} \end{aligned} \tag{2.91}$$

then transferring derivatives or integrals, can be expressed by

$$\int \left(\partial_t^m f(t)\right) \frac{1}{\sigma}\psi\left(\frac{t-b}{\sigma}\right)\mathrm{d}t = (-1)^m \int f(t)\; \partial_t^m \frac{1}{\sigma}\psi\left(\frac{t-b}{\sigma}\right)\mathrm{d}t, \tag{2.92}$$

where, due to the decay properties of the wavelet, the stock terms vanish. Equation (2.92) is only valid if the wavelet ψ is sufficiently many times differentiable or integrable.

Example 2.4 *Suppose a distribution f has a local singularity at $t = t_0$ characterized by a Hölder exponent of $-1 < \alpha < 0$. For $m > 0$ the distribution $\partial_t^m f$ has a local negative singularity of $\alpha - m$ at t_0. To be able to analyze the distribution $\partial_t^m f$ via the wavelet transform, the m derivatives has to be passed over to the wavelet according to equation (2.92). Hence, the wavelet has to be sufficiently many times differentiable. On the other hand, for $m > 0$ the distribution $\partial_t^{-m} f$ has a local positive singularity of $\alpha + m$ at t_0. To be able to analyze the underlying singularity the integrations have to be passed over to the wavelet. Hence, the wavelet has to be sufficiently many times integrable which is guaranteed by choosing a wavelet with enough vanishing moments.*

temporal domain	spectral domain	analyzing effect
decay	regularity	resolving power
regularity	decay at $\omega \to \infty$	differentiability $\to$ negative singularities
vanishing moments	decay at $\omega \to 0$	integrability $\to$ positive singularities

Table 2.1 *Relation between wavelet properties in the temporal domain and the spectral domain, and their significance for the ability of wavelets to analyze singularities.*

It may be concluded that the choice of the wavelet determines the range of singularities that can be observed. On first thought, one is tempted to choose as much vanishing moments and as much smoothness as possible. However, both more vanishing moments and more smoothness deteriorate the decay behavior in the temporal domain. Consequently, there is a trade-off between the decay behavior on the one hand and the number of vanishing moments and the smoothness on the other hand. Table 2.1 summarizes the observations; for examples and more details the reader is referred to chapter 5 or to Herrmann (1997).

2.5 Frames and multiresolution approximations

2.5.1 General background

It may be concluded from the previous sections that the orthonormal representations related to self-adjoint operators, here especially the time representation and the frequency representation, are fundamentally different from the highly redundant coherent state representations, i.e. the Gabor representation and the wavelet representation. The redundancy of the coherent states suggests that a suitably chosen subset of the parameters defining the representation can be chosen such that the original function is still uniquely represented. In the general notation of section 2.3.1, the question is whether it is possible to find a $a_0 \in \mathbb{R}^n$ such that

1. the coefficients $\langle f, g_{ma_0} \rangle$, with m chosen in a suitable discrete set $J \in \mathbb{Z}^n$, are uniquely related to the function f. Hence, $\langle f_1, g_{ma_0} \rangle = \langle f_2, g_{ma_0} \rangle \ \forall m$ if and only if $f_1 = f_2$;

2. a reconstruction or dual function $\underline{g}_{ma_0}$ exists with the property that

$$f = \sum_m \langle f, g_{ma_0} \rangle \, \underline{g}_{ma_0}. \tag{2.93}$$

Equation (2.93) can be seen as the discrete version of equation (2.48).

It is important to note that (1) g_{ma_0} is not a set of coherent states anymore; (2) the discretization of the parameters of the coherent states does not necessarily imply

a discrete representation of f; both the coherent state representation of (2.48) and the representation (2.93) can be discretized in the parameters of f; (3) the so-called dual function $\underline{g}_{ma_0}$ is not necessarily equal to g_{ma_0}. The term dual function refers to the fact that $\underline{g}_{ma_0}$ can also be seen as the decomposition function and g_{ma_0} as the reconstruction function, according to

$$f = \sum_m \langle f, \underline{g}_{ma_0} \rangle \, g_{ma_0}. \tag{2.94}$$

The condition which guarantees that the two aforementioned issues can be positively answered is the frame condition:

Definition 2.2: Frames (Daubechies, 1992)
A family of functions $\{g_{ma_0}, m \in J\}$ *in a Hilbert space is called a frame if there exist* $A > 0$, $B < \infty$ *so that, for all* f *in the Hilbert space,*

$$A\|f\|^2 \leq \sum_{m \in J} |\langle f, g_{ma_0} \rangle|^2 \leq B\|f\|^2. \tag{2.95}$$

A and B are called the frame bounds.

If a family of functions $\{g_{ma_0}\}$ forms a frame, functions $\underline{g}_{ma_0}$ can be found such that (2.93) is true. The role of the inequalities in equation (2.95) can be intuitively understood. The second inequality denotes that the sum of coefficients $\langle f, g_{ma_0} \rangle$ should be finite. The first inequality is related to stability: for a stable reconstruction a logical requirement is that the coefficients $\langle f, g_{ma_0} \rangle$ are only zero for all $m \in J$ if $f = 0$.

Generally, a discrimination is made between four frame types (Daubechies, 1992).

1. For $A = B \neq 1$ the set $\{g_{ma_0}\}$ is called a tight frame. The reconstruction (2.93) reads

$$f = A^{-1} \sum_m \langle f, g_{ma_0} \rangle \, g_{ma_0}, \tag{2.96}$$

   at least in the weak sense. Hence the reconstruction or dual function and the decomposition function are related by $\underline{g}_{ma_0} = A^{-1} g_{ma_0}$. Note that a tight frame is generally not equal to an orthogonal basis. The value of A is a measure for the redundancy of a tight frame.

2. For $A = B = 1$ and $|g_{ma_0}| = 1$, $\{g_{ma_0}\}$ forms an orthonormal basis and consequently

$$f = \sum_m \langle f, g_{ma_0} \rangle \, g_{ma_0}. \tag{2.97}$$

3. For $A \neq B$, and B/A close to 1, $\{g_{ma_0}\}$ is called a snug frame. For snug frames $\underline{g}_{ma_0} \neq g_{ma_0}$, but they are close. The reconstruction can be written by

$$f = \frac{2}{A+B} \sum_m \langle f, g_{ma_0} \rangle \, g_{ma_0} + \text{small rest term.} \tag{2.98}$$

4. For $B/A \gg 1$, the dual function $\underline{g}_{ma_0}$ can be very different from the decomposition function[31]. For a larger ratio B/A, the reconstruction becomes less stable.

In section 2.5.2 the theory of frames will be applied to the Gabor representation. Section 2.5.3 is devoted to general wavelet frames. In section 2.5.4 the theory of multiresolution approximations will be introduced. A multiresolution approximation gives naturally rise to a special class of wavelet frames, namely orthonormal wavelet bases. The latter will lead to the discrete wavelet transform as opposed to the continuous wavelet transform discussed in section 2.3.3.

2.5.2 Gabor frames

For Weyl-Heisenberg coherent states the frame decomposition takes the following form (cf. equation 2.56)

$$\mathfrak{G}\{f, g\}(ma_0, nb_0) = \langle f, g_{mn} \rangle = \int f(t) \, g^*_{mn}(t) \, \mathrm{d}t, \tag{2.99}$$

where $g_{mn}(t) \triangleq e^{jma_0 t} g(t - nb_0)$. Here, I will work with a Gaussian window function $g(t)$ given in equation (2.65)[32]. The fact that Weyl-Heisenberg coherent states form a redundant set, was already known to Gabor (1946). The regular frame decomposition he chose, was based on the invariant shape of the functions $g_{ab}(t)$ in the time-frequency plane and on Nyquist sampling density[33]. For this choice, the first of the requirements on page 45 is indeed fulfilled. Hence, any square-integrable function is uniquely described by its expansion in Gabor coefficients. However, a stable reconstruction is not possible, because the dual function $\underline{g}_{mn}$ is not square-integrable (Bastiaans, 1980; Daubechies, 1990). Daubechies showed that for

$$a_0 b_0 \leq 2\pi \tag{2.100}$$

the set $\{g_{mn}\}$ constitutes a frame, the so-called Gabor frame. Loosely speaking, for $a_o b_0 > 2\pi$ the time-frequency plane is not sampled dense enough to represent all

[31] Daubechies (1990, 1992) shows that the dual function can be computed via the frame operator $\mathcal{T}$, according to $\underline{g}_{ma_0} = (\mathcal{T}^\dagger \mathcal{T})^{-1} g_{ma_0}$, where $\mathcal{T} f = \langle f, g_{ma_0} \rangle$ and its adjoint $\mathcal{T}^\dagger c = \sum_m c_m g_{ma_0}$.

[32] For other window functions, frames can be constructed as well (Daubechies, 1992).

[33] For a function with bandwidth $[-\Omega, \Omega]$ and with duration T, Nyquist sampling density is $\Omega T/\pi$, corresponding to the choice $a_0 b_0 = 2\pi$. In the sixties it was shown that Nyquist sampling density actually denotes the number of possible independent functions in this time-frequency area (see Daubechies, 1992).

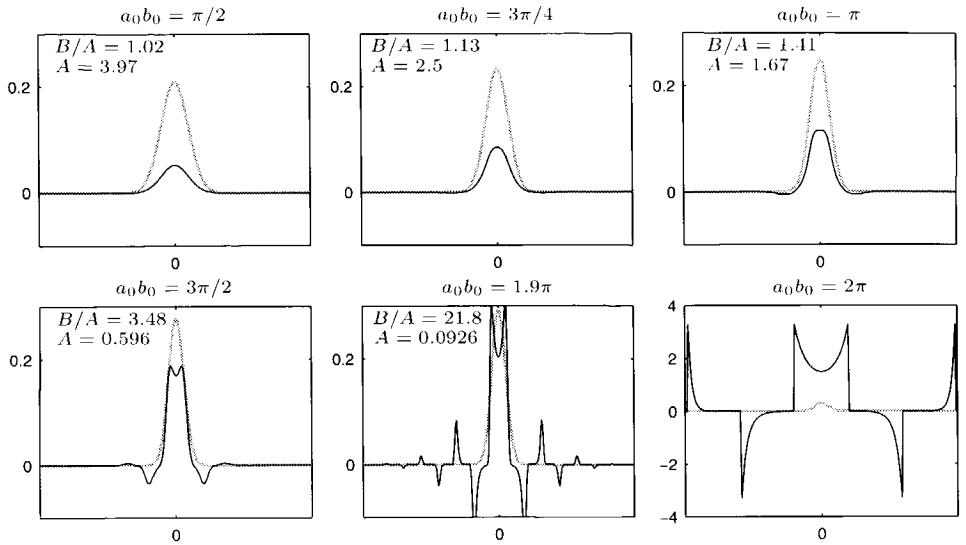


Fig. 2.9 *Reconstruction functions $\underline{g}_{00}$ for different values of a_0b_0, with $a_0 = \sqrt{a_0b_0}/\gamma$, hence $b_0 = \sqrt{a_0b_0}\,\gamma$. The gray lines denote the decomposition function g_{00}, which is chosen equal for the six values of a_0b_0.*

functions. The frame bounds A and B, and the dual function $\underline{g}_{mn}$ can be numerically estimated[31]. Figure 2.9 shows a number of numerically estimated dual functions. For a given value of γ in equation (2.65), and for a number of values for the product a_0b_0, the decomposition and corresponding reconstruction function are shown for $m = n = 0$. For smaller a_0b_0 the redundancy increases. The ratio $\frac{2\pi}{a_0b_0}$ denotes the redundancy of the frame. For a large redundancy a snug frame is obtained and the reconstruction or dual function is almost equal to the decomposition function. For the frame parameters chosen according to $a_0b_0 < 2\pi$, the set of functions $\{g_{mn}\}$ forms a stable frame and the reconstruction of f from its coefficients $\langle f, g_{mn}\rangle$ is possible according to equation (2.93).

The representation of an operator $\mathcal{K}$ in the Gabor domain reads (cf. equation 2.50)

$$\hat{\mathrm{K}}(m, n; m', n') = \langle \mathcal{K} g_{m'n'}, g_{mn} \rangle\,, \tag{2.101}$$

which can also be written as the transformation from a temporal (or spatial) representation to the Gabor domain representation (cf. equation 2.23), according to

$$\hat{\mathrm{K}}(m, n; m', n') = \int_t \int_{t'} g^*_{mn}(t)\, \mathrm{K}(t, t')\, g_{m'n'}(t')\, \mathrm{d}t\, \mathrm{d}t'. \tag{2.102}$$

[34]For $a_0b_0 = 2\pi/n, n \in \mathbb{N}, n > 1$ the frame bounds can be computed analytically (Daubechies, 1990).

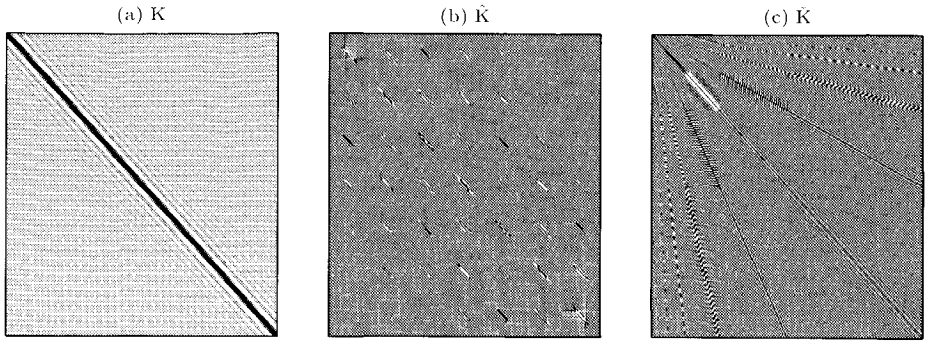


Fig. 2.10 *(a) Spatial representation of a monochromatic Green's operator. (b) Gabor domain representation of a monochromatic Green's operator. (c) Wavelet domain representation of a monochromatic Green's operator.*

For appropriately chosen frame parameters a_0 and b_0 the elements of the transformed kernel $\hat{\mathrm{K}}(m, n; m', n')$ completely represent the operator $\mathcal{K}$. For a monochromatic Green's operator in a homogeneous medium, the Gabor domain kernel is shown in Figure 2.10. The interpretation of the elements of $\hat{\mathrm{K}}(m, n; m', n')$ will be left to the next chapter.

2.5.3 Wavelet frames

For wavelet frames a general limitation as for the Gabor frames in equation (2.100) does not exist. Looking at the way wavelets cover the time-frequency plane, a natural choice for the frame parameters for wavelets is

$$\sigma = \sigma_0^m \quad b = nb_0\sigma_0^m, \quad \text{for} \quad m, n \in \mathbb{Z}, \tag{2.103}$$

with $\sigma_0 > 1$ and $b_0 > 0$. This yields for the wavelet representation of equation (2.72)

$$\mathfrak{W}\{f, \psi\}(\sigma_0^m, nb_0\sigma_0^m) = \langle f, \psi_{mn} \rangle = \int f(t)\psi_{mn}^*(t)\,\mathrm{d}t, \tag{2.104}$$

with

$$\psi_{mn}(t) \triangleq \sigma_0^{-m/2}\psi(\sigma_0^{-m}t - nb_0). \tag{2.105}$$

By adjusting the discretization interval of b to the actual value of σ, one takes into account that for larger σ the wavelet is more spread out than for smaller σ (cf. equation 2.78). Using the localization properties of section 2.3.3, the coverage of the time-frequency plane takes the form shown in Figure 2.11. Here, the scale parameter is chosen to vary according to $\sigma_0 = 2$, giving rise to a dyadic wavelet representation. It is compared with the regular time-frequency sampling natural for

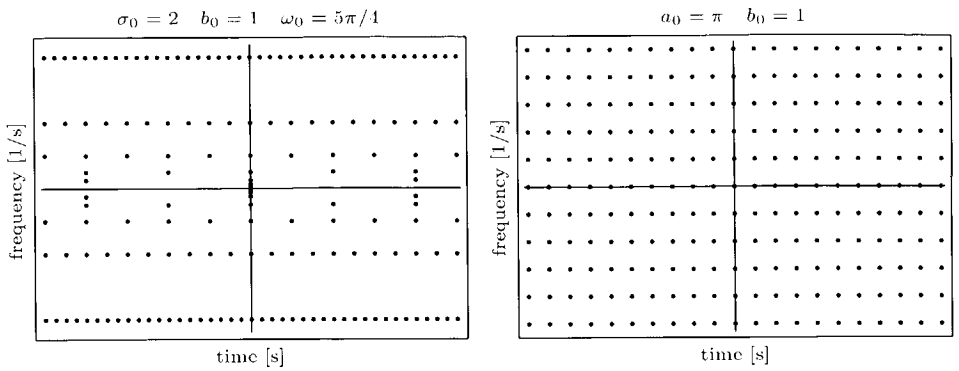


Fig. 2.11 *Natural coverage of the time-frequency plane for wavelet frames (left) and Gabor frames (right).*

Gabor frames. Daubechies (1990) showed that for a wide range of values for σ_0 and b_0 the wavelets $\psi_{mn}(t)$ constitute a wavelet frame, in which cases the reconstruction of f from its wavelet representation $\mathfrak{W}\{f, \psi\}(\sigma_0^m, nb_0\sigma_0^m)$ can be carried out.

For applications which require both efficiency and good temporal analysis properties, alternative frame samplings are sometimes used. Of special interest is a frame decomposition with a time sampling independent of the scale sampling, according to

$$\sigma = \sigma_0^m \qquad b = nb_0 \tag{2.106}$$

yielding a so-called non-downsampled dyadic wavelet transform (Mallat and Zhong, 1992; Saito, 1994). The efficiency gain with respect to the coherent state decomposition is obtained by coarsely sampling the scale parameter. In chapter 6 this specific discretization will be discussed in relation to an efficient 3-D image analysis procedure of migrated data.

In contrast to the Gabor frames with a Gaussian ground state, wavelet frames can be chosen such that $A = B = 1$, and hence, such that the set $\{\psi_{mn}\}$ forms an orthonormal basis. This property is known for a long time already. Whereas the term wavelets was only coined in the eighties, wavelets have been used in applied mathematics for a much longer time (see e.g. Meyer, 1993). One of the earliest examples is the Haar wavelet (Haar, 1910). The Haar wavelet is given by

$$\psi(t) = \begin{cases} 1 & 0 \leq t < 1/2 \\ -1 & 1/2 \leq t < 1 \\ 0 & \text{otherwise.} \end{cases}$$

If one chooses in addition $\sigma_0 = 2$ and $b_0 = 1$ in equation (2.105) then the set $\{\psi_{mn}\}$

forms an orthonormal basis for $L^2(\mathbb{R})$[35]. The characteristic features of the Haar wavelet are the discontinuity and the compactness of its support. The compactness is an attractive property; the discontinuity is often a disadvantage because of the limitation in analyzing power (see section 2.4.2). Other orthonormal wavelets were systematically pursued in the second half of the eighties. Within the theory of multiresolution approximations, orthonormal wavelets can be most easily dealt with.

Let me briefly describe the wavelet representation of an operator before I turn to the multiresolution approximations in the subsequent section. In the case that the frame parameters σ_0 and b_0 are chosen such that the set $\{\psi_{mn}\}$ constitutes a frame, an operator can be uniquely represented by (cf. equation 2.50 and 2.101)

$$\check{\mathrm{K}}(m,n;m',n') = \langle \mathcal{K}\psi_{m'n'}, \psi_{mn} \rangle , \tag{2.107}$$

which can also be written as the transformation from a temporal (or spatial) representation to the wavelet representation, i.e.

$$\check{\mathrm{K}}(m,n;m',n') = \int_t \int_{t'} \psi^*_{mn}(t)\, \mathrm{K}(t,t')\, \psi_{m'n'}(t')\, \mathrm{d}t\, \mathrm{d}t'. \tag{2.108}$$

For appropriately chosen frame parameters σ_0 and b_0 the wavelet domain representation $\check{\mathrm{K}}(m,n;m',n')$ completely represents the operator $\mathcal{K}$. In Figure 2.10c the representation of a monochromatic Green's operator in the wavelet transform domain is shown. The interpretation of the coefficients of the kernel $\check{\mathrm{K}}(m,n;m',n')$ will be left to the next chapter. There, I will also discuss other operator representations.

2.5.4 Multiresolution approximation

The idea of a multiresolution approximation[36] is to write a function $f(t)$ as a limit of successive approximations in a series of nested subspaces V_m. Each of the approximations can be seen as a smoothed version of the original function. The approximation in the subspace V_m can be obtained by a projection on this subspace. The approximation is said to be at resolution m or at scale 2^m. The difference between two successive approximations is the detail at resolution m. Note that in section 2.4.1 a similar reasoning yielded the continuous wavelet transform as the difference between two approximations at an infinitesimal small different scale. The multiresolution approximation can be seen as a discrete counterpart of the diffusion equation, although the theory is quite different.

Definition 2.3: Multiresolution approximation (Adapted from Mallat, 1989b)
A multiresolution approximation consists of a sequence of successive approximation spaces V_j satisfying

[35] It can be shown that the Haar wavelet family forms an orthonormal basis for $L^{\mathrm{p}}(\mathbb{R})$, $1 < \mathrm{p} < \infty$ (Daubechies, 1992, chapter 9).

[36] The term *multiresolution analysis* is nowadays more often used. The term *multiresolution approximation*, however, better refers to its actual significance.

1. $\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \subset \cdots$
2. $\bigcup_{m \in \mathbb{Z}} V_m$ *is dense in* $L^2(\mathbb{R})$
3. $\bigcap_{m \in \mathbb{Z}} V_m = \{0\}$
4. $f(t) \in V_0 \iff f(2^{-m}t) \in V_m$
5. $f(t) \in V_0 \iff f(t-n) \in V_0 \quad \forall\, n \in \mathbb{Z}$
6. *there is a function* g *in* V_0 *such that the set* $\{g(x-k),\, k \in \mathbb{Z}\}$ *forms a Riesz basis of* V_0.

Let me briefly interpret these properties. The first property denotes that V_m is a causal sequence. A projection of f in V_m contains all the necessary information to compute an approximation in a coarser subspace V_{m+1}. This can be seen as a semi-group property. When computing an approximation of f in the subspace V_m some information of f is lost. However the difference between f and its approximation can be made arbitrarily small by choosing smaller values for m. This is what property two says. On the other hand, for larger m the projection of f on V_m contains less and less detail and it will finally be zero, which is stated in the third property. The fourth property denotes that V_m is a scaled version of V_0. The fifth property expresses that V_0 is invariant under integer translations, and due to property four, V_m is invariant under translations by $2^{-m}n$. Properties five and six together give rise to a so-called sampling space (Holschneider, 1995). The sinc-function, for example, gives rise to a sampling space, according to Shannon's sampling theorem (Shannon, 1949).

It can be shown (Mallat, 1989b) that the presence of a Riesz basis[37] guarantees that a multiresolution approximation gives rise to a unique scaling function $\phi(t)$ such that $\phi_{mn}(t)$, defined by

$$\phi_{mn}(t) \triangleq 2^{-m/2}\phi(2^{-m}t - n) \tag{2.109}$$

with $(m,n) \in \mathbb{Z}^2$, forms an orthonormal basis of V_m. The projection of f in the subspace V_m, denoted by $\mathcal{A}_m\{f,\phi\}(t)$, now reads (cf. equation 2.84)

$$\mathcal{A}_m\{f,\phi\}(t) = \sum_n \langle f, \phi_{mn}\rangle\, \phi_{mn}(t). \tag{2.110}$$

In the case that no doubt exists with respect to the scaling function ϕ, $\mathcal{A}_m f(t)$ will be used as a shorthand notation for $\mathcal{A}_m\{f,\phi\}(t)$. The inner products of f with the scaling functions ϕ_{mn}, with $n \in \mathbb{Z}$, completely characterize the approximation in the subspace V_m. In addition, a measurement is normally given as a set of discrete data. Hence, it is appealing to refer to a set of measurements at resolution m as

$$\mathcal{A}_m^d f \triangleq \{\, \langle f, \phi_{mn}\rangle\,,\ n \in \mathbb{Z}\}. \tag{2.111}$$

[37] For a general introduction to different types of bases the reader is referred to Young (1980).

$\mathcal{A}_m^d f$ is called the discrete approximation of f at the resolution m.

Let me discuss two eye-catching aspects of a multiresolution approximation. First, due to property 1 of Definition 2.3 and due to the fact that $\{\phi_{m-1,n}, n \in \mathbb{Z}\}$ is an orthonormal basis of V_{m-1}, any function in V_m can be expressed as a linear combination of the basis functions of V_{m-1}, in particular

$$\phi_{mn} = \sum_k h_{2n-k} \phi_{m-1,k}, \tag{2.112}$$

where

$$h_n \triangleq \langle \phi, \phi_{-1,n} \rangle . \tag{2.113}$$

Consequently, the discrete approximation in the subspace V_m can be derived from the discrete approximation in the subspace V_{m-1} according to

$$\langle f, \phi_{mn} \rangle = \sum_k h^*_{2n-k} \langle f, \phi_{m-1,k} \rangle , \tag{2.114}$$

or more symbolically

$$\mathcal{A}_{m-1}^d f \xrightarrow{\mathsf{h}^*} \mathcal{A}_m^d f, \tag{2.115}$$

where h refers to the set of discrete filter elements h_n. In words, $\mathcal{A}_m^d f$ can be computed from $\mathcal{A}_{m-1}^d f$ by a convolution with the filter coefficients of h and, subsequently, retaining every other sample.

Secondly, the difference between two successive approximations of f in V_{m-1} and V_m is the projection of f on the orthogonal complement of V_m in V_{m-1}. The orthogonal complement is denoted by O_m. Mathematically, this is expressed by

$$O_m \perp V_m \tag{2.116}$$

$$O_m \oplus V_m = V_{m-1}. \tag{2.117}$$

To compute the projection on the orthogonal complement O_m, a basis is required. Mallat (1989a,b) shows that within a multiresolution approximation a function $\psi(t)$ exists, such that the set of functions $\{\psi_{mn},\ n \in \mathbb{Z}\}$ forms an orthonormal basis of O_m. Not unexpectedly, the function $\psi(t)$ turns out to be a wavelet, because the difference between two (low-pass) approximations is clearly of a band-pass nature. The functions $\psi_{mn}(t)$ are defined according to equation (2.105) with $\sigma_0 = 2$ and $b_0 = 1$. The projection of a function f on the subspace O_m is given by

$$\mathcal{D}_m\{f, \psi\}(t) = \sum_n \langle f, \psi_{mn} \rangle \psi_{mn}(t). \tag{2.118}$$

In the case that there is no confusion about the nature of the wavelet ψ, $\mathcal{D}_m\{f, \psi\}(t)$ is also denoted by $\mathcal{D}_m f(t)$. The discrete detail in the subspace O_m is denoted by $\mathcal{D}_m^d f$, and it is defined by

$$\mathcal{D}_m^d f \triangleq \{ \langle f, \psi_{mn} \rangle ,\ n \in \mathbb{Z} \}. \tag{2.119}$$

In analogy with the reasoning in the previous paragraph, it is easy to derive an expression for the discrete detail in terms of the discrete approximation. It follows from (2.117) that $O_m \subset V_{m-1}$. Due to the fact that $\{\phi_{m-1,n}, n \in \mathbb{Z}\}$ is an orthonormal basis of O_{m-1}, any function in V_m can be expressed as a linear combination of the basis of V_{m-1}, in particular

$$\psi_{mn} = \sum_k g_{2n-k} \phi_{m-1,k}, \tag{2.120}$$

where

$$g_n = \langle \psi, \phi_{-1,n} \rangle . \tag{2.121}$$

Consequently, the discrete detail in the subspace O_m can be derived form the discrete approximation in the subspace V_{m-1} according to

$$\langle f, \psi_{mn} \rangle = \sum_k g^*_{2n-k} \langle f, \phi_{m-1,k} \rangle , \tag{2.122}$$

or more symbolically

$$\mathcal{A}^d_{m-1} f \xrightarrow{\mathrm{g}^*} \mathcal{D}^d_m f, \tag{2.123}$$

where g refers to the set of discrete filter elements g_n. In words, $\mathcal{D}^d_m f$ can be computed from $\mathcal{A}^d_{m-1} f$ by a convolution with the filter coefficients of g and, subsequently, retaining every other sample.

Based upon the aforementioned theory, a function f in $L^2(\mathbb{R})$ can be represented, completely or up to a certain fine resolution, in various ways:

1. Due to property two of Definition 2.3 of a multiresolution approximation, and using equation (2.117), it can be seen that $\{\psi_{mn},\ m, n \in \mathbb{Z}\}$ forms an orthonormal basis of $L^2(\mathbb{R})$. Hence, any function in $L^2(\mathbb{R})$ can be represented according to

$$f(t) = \sum_m \sum_n \langle f, \psi_{mn} \rangle \psi_{mn}(t), \tag{2.124}$$

or, using equation (2.118),

$$f(t) = \sum_m \mathcal{D}_m f(t). \tag{2.125}$$

Consequently, using equation (2.119) it can be concluded that the set

$$\{\mathcal{D}^d_m f,\ m \in \mathbb{Z}\} \tag{2.126}$$

completely characterizes the function f.

2. It is not necessary to do a decomposition up to an infinitely coarse resolution. It is allowed to stop a decomposition at a certain resolution and to use a coarse approximation in the subspace V_M to represent the remaining part, i.e.

$$f(t) = \sum_{m \leq M} \sum_{n} \langle f, \psi_{mn} \rangle \, \psi_{mn}(t) + \sum_{n} \langle f, \phi_{Mn} \rangle \, \phi_{Mn}(t), \tag{2.127}$$

or, using equations (2.110) and (2.118),

$$f(t) = \sum_{m \leq M} \mathcal{D}_m f(t) + \mathcal{A}_M f(t). \tag{2.128}$$

Consequently, the discrete representation of a function f in a combined wavelet and approximation space reads

$$\{\ldots, \mathcal{D}^d_{M-2} f, \mathcal{D}^d_{M-1} f, \mathcal{D}^d_M f, \mathcal{A}^d_M f\}. \tag{2.129}$$

3. If a representation of f at a certain finite resolution, which is for convenience fixed at $m = 0$, is available, i.e. a projection of f in the subspace V_0, then it is easy to see that

$$\mathcal{A}_0 f(t) = \sum_{1 \leq m \leq M} \sum_{n} \langle f, \psi_{mn} \rangle \, \psi_{mn}(t) + \sum_{n} \langle f, \phi_{Mn} \rangle \, \phi_{Mn}(t), \quad M > 0, \tag{2.130}$$

or in the compact notation of equation (2.128)

$$f(t) = \sum_{1 \leq m \leq M} \mathcal{D}_m f(t) + \mathcal{A}_M f(t). \tag{2.131}$$

It is now clear that the discrete approximation $\mathcal{A}^d_0 f$ is completely equivalent to $\mathcal{D}^d f$, which is given by

$$\mathcal{D}^d f \triangleq \{\mathcal{D}^d_1 f, \mathcal{D}^d_2 f, \ldots, \mathcal{D}^d_M f, \mathcal{A}^d_M f\}. \tag{2.132}$$

Using equations (2.114) and (2.122) one can see that the discrete representation of (2.132) can be easily obtained from the discrete approximation $\mathcal{A}^d_0 f$, according to the scheme in Figure 2.12. This decomposition can be seen as a unitary matrix operation. It can be carried out very efficient in $\mathcal{O}(N)$ operations, where N is the number of elements in the discrete approximation $\mathcal{A}^d_0 f$. The set $\mathcal{D}^d f$ is often referred to as the discrete wavelet transform of $\mathcal{A}^d_0 f$. The action of going from the discrete approximation $\mathcal{A}^d_0 f$ to $\mathcal{D}^d f$ is referred to as the discrete wavelet transformation or discrete wavelet transform[38]. Since the discrete wavelet transform is unitary, it is easily invertible.

[38] The terminology with respect to transform and transformation is elucidated in the introductory chapter on page 5.

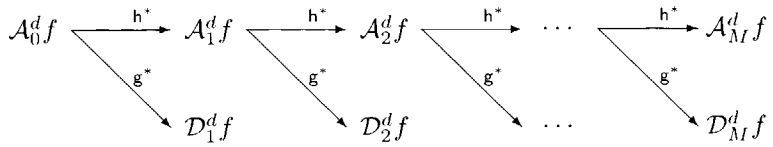


Fig. 2.12 *Decomposition scheme from the discrete approximation $\mathcal{A}_0^d f$ in the subspace V_0 to discrete details and discrete approximations in coarser approximation and detail subspaces. This decomposition is referred to as the discrete wavelet transformation.*

The filter coefficients of h and g, also called quadrature mirror filters, turn out (Daubechies, 1988, 1992) to be related according to

$$g_n = (-1)^n h^*_{-n+1}. \tag{2.133}$$

The multiresolution approximation is completely determined by the scaling function ϕ, but also by a specification of either one of the filters h and g. Up to now nothing has been said about the type of multiresolution approximation. In contrast to the affine coherent states, which allow a continuous variation of the scale and translation parameter, or the wavelet frames, which allow a general sampling of the scale-time plane, the sampling of the scale-time plane in a multiresolution approximation is fixed. There is, however, some freedom of choice with respect to the type of wavelet. The ground state of an affine coherent state can be characterized by looking at its decay properties, the number of vanishing moments and its regularity (see section 2.4.2). These properties do apply to the wavelet of a multiresolution approximation as well. Besides, requirements with respect to symmetry, boundary treatment and efficiency further determine the choice for a specific multiresolution approximation. Particular user requirements determine which properties prevail, since a multiresolution approximation does not allow a complete freedom of choice. The wavelet of a multiresolution approximation can not be both C^∞ (i.e. infinitely regular) and exponentially decaying, whereas, for example, the derivative of a Gauss function combines these properties. The derivate of a Gauss, however, can never be part of a multiresolution approximation. Moreover, within a multiresolution approximation the regularity is proportional to the number of vanishing moments (Daubechies, 1992, Corrolary 5.5.2 and 5.5.3). The treatment of all possible filter types is far beyond the scope of this thesis. In the applications in chapter 4, I will make use of the compactly supported wavelets of Daubechies (1988), a choice which will be further clarified in that chapter. Compactly supported wavelets are very efficient, but they have a limited regularity. In addition, they can not be symmetric, except for the Haar wavelet. Figure 2.13 shows some wavelets and scaling functions in the class of orthogonal compactly supported wavelets.

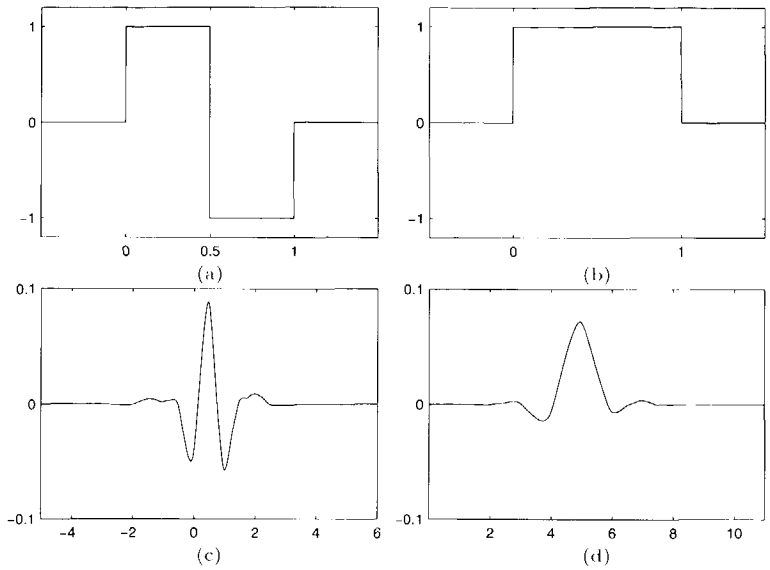


Fig. 2.13 *Examples of wavelet filters and scaling functions. (a) The Haar wavelet; (b) the Haar scaling function; (c) a compactly supported Daubechies wavelet with 6 vanishing moments and as symmetric as possible; (d) corresponding scaling function.*

2.6 Applications of the wavelet transform

The rather theoretical framework that has been set down in the previous sections requires a seismic veneer in order to be fully appreciated. The veneer will take the form of a discussion of the five main groups of applications, in relation to the seismic method, as I have come to see them in the course of the past years. The main groups of applications are: compression, time-frequency analysis, operator representation, singularity analysis, and image analysis. Three of the five applications (operator representation, singularity analysis and image analysis) are the subject of separate chapters (chapter 4, 5 and 6 respectively). The reason for not discussing the compression application in a separate chapter is mainly the fact that considerable success has already been achieved without my contribution. The remaining problems related to the compression application are not specifically associated to the wavelet transform anymore. The reason for not discussing the time-frequency analysis in a separate chapter is the fact that I have not been able to make a significant contribution in this area[39]. In this section I will start with a description of the guiding principle for the five application groups. I will subsequently deal with

[39]In my opinion the wavelet transform as a tool to carry out time-frequency analysis (especially for the separation of different types of events) has not proven its usefulness for the seismic method yet. Especially the coherency of seismic reflection data over multiple traces is not easily dealt with.

the different applications in five subsections. For the applications that will be dealt with in separate chapters, I will only discuss the basic idea and refer to the sections of the present chapter that should be used as background material. The other two subjects will be dealt with in more detail in the present section.

The five applications, compression, time-frequency analysis, operator representation, singularity analysis and image analysis are quite different, but note that they are neither completely independent nor that they are put in a random order. The common part of the applications is the fact that they rest on the principle that a wavelet is "blind" to something that is slowly varying, and on the principle that a wavelet transform allows for a natural subdivision in details and approximations (at multiple scales). These principles have been given a mathematical basis in the preceding sections. All the applications rest at least on a manifestation of one of the two principles. The actual form and the importance will be discussed in each of the subsections. The significance of the chosen order comes across if it is put next to a standard seismic processing sequence as has been done in the introductory chapter (Figure 1.3). A seismic processing sequence starts with the seismic reflection data (compression), and via a preprocessing step (time-frequency analysis), an imaging step (operator representation), and an optional characterization step (singularity analysis), one will finally get an image of the subsurface (image analysis).

2.6.1 Compression

Seismic data volumes are rapidly growing due to ever increasing 3-D data acquisition surveys. In spite of the fact that storage and transmission capacities are growing as well, there is an increasing demand for efficient and effective seismic data compression techniques to enhance the seismic processing sequence. The enhancement can be obtained for example by an easier quality control during acquisition, or by an easy accessibility of the data during processing. Large compression ratios (up to a factor of 100) can be obtained with the compression techniques at the cost of some loss of information. In this subsection I will briefly describe the compression algorithm with a focus on the role of the wavelet transform[40]. At the end, the main issues related to seismic data compression will be pointed out. The main issues are not related to the wavelet transform anymore. Therefore, I did not devote a separate chapter to this subject.

A compression algorithm generally consists of three steps: a transformation T, a quantization Q and an entropy coding E (Vetterli and Kovačević, 1995). The first two steps aim at reducing the entropy of the (possibly more dimensional) signal, the entropy coding can be seen as an efficient reordering of the data on a bit-level. Figure 2.14 shows the compression and decompression scheme. The next

[40]The contents of the present subsection is extracted from a Technical Report by Dessing and Hoekstra (1997), which deals with the effect of seismic data compression on (angle-dependent) imaging.

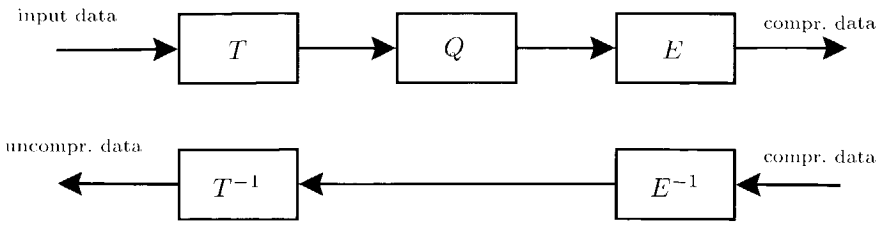


Fig. 2.14 *Scheme of a data compression algorithm. The first step is a transformation T, the second step a quantization Q and the last step an entropy coding E. Generally, this is an open-loop system, because the middle step is irreversible and introduces errors. The inverse scheme is shown in the bottom graph. For the lossy compression scheme discussed here the uncompressed data is in general not equal to the input data.*

three paragraphs subsequently deal with the transformation step, the quantization step and the entropy coding step.

The reduction of the entropy through the transformation step can be viewed upon as trying to realize an optimal decorrelation or an optimal concentration of energy. A well-known result from information theory is the fact that the Karhunen-Loève transform can realize an optimal packing of the energy, i.e. the minimum entropy (Coifman and Wickerhauser, 1992; Vetterli and Kovačević, 1995). However, a major drawback concerning this transform is its inefficiency. Quite a few alternatives exist. For seismic reflection data the discrete wavelet transform applied to the time dependent coordinate as well as to the spatial coordinates turns out to be very effective (Bosman and Reiter, 1993; Ergas et al., 1996; Reiter, 1996). It is able to decorrelate the stream of input data, and consequently to reduce the entropy of the input stream. The filter pairs that are generally used correspond to biorthogonal wavelets (Daubechies, 1992), which have the advantage to be both of compact support and symmetric. These filter pairs are able to effectively capture a wide range of reflection events and to make partially use of the trace-to-trace coherence of seismic reflection data.

Although a transformation can help in packing the energy in a signal, a real data reduction has not been obtained after the transformation step. The energy is concentrated in less samples, but the same number of samples is still in play with a wide range of output levels. Quantization reduces the number of output levels. Unlike the transformation (and the entropy coding to be described in the next paragraph), quantization results in a loss of information. It is an irreversible non-linear step, that contributes largely and flexibly to the decrease in entropy. Quantization consists of a mapping of a set of real valued samples into a discrete set, i.e. it maps the real line into a countable discrete alphabet (Vetterli and Kovačević,

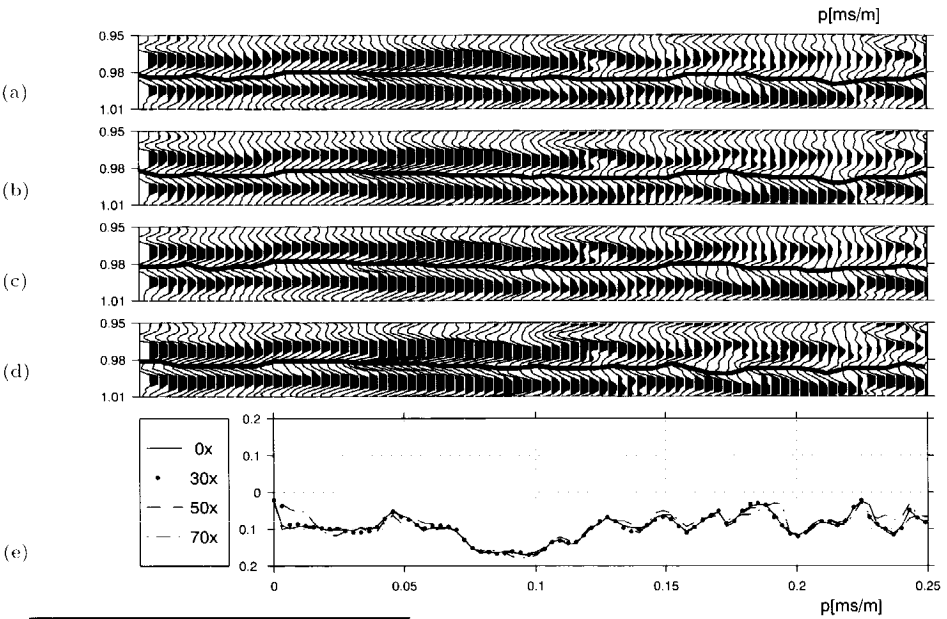


Fig. 2.15 *Effect of compression on angle-dependent imaging. The shot records of a real data set have been compressed and decompressed with a lossy compression scheme as described in the text. After decompression an angle-dependent imaging procedure has been carried. Around the target level of this specific data set the angle-dependent behavior is shown. (a) Angle-dependent behavior without compression. (b)-(d) Angle-dependent behavior after compression/decompression with a compression ratio 30 (b), compression ratio 50 (c) and compression ratio 70 (d). (e) The amplitudes along the tracked lines. The differences are small. The amplitudes have been tracked with an automatic tracking algorithm based on the wavelet transform modulus maxima representation. The compression/decompression has been carried out with commercial software (Courtesy Aware Inc., USA). This example has been extracted from Dessing and Hoekstra (1997).*

1995). A well-known but primitive form of quantization is thresholding. Generally, more advanced quantizations are utilized (Gersho and Gray, 1992).

Entropy coding is the final (reversible) step in a compression procedure. Entropy is the measure for the amount of concentration of energy. In the field of information theory, it has been introduced by Shannon (1949). It consists of a mapping from the output set of the quantization step to a new set such that the average number of bits per sample is minimized. Vetterli and Kovačević (1995) write: "A historical example is the Morse code which assigns short codes to letters that appear frequently in English language while preserving long codes to less frequent ones." As such, Morse code is one of the first examples of entropy coding.

The compression procedure itself is regarded to be quite mature. As an example Figure 2.15 is included. It shows the effects of the compression of shot records on the AVO-behavior after imaging. Questions are especially related to the a priori estimation of the propagation of errors to the structural image, to the AVO attributes and to other amplitude- and phase-dependent features after imaging. Some relatively small questions are left open with respect to the wavelet transform. How to further optimize the discrete filter pair? How to choose the sub-band division? How can the efficiency of the forward and inverse transformation be improved? It is conjectured that the recently developed Lifting Scheme might provide answers to some of these questions. The Lifting Scheme is a very efficient and flexible variant of the discrete wavelet transform (Sweldens, 1996).

2.6.2 Time-frequency analysis

The application of the wavelet transform for time-frequency analysis of seismic reflection data can be subdivided in two areas: (1) the analysis of the time, shape, phase, etc. of reflection events; (2) the separation of different types of events, such as primaries, multiples, ground-roll, noise, etc.

In the first application area one is aiming at extracting information concerning the propagation and reflection process, from the seismic data. Here, the wavelet transform is an analysis tool with little concern with respect to reconstruction and efficiency. The results of Morlet et al. (1982) can be placed in this application area. Since I have explicitly focussed on the relation between the subsurface complexity and the structure of the reflection events I have ranged this time-frequency application under the singularity analysis in chapter 5. For this application area I see sufficient opportunities.

The second application area has attracted some attention over the past years (Rappin and de Bazelaire, 1992; Grubb and Walden, 1994; Foster et al., 1994), which were however not particularly successful. The requirements here are related to efficiency and reconstruction. One is consequently forced to appeal to the discrete wavelet transform. The discrete wavelet transform however has three disadvantages: it is not translation invariant, it provides a relatively bad temporal resolution, and it does not sufficiently take the trace-to-trace coherency of reflection events into account. These limitations do not stride well with the analysis properties required for a good separation of events.

2.6.3 Operator representation

A model of a physical process given in terms of operators is abstract. However, such an operator model allows for a clear separation of different aspects of the physical model. Moreover, it allows for a conceptual treatment of the problem under consideration, using a generalized formulation of the operators. As soon as it comes

to actually solving a specific problem, a specific representation of the operators in a particular domain has to be utilized. In chapter 4, such an approach will be followed while dealing with the migration problem.

Migration is the most computer intensive part in a seismic processing sequence. The seismic community is continuously looking for techniques that can improve the efficiency. In chapter 4 the opportunities provided especially by the wavelet transform will be worked out. The point of departure is the monochromatic representation of seismic reflection data, which will be introduced in full detail in chapter 3. In its basic form the data representation takes the following operator form (see also chapter 1):

$$P^- = \mathcal{W}^- \mathcal{R}^+ \mathcal{W}^+ S.$$

Following the formulation of Berkhout (1982), in seismic migration the aim is to remove the effects of the propagators $\mathcal{W}^{\pm}$ and to estimate the reflection operator $\mathcal{R}^+$. In chapter 4 the possibilities to efficiently render the propagators in the wavelet transform domain will be presented. Besides, a method will be shown that allows for a coarse-to-fine reconstruction of the subsurface image via a division of the reflection operator in a coarse and a detail part.

The two guiding principles for this application are the fact that wavelets are "blind" to smooth parts of a kernel and the fact that the multiresolution approximation gives naturally rise to a division into a coarse part and a detail part. The background material for chapter 4 is provided by sections 2.5.3 and 2.5.4.

2.6.4 Singularity analysis

The subsurface structure is highly complex. Reflection is the result of a local interaction process of the complex subsurface with a wave, propagating into the earth. The interaction process takes place over the effective size of the probing seismic wave. A good understanding of this interaction process can help in characterizing relevant geologic boundaries. In chapter 5 analytic studies are carried out to analyze how the complexity in the subsurface is transferred to the reflected seismic wave field both for normal and oblique incidence. The subsurface complexity is considered to consist of homogeneous distributions on top of a smooth background. The wavelet transform is used here as a tool to characterize the subsurface complexities both directly, and indirectly via the reflected seismic wave field.

The guiding principle for this application is the fact that the wavelet transform senses rapid variations, the singularities. With the wavelet transform one can relatively easy characterize the local singularities. The background material for chapter 5 is mainly provided by section 2.4.

2.6.5 Image analysis

Nowadays 3-D acquisition surveys are used routinely to collect seismic reflection data. Such 3-D acquisition surveys naturally lead to 3-D migrated cubes. Whereas the task of a seismic interpreter was surveyable for 2-D sections, the task gets much more complicated as soon as it comes to 3-D migrated cubes, despite the fact that powerful visualization tools have been developed in recent years. The interpreter is faced with the task to extract stratigraphic features and faults from the 3-D data sets, i.e. he or she has to integrate information from multiple slices and sections. Automatic extraction procedures which require an image analysis preprocessing step will certainly be of advantage to the seismic interpreter. In chapter 6 the opportunities provided by the 3-D wavelet transform for the delineation of interesting geological features will be pointed out.

The guiding principles for this application are based upon the fact that the wavelet transform senses rapid variations at the scale of the wavelet, and the fact that the full 3-D data coherence or incoherence is utilized. The background material for chapter 6 is provided by section 2.4.1 and by sections 2.5.3 and 2.5.4.

2.7 Summary

In this chapter I have presented an overview of linear representations and transformations to be utilized in subsequent chapters. A representation can be seen as the realization of a function or operator in a particular domain. A transformation is an operation that links two different representations. The representations have been subdivided in two main classes: one is associated with self-adjoint operators, the other with the irreducible representation of groups, yielding the coherent states. The former class comprises the temporal domain representation, the Fourier domain representation and the Mellin domain representation. The latter class generates the Gabor representation or windowed Fourier representation and the wavelet representation. Table 2.2 summarizes the symbols associated with these domains. Whereas self-adjoint operators generate orthonormal bases, coherent states generate a highly redundant set of functions. For analysis purposes the redundancy is advantageous, however for efficiency purposes a subset is looked for. The theory of frames provides methods to extract subsets that still represent the complete space under consideration. For the Gabor representation it is not possible to find complete and orthogonal subsets. However, within the theory of multiresolution approximations it can be seen that orthogonal and complete subsets of wavelets can be easily found.

A joint time-frequency plane has been introduced to better understand the functionality of the different representations. The temporal representation provides information on the temporal development of a variable, but it does not give information on the frequency behavior. The Fourier representation provides the frequency behavior, but it does not provide temporal information. The Gabor representa-

symbol	domain
g	time
$\tilde{g}$	frequency
$\hat{g}$	Gabor
$\check{g}$	wavelet
$\breve{g}$	Mellin

Table 2.2 *Shorthand transformation symbol list for the domains introduced in this chapter for a general function g.*

tion provides both temporal and frequency information within a window function of constant shape and width. The wavelet representation is a decomposition with constant-Q filters, i.e. for higher frequencies the temporal resolution improves. Because the analyzing functions have a constant shape and a variable width, the wavelet transform acts as a mathematical microscope. A further understanding is obtained by looking at the behavior of the different representations under the action of parameterized unitary operators, such as shifting, modulation and dilation. This procedure yields invariance and covariance properties.

Whereas the wavelet transform can be seen as just another division of the time-frequency plane with a number of properties, its fundamental significance deserves a better motivation. The wavelet transform naturally arises as a result of an observation at different scales, or as a result of an observation at different times. The changes in the scale direction or in the time direction are exactly the elements that a wavelet representation is registering. This interpretation of the wavelet representation, makes the link with distribution theory manifest. Moreover, understanding that the wavelet representation expresses the changes of a distribution under infinitesimal scale changes makes the introduction of a multiresolution approximation just a logical step. A multiresolution approximation is a formal system that relates approximations of functions at different (dyadic) scales. Due to the fast wavelet transform algorithms induced by a multiresolution approximation, this specific interpretation of a multiresolution approximation is in danger of being overlooked.

The present chapter is finalized with an overview of the different application areas for the wavelet transform in relation to the seismic method. The following application groups are discriminated: compression, time-frequency analysis, operator representation, singularity analysis, and image analysis. All these applications are based upon two guiding principles: (1) a wavelet is "blind" to something that is slowly varying; (2) the wavelet transform allows for a natural subdivision in details and approximations at multiple scales.

Chapter 3

One-way representation of seismic data

The content of this chapter is a result of work done in collaboration with Grimbergen and Wapenaar. During his stay at our laboratory as an M.Sc. student, Grimbergen was asked to look for accurate solutions of the monochromatic wave equation in laterally varying medium configurations. Together with Wapenaar and the author he came up with an elegant solution based upon a modal decomposition, which is well-known in quantum physics, but whose usefulness for seismic applications seemed to be overlooked. Grimbergen wrote a well received Master's thesis (Grimbergen, 1995) and later on the work has been accepted for publication (Grimbergen et al., 1998). In this chapter the line of the accepted article will be roughly followed. The emphasis will be put on a refined analysis of the spectrum of the Helmholtz operator. Moreover, a more extensive introduction of the basic equations leading to the monochromatic one-way wave equation and to the one-way data representation will be given for further use in subsequent chapters. Applications in wave field extrapolation and migration will be briefly discussed.

3.1 Introduction

As described in the introductory chapter, the most important source of information that an exploration geophysicist has at his disposal is seismic reflection data. Seismic data carry the information of the structure and character of the subsurface in a disguised form. The disguise is mainly due to propagation effects to and from the reflecting boundaries and due to complex reflections at all reflecting boundaries. These two effects cause the seismic data to be retarded and dispersed. As one is mainly interested in the location and character of the seismic reflectors, one would preferably like to have sources and receivers directly on top of all the reflecting

boundaries. Thereby, one has actually solved the problem called migration. The task of migration is to put sources and receivers virtually on top of the reflecting boundaries. Which tools are required to relocate the sources and receivers from the surface to the reflecting boundaries in the subsurface?

Firstly, an adequate model or theory describing the physical process under consideration is indispensable. Generally, a physical theory can be casted in a differential form according to (cf. equation 2.1)

$$\mathcal{A}u = f, \tag{3.1}$$

where f is the source function that brings the process into action, u is the actual field caused by the action f, and where the operator $\mathcal{A}$ relates infinitesimal changes in u to the source field f. In the case of acoustical remote sensing a theory is looked for[1] describing the propagation and reflection of elastic waves in solids and fluids; hence, I am looking for a wave operator $\mathcal{A}$ coupling the source wave field f with the wave field u in the medium. The wave theory that I am referring to has been developed by some great scientists living in the 17th, 18th and 19th century. Although giving a full historical account of the scientific development of the theory of elastic and acoustic waves falls outside the scope of the thesis, I would like to point out a few highlights. Noteworthy to be mentioned is D'Alembert who wrote down the first wave operator $\mathcal{A}$ for a vibrating string in 1747, but also Navier, Poisson and Cauchy who developed in competition the basic theory of elastic waves in solids and, finally, Lord Rayleigh who gave a rather thorough treatment of the theory of sound in his two volume masterpiece (Rayleigh, 1877, 1878). Realizing that a shorter history of a scientific theory has almost never been written, I would like to refer the reader to Rayleigh (1877) and Love (1927) for a more elaborate historical perspective. In this thesis the theory of the propagation and reflection of compressional waves in fluids will be employed. The functional dependence of the wave operator or D'Alembertian in equation (3.1) can be expressed by $\mathcal{A}(\rho, \kappa, \partial_t, \partial_k)$, where ρ is the volume density of mass [kg m^{-3}], and κ is the compressibility [Pa^{-1}]. The other two symbols are introduced in section 1.5. Further details concerning the actual form of the wave operator follow in section 3.2.

Secondly, a method is required to reduce the set of observations, i.e. the seismic data, such that one is able to pronounce upon the density and compressibility as a function of the spatial coordinates. Is this a simple task? Not at all. First of all, one has to be confident that the forward problem, symbolically represented by

$$\rho, \kappa, f \longrightarrow u,$$

[1] Note that as soon as a theory has been successfully formulated, one can easily say what the scientist has been looking for. If a theory is not present, one is not aware of the actual missing link.

and consisting of equation (3.1) together with appropriate Dirichlet or Neumann boundary conditions and Cauchy conditions (see e.g. Dautray and Lions, 1992), is well-posed[2] and valid for the actual configuration under consideration. Assuming that the forward problem is well-posed the inverse problem can be formulated:

$$u, f \longrightarrow \rho, \kappa. \tag{3.2}$$

Generally, this is however an ill-posed problem. To mention three reasons for the ill-posedness: (1) seismic data are only acquired at the surface; (2) the seismic data is band limited; (3) waves are not only propagating but are also partially diffused due to the small scale heterogeneities, which is a semi-group action.

In the present thesis, the inverse problem of equation (3.2) is approached via a one-way or directional decomposition of the wave fields. The keynote of the directional decomposition is the explicit discrimination between propagation in the direction of preference and scattering in the direction of preference. This procedure leads to a 3-D one-way forward model of seismic data related to reflections at all depth levels $x_3 > 0$ in the subsurface according to (see also Figure 1.2)

$$P^- = \int_{x_3>0} \mathcal{W}^- \mathcal{R}^+ \mathcal{W}^+ S^+ \, \mathrm{d}x_3. \tag{3.3}$$

Here, S^+ is a representation of the source distribution, $\mathcal{W}^+$ is an operator which brings the wave field from the surface to a specific depth level x_3, $\mathcal{R}^+$ is an operator denoting reflection at depth level x_3, $\mathcal{W}^-$ denotes the upward propagation operator from depth level x_3 back to the surface, and P^- is a representation of the seismic data measured at the surface related to the reflection at all depth levels. The operators $\mathcal{W}^+$ carry out a (generalized) spatial convolution. Equation (3.3) is valid for general 3-D configurations. In section 3.2 the data model of equation (3.3) will be dealt with in more detail. For this introduction the important points are:

1. The data model for P^- is a first order equation in the reflection operator $\mathcal{R}^+$. Surface related multiples are not included in the model on the right-hand side of equation (3.3); it is assumed that they have been eliminated in the data at the left-hand side of equation (3.3). Hence, $\mathcal{R}^-$ is not considered. Therefore, the shorthand notation $\mathcal{R}$ is used for the subsurface reflection operator $\mathcal{R}^+$ in the sequel.

2. Inversion of equation (3.3) yields an estimation of the (angle-dependent) reflectivity operator, which is a function of the rapid and significant changes in the

[2] A problem is well-posed if (1) there exists a solution, (2) the solution is unique and (3) the solution depends continuously on the data. A problem is ill-posed if it is not well-posed (Renardy and Rogers, 1993).

subsurface parameters[3]. In the seismic literature, the estimation of the reflectivity is called migration. As mentioned before migration does not make any assumptions about the reflectivity $\mathcal{R}$. Therefore, migration is a less ambitious process than the computation of the full profiles of the constitutive parameters, which is generally the aim of the inverse scattering approach (e.g. Bleistein and Cohen, 1979). The distinction between migration and the inverse scattering approach is discussed in considerable detail in Berkhout (1982), Tarantola (1984), Stolt and Weglein (1985) and Wapenaar (1996a).

3. The problem described in the first paragraph of this section of finding the reflecting boundaries has been "reduced" to the computation of the operators $\mathcal{W}^{\pm}$ and their inverses in a propagation or background medium characterized by ρ_0 and κ_0[4]. Hence, the inverse problem of equation (3.2) can now be reformulated as

$$P^-, S^+ \xrightarrow{(\mathcal{W}^{\pm})^{-1}} \mathcal{R}, \tag{3.4}$$

which represents a (generalized) spatial and temporal deconvolution process. Given P^- and S^+, the reflectivity $\mathcal{R}$ can be extracted by appropriately applying the operators $(\mathcal{W}^{\pm})^{-1}$.

4. Finally, the estimated subsurface reflection operator $\mathcal{R}$ can be used as input for a lithologic inversion procedure in order to find the rock and pore parameters or in order to characterize the subsurface boundaries.

The third point touches upon the heart of the reason to write this chapter: the computation of the operators $\mathcal{W}^{\pm}$. The operators $\mathcal{W}^{\pm}$ describe propagation from one depth level $x_{3,m}$ to another depth level $x_{3,n}$. Henceforth, the operators $\mathcal{W}^{\pm}$ will be referred to as propagators[5].

The computation of the propagators basically boils down to the computation of

$$\exp(-j\sqrt{\mathcal{H}_2}\,\Delta x_3), \tag{3.5}$$

where $\mathcal{H}_2$ is the Helmholtz operator, which is the temporal Fourier transform of the D'Alembertian; $\Delta x_3 = |x_{3,n} - x_{3,m}|$ is the depth step over which the wave field has

[3]The notions rapid and significant are not absolute quantities. Rapid changes are rapid with respect to the resolution of the probing seismic wave field. Significant changes are able to reflect a considerable amount of energy in a coherent way. Significant changes are not necessarily rapid and the other way around.

[4]In this thesis the question how to estimate the parameters ρ_0 and κ_0 will not be addressed. It is however an important and not completely solved issue; see, for example, Kabir (1997) for a recent treatment of this issue.

[5]Normally, the propagator or evolution operator is referring to a development of a physical system as a function of time, not as a function of one of the spatial coordinates (see e.g. Dautray and Lions, 1992).

to be extrapolated. The medium parameters are assumed not to vary as a function of the x_3-coordinate in the interval $(x_{3,m}, x_{3,n})$. A function of an operator requires special care, because a simple evaluation as with $\exp(x)$ is not allowed, unless the operator is a diagonal operator. Although analytic solutions exist for very specific medium configurations (Fishman, 1992), they are in general not applicable in the highly heterogeneous earth. Consequently, one is driven back on numerical methods. The numerical methods are derived from two competing formally exact expressions for the propagator in (3.5). The first one is based upon a Hamiltonian path integral representation (Feynman and Hibbs, 1965; DeWitt-Morette et al., 1979; Fishman and McCoy, 1984a; Fishman, 1992; de Hoop, 1992) and the second one is based upon a normal mode or eigenvalue decomposition of the Helmholtz operator (Reed and Simon, 1979; Dautray and Lions, 1990).

The first is the most general method in the sense that the depth step Δx_3 in which the parameters are assumed not to vary can be made arbitrarily small in a natural way. Moreover, Fishman and McCoy (1984b) stated that the path integral representation can be seen as a starting point to derive solutions, such as rational approximations (Claerbout, 1971; Berkhout, 1982; de Hoop, 1992; van Stralen, 1997), such as the optimized operator methods (Holberg, 1988; Blacquière et al., 1989; Hale, 1991; Thorbecke, 1997), or such as the uniform expansion methods currently being developed (Fishman et al., 1997). While performing well as long as the lateral medium variations do not take place within the size of the derived operators, the rational approximations and optimized operator methods fail to perform well if significant lateral changes take place within the size of the operators: unreliable or sometimes even unstable results (Etgen, 1994) may be expected.

The second method, based upon a modal decomposition of the Helmholtz operator, provides the possibility of an improved handling of lateral medium variations. From optics, shallow-water acoustics, seismology and also in specific seismic applications, it is known that an expansion of the wave field into wave modes proves to be an appropriate method of dealing with predominantly laterally varying media (e.g. Weinberg and Burridge, 1974; Blok, 1995; Ernst and Herman, 1995). These applications are usually limited to wave guides or other structures with comparatively small variations in the direction of preference. Moreover, in those applications the medium parameters vary in such a way, that the guided wave modes dominate the wave field. Clearly, this is not the case in reflection seismics, where wave guiding situations seldom occur as a result of lateral variations. Here, the radiating part of the wave field is generally more important than the guided wave modes.

As will be shown in this chapter, an extrapolation scheme based on a modal expansion of the wave field into both guided and radiating wave field constituents, significantly increases the lateral resolution of the result. The method is tested in a synthetic migration example. The subsurface model of the migration example contains a high velocity domal structure (salt) and a number of faults. It is shown

that by using the modal expansion for the construction of the one-way wave field operators, significant improvements can be achieved compared to the more local methods. Besides the improved handling of lateral medium variations, the modal decomposition has another advantageous property. It provides insight in the physical structure of the problem at hand through a decomposition in orthogonal wave field constituents.

Close links exist between the method presented in this chapter and the work of Pai (1985) and Kosloff and Kessler (1987). In Pai (1985), a modal decomposition in the wavenumber-frequency domain is carried out for laterally-varying media and applied to the two-way wave equation. For laterally invariant media the resulting extrapolation operator reduces to the phase-shift operator. Kosloff and Kessler (1987) mention the possibility of a modal decomposition applied to the two-way wave equation in the space-frequency domain, but they choose to make use of Chebyshev polynomials as an approximation. Both references take the discretized wave equation as a point of departure, whereas in the present chapter a derivation from the continuous formulation is carried out. This derivation leans upon functional calculus and it provides better insight in the physical nature of the problem at hand.

A general motivation has been given for the present chapter and I went by seven-league strides through what will follow. Let me conclude this introduction with an overview of the flow of the chapter. The chapter will continue in section 3.2 with the basic equations ruling the behavior of compressional waves. In the same section the basic equations will be recasted such that the directional decomposition can be carried out. The directional decomposition yields the so-called one-way wave equations and the primary data representation. Section 3.3 is devoted to the modal decomposition of the Helmholtz operator. Special attention will be paid to the definition of the mathematical space in which the Helmholtz operator is working, the definition of the spectrum, and the derivation of the kernel of the propagator. The attention in section 3.4 is dedicated to examples. In the light of the presented examples, the proposed method is discussed in section 3.4.3, its advantages and disadvantages; section 3.4.3 is as such the overture for chapter 4. Finally, section 3.5 summarizes what has been found.

3.2 The wave equation

The two linearized partial differential equations forming a hyperbolic system and governing the acoustic wave field, are given in a right-handed orthogonal coordinate system by[6]

$$\rho\partial_t v_k + \partial_k p = f_k, \qquad k = 1,2,3, \tag{3.6}$$

$$\kappa\partial_t p + \partial_r v_r = q, \tag{3.7}$$

[6]The notational conventions and frequently used definitions are put together in section 1.5.

where

$p = p(\boldsymbol{x}, t)$ is the acoustic pressure [Pa],
$v_k = v_k(\boldsymbol{x}, t)$ is the particle velocity in the $\boldsymbol{i}_k$-direction [$\mathrm{m\,s^{-1}}$],
$\rho = \rho(\boldsymbol{x}, t)$ is the volume density of mass [$\mathrm{kg\,m^{-3}}$],
$\kappa = \kappa(\boldsymbol{x}, t)$ is the compressibility [$\mathrm{Pa^{-1}}$],
$f_k = f_k(\boldsymbol{x}, t)$ is the volume source density of force in the $\boldsymbol{i}_k$-direction [$\mathrm{N\,m^{-3}}$],
$q = q(\boldsymbol{x}, t)$ is the volume source density of volume injection rate [$\mathrm{s^{-1}}$].

It is assumed that the constitutive parameters ρ and κ are sufficiently smooth functions of the spatial coordinates, that they are constant outside a sphere of finite radius, and, finally, that they are isotropic and time independent. Equation (3.6) is a manifestation of Newton's second law of motion and equation (3.7) is the acoustic version of Hooke's law[7]. Substitution of the former into the latter yields

$$\partial_t^2 p - c^2 \rho \partial_k \frac{1}{\rho} \partial_k p = h, \tag{3.8}$$

where $c = 1/\sqrt{\rho\kappa}$ is identified as the velocity with which the compressional waves are traveling and h is a notional source function depending on f_k and q. The operator $\Box = \partial_t^2 - c^2 \rho \partial_k \frac{1}{\rho} \partial_k$ is recognized as the wave operator or D'Alembertian. A solution can be obtained if initial or Cauchy conditions, and boundary conditions are specified. General solutions and general characteristics of the resulting Cauchy problem[8], such as reversibility, no smoothing action, finite propagation speed, etc., can be found for example in Dautray and Lions (1992).

3.2.1 The two-way wave equation

The time independence of the constitutive parameters favors the use of the temporal Fourier transform, which makes the time independence of the constitutive parameters manifest. The application of the temporal Fourier transform (cf. equation 2.34) to equations (3.6) and (3.7) yields

$$j\omega\rho V_k + \partial_k P = F_k, \tag{3.9}$$

$$j\omega\kappa P + \partial_r V_r = Q, \tag{3.10}$$

[7] Hooke formulated his law of proportionality of stress and strain in 1678. He used the simple words: "Ut tensio sic vis".

[8] The Cauchy problem associated with the wave equation consists of equation (3.8) together with Dirichlet or Neumann boundary conditions and Cauchy conditions, which can take the following form

$$\begin{cases} p(\boldsymbol{x}, t)_{|\boldsymbol{x}\in\Gamma} = 0 \ \text{(Dirichlet condition)} \ \text{or} \ \nabla p(\boldsymbol{x}, t) \cdot \boldsymbol{n}_{|\boldsymbol{x}\in\Gamma} = 0 \ \text{(Neumann condition)} \\ p(\boldsymbol{x}, 0) = p_0(\boldsymbol{x}) \qquad \partial_t p(\boldsymbol{x}, 0) = p_1(\boldsymbol{x}) \ \text{(Cauchy condition).} \end{cases}$$

where $V_k = V_k(\boldsymbol{x}, \omega)$, $P = P(\boldsymbol{x}, \omega)$, $F_k = F_k(\boldsymbol{x}, \omega)$, and $Q = Q(\boldsymbol{x}, \omega)$ are the temporal Fourier transforms of v_k, p, f_k and q, respectively. A subsequent elimination of the horizontal components of the particle velocity V_1 and V_2 yields a coupled system of partial differential equations for the pressure P and the vertical component of the particle velocity V_3 as stated by (Claerbout, 1971; Wapenaar and Berkhout, 1989)

$$\partial_3 \boldsymbol{Q} + j\boldsymbol{\mathcal{A}}\boldsymbol{Q} = \boldsymbol{D}, \tag{3.11}$$

which is a special case of the general matrix system discussed by Volterra, as Gilbert and Backus (1966) point out. Equation (3.11) is referred to as the two-way wave equation. The wave vector $\boldsymbol{Q}$ and the notional source vector $\boldsymbol{D}$ are defined by

$$\boldsymbol{Q} = \begin{pmatrix} P \\ V_3 \end{pmatrix}(\boldsymbol{x}, \omega), \quad \boldsymbol{D} = \begin{pmatrix} F_3 \\ Q - \frac{1}{j\omega}\partial_\mu \left(\frac{F_\mu}{\rho}\right) \end{pmatrix}(\boldsymbol{x}, \omega),$$

and the two-way (matrix) wave operator $\boldsymbol{\mathcal{A}}$ is given by

$$\boldsymbol{\mathcal{A}} = \begin{pmatrix} 0 & \omega\rho \\ \frac{1}{\omega\rho^{1/2}}(\mathcal{H}_2 \rho^{-1/2} \cdot) & 0 \end{pmatrix},$$

where the Helmholtz operator $\mathcal{H}_2$ can be written as

$$\mathcal{H}_2 = k^2(\boldsymbol{x}) + \partial_\mu \partial_\mu, \tag{3.12}$$

with the modified wavenumber $k(\boldsymbol{x})$ given by[9]

$$k^2(\boldsymbol{x}) = \left(\frac{\omega}{c(\boldsymbol{x})}\right)^2 - \frac{3(\partial_\mu \rho)(\partial_\mu \rho)}{4\rho^2} + \frac{(\partial_\mu \partial_\mu \rho)}{2\rho}.$$

If the density does not vary with respect to the lateral coordinates $\boldsymbol{x}_L = (x_1, x_2)$, the modified wavenumber reduces to its normal form according to

$$k(\boldsymbol{x}) = \frac{\omega}{c(\boldsymbol{x})}.$$

For further details concerning the substitutions required for the nice functional form of $\mathcal{H}_2$ the reader should consult (Brekhovskikh, 1960, p. 162), Wapenaar and Berkhout (1989) or de Hoop (1992).

Remark 3.1: Operator notation

Except for the operator ∂_3, in this chapter an operator is working in the plane perpendicular to the direction of preference. The action of an operator $\mathcal{A}$ on the field quantity $\psi(\boldsymbol{x}_L)$ can be explicitly written as

$$(\mathcal{A}\,\psi)(\boldsymbol{x}_L) = \int_\Omega \mathrm{A}(\boldsymbol{x}_L; \boldsymbol{x}'_L)\, \psi(\boldsymbol{x}'_L)\, \mathrm{d}^2\boldsymbol{x}'_L, \tag{3.13}$$

[9] The introduction of the modified wavenumber is not strictly required. However, it is advantageous since it results in the nice functional form of the Helmholtz operator.

where the kernel $\mathrm{A}(\boldsymbol{x}_L;\boldsymbol{x}'_L)$ *is the spatial representation of the operator* $\mathcal{A}$ *(cf. section 2.2). In this chapter the field* $\psi(\boldsymbol{x}_L)$ *represents a monochromatic wave field at a fixed depth level. The operator* $\mathcal{A}$ *and its kernel are symbolically related by*

$$\mathrm{A}(\boldsymbol{x}_L;\boldsymbol{x}'_L) = \mathcal{A}\,\delta(\boldsymbol{x}_L - \boldsymbol{x}'_L). \tag{3.14}$$

What does the solution of the two-way wave equation (3.11) look like? Consider the homogeneous Cauchy problem associated with (3.11), which is given by

$$\begin{cases} \partial_3 \boldsymbol{Q} + j\boldsymbol{\mathcal{A}}\boldsymbol{Q} = 0 \\ \boldsymbol{Q}_{|x_3=0} = \boldsymbol{Q}_0. \end{cases} \tag{3.15}$$

In the case that the two-way wave operator is independent of the x_3-coordinate, which denotes the direction of preference, the solution of the homogeneous Cauchy problem can be written as

$$\boldsymbol{Q}(x_3) = \exp(-j\boldsymbol{\mathcal{A}}x_3)\boldsymbol{Q}_0. \tag{3.16}$$

Since the operator matrix $\boldsymbol{\mathcal{A}}$ is neither self-adjoint nor normal[10], it is not a priori clear that a meaning can be given to this expression in the sense of functional calculus. A straightforward evaluation of the exponent in equation (3.16) yields a structure well-known in time evolution problems (Reed and Simon, 1979; Wilcox, 1984; Dautray and Lions, 1990; Faris, 1995)

$$\boldsymbol{Q}(x_3) = \begin{pmatrix} \rho^{1/2}\cos(\sqrt{\mathcal{H}_2}\,x_3)\rho^{-1/2} & -j\omega\rho^{1/2}\sqrt{\mathcal{H}_2}^{-1}\sin(\sqrt{\mathcal{H}_2}\,x_3)\rho^{1/2} \\ \frac{1}{j\omega\rho^{1/2}}\sqrt{\mathcal{H}_2}\sin(\sqrt{\mathcal{H}_2}\,x_3)\rho^{-1/2} & \rho^{-1/2}\cos(\sqrt{\mathcal{H}_2}\,x_3)\rho^{1/2} \end{pmatrix}\boldsymbol{Q}_0,$$

but this expression does not result in a stable marching algorithm. The question whether or not the functions of operators arising in this equation exist and how they can be computed will be addressed in the next section where an alternative and stable approach is followed. The important point to notice with respect to the last equation is the fact that functions of the Helmholtz operator are required in order to evolve the wave field from one depth level to another. What should one do in case $\mathcal{H}_2$ is varying as a function of depth? In that case one would like to separate propagation and scattering due to medium variations in the depth direction such that the wave field can still be explicitly extrapolated. This objective can be achieved with the help of the one-way wave equation to be derived in section 3.2.2. It turns out that functions of the Helmholtz operator are essential in the one-way wave equation as well. Consequently, in section 3.3 the attention will be focussed on the interpretation of the Helmholtz operator and functions of the Helmholtz operator.

[10] A normal operator $\mathcal{A}$ commutes with its adjoint according to $[\mathcal{A}\,,\,\mathcal{A}^\dagger] = 0$, where the commutator is defined in equation (2.29).

3.2.2 The one-way wave equation

As stated at the end of the previous subsection, a method is looked for which explicitly separates the scattering process along the direction of preference and the propagation process in the direction of preference. This separation can be achieved by introducing a composition operator $\boldsymbol{\mathcal{L}}$ and a decomposition operator $\boldsymbol{\mathcal{L}}^{-1}$ which diagonalize the two-way wave operator $\boldsymbol{\mathcal{A}}$ in equation (3.11). Analogous to the decomposition approach in horizontally layered media (see Ursin, 1983, for an overview), operators $\boldsymbol{\mathcal{L}}$ and $\boldsymbol{\mathcal{H}}$ can be found such that (Claerbout, 1971; Wapenaar and Berkhout, 1989; de Hoop, 1992)

$$\boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{L}}\boldsymbol{\mathcal{H}}\boldsymbol{\mathcal{L}}^{-1}, \tag{3.17}$$

where $\boldsymbol{\mathcal{H}}$ is the required diagonal operator given by

$$\boldsymbol{\mathcal{H}} = \begin{pmatrix} \mathcal{H}_1 & 0 \\ 0 & -\mathcal{H}_1 \end{pmatrix}.$$

The operator $\mathcal{H}_1$ is related to the Helmholtz operator $\mathcal{H}_2$ via

$$\mathcal{H}_2 = \mathcal{H}_1\mathcal{H}_1, \tag{3.18}$$

and it is referred to as the square-root operator. Note that $\mathcal{H}_1$ is a pseudodifferential operator (Fishman and McCoy, 1984a,b; Shubin, 1987; de Hoop, 1992). A whole class of composition operators $\boldsymbol{\mathcal{L}}$, all leading to different representations of the scattering process in the horizontal space, exist. Here, the vertical-acoustic-power-flux normalization is utilized, which amounts to the following composition and decomposition operators

$$\boldsymbol{\mathcal{L}} = \begin{pmatrix} \mathcal{L}_1 & \mathcal{L}_1 \\ \mathcal{L}_2 & -\mathcal{L}_2 \end{pmatrix} \qquad \boldsymbol{\mathcal{L}}^{-1} = \frac{1}{2}\begin{pmatrix} \mathcal{L}_1^{-1} & \mathcal{L}_2^{-1} \\ \mathcal{L}_1^{-1} & -\mathcal{L}_2^{-1} \end{pmatrix}$$

with

$$\mathcal{L}_1 = \left(\frac{\omega\rho}{2}\right)^{1/2}\mathcal{H}_1^{-1/2}, \quad \text{and} \quad \mathcal{L}_2 = \left(\frac{1}{2\omega\rho}\right)^{1/2}\mathcal{H}_1^{1/2}.$$

Other well-known normalizations are the acoustic-pressure normalization and the vertical-particle-velocity normalization (de Hoop, 1992). They may lead to different intermediate results but the expressions P and V_3 will not depend on the choice of the normalization. Substituting equation (3.17) in equation (3.11) and letting the operator $\boldsymbol{\mathcal{L}}^{-1}$ work on the left-hand and right-hand side yield a coupled system of partial differential equations

$$\partial_3\boldsymbol{P} + j\boldsymbol{\mathcal{H}}\boldsymbol{P} = \boldsymbol{S} + \boldsymbol{\Theta}\boldsymbol{P}. \tag{3.19}$$

which is hereafter referred to as the one-way wave equation. The wave vectors $\boldsymbol{Q}$ and $\boldsymbol{P}$ and the notional source vectors $\boldsymbol{D}$ and $\boldsymbol{S}$ are related by

$$\boldsymbol{Q}(\boldsymbol{x}_L, x_3, \omega) = \boldsymbol{\mathcal{L}}\boldsymbol{P}(\boldsymbol{x}_L, x_3, \omega), \quad \text{and} \quad \boldsymbol{D}(\boldsymbol{x}_L, x_3, \omega) = \boldsymbol{\mathcal{L}}\boldsymbol{S}(\boldsymbol{x}_L, x_3, \omega).$$

Note that in the notation of the spatial coordinates, the separation between the horizontal coordinates and the vertical coordinate is made explicit by writing $(\boldsymbol{x}_L, x_3)$ instead of $\boldsymbol{x}$. In the remainder of this chapter the dependence on the angular frequency ω will be suppressed. The second term on the right-hand side of equation (3.19) is defined by

$$\boldsymbol{\Theta}\boldsymbol{P} = \left(-\boldsymbol{\mathcal{L}}^{-1}\partial_3\boldsymbol{\mathcal{L}}\right)\boldsymbol{P} \triangleq \begin{pmatrix} \mathcal{T} & -\mathcal{R} \\ -\mathcal{R} & \mathcal{T} \end{pmatrix}\boldsymbol{P},$$

and can be seen as a secondary source term accounting for scattering along the direction of preference, i.e for the coupling between downgoing and upgoing waves. The operator $\mathcal{T}$ is the transmission operator and $\mathcal{R}$ is the reflection operator. They are given by

$$\mathcal{T} = -1/2\left(\mathcal{L}_1^{-1}\partial_3\mathcal{L}_1 + \mathcal{L}_2^{-1}\partial_3\mathcal{L}_2\right),$$

and

$$\mathcal{R} = 1/2\left(\mathcal{L}_1^{-1}\partial_3\mathcal{L}_1 - \mathcal{L}_2^{-1}\partial_3\mathcal{L}_2\right).$$

In the absence of vertical variations in the medium parameters the secondary source term vanishes. Hence, the objective I have set myself, namely to separate the propagation in the direction of preference and the scattering in the direction of preference, is fulfilled. It is therefore justified to explicitly identify the elements of the one-way wave vector $\boldsymbol{P}$ as a down- and upgoing wave field, i.e.

$$\boldsymbol{P} = \begin{pmatrix} P^+ \\ P^- \end{pmatrix},$$

where the plus-sign denotes a downgoing wave field and the minus-sign an upgoing wave field. In a similar fashion the notional source vector $\boldsymbol{S}$ can be written as $\boldsymbol{S} = (S^+, S^-)^T$, where S^+ (S^-) is the notional source that initiates the downgoing (upgoing) wave field. The different roles the operator $j\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\Theta}$ play in equation (3.19), are summarized in Table 3.1. The operator $j\boldsymbol{\mathcal{H}}$ accounts for (downward/upward) propagation and the operator $\boldsymbol{\Theta}$ for scattering due to vertical variations of the constitutive parameters. Both the operator $j\boldsymbol{\mathcal{H}}$ and $\boldsymbol{\Theta}$ account implicitly for scattering due to horizontal variations in the constitutive parameters.

Notwithstanding the fact that the one-way wave equation separates propagation and scattering in the direction of preference, nothing has been actually solved so

	$j\mathcal{H}$	Θ
Propagation	×	
Vertical scattering		×
Horizontal scattering	×	×

Table 3.1 *Downward/upward propagation is accounted for by the operator $j\mathcal{H}$. Vertical scattering is accounted for by the operator Θ. Horizontal scattering is accounted for by both operators (Wapenaar and Grimbergen, 1996).*

far: the two-way wave equation (3.11) has only been replaced by the one-way wave equation (3.19). The solution of the one-way wave equation starts with the solution of the homogeneous one-way wave equation and with the all-important Green's function. The Green's function or elementary solution $\mathbf{G}$ of equation (3.19) is the solution in distributional sense of

$$\partial_3 \mathbf{G} + j\mathcal{H}\mathbf{G} = \delta(\boldsymbol{x}_L - \boldsymbol{x}'_L)\delta(x_3 - x'_3)\mathbf{I}, \tag{3.20}$$

where

$$\delta(\boldsymbol{x}_L - \boldsymbol{x}'_L) \triangleq \delta(x_1 - x'_1)\delta(x_2 - x'_2), \quad \text{and} \quad \mathbf{I} \triangleq \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In equation (3.20) the coupling of downgoing and upgoing waves due to medium variations in the direction of preference has been set to zero. The matrix kernel

$$\mathbf{G}(\boldsymbol{x}_L, x_3; \boldsymbol{x}'_L, x'_3) = \begin{pmatrix} \mathrm{G}^+ & 0 \\ 0 & \mathrm{G}^- \end{pmatrix} (\boldsymbol{x}_L, x_3; \boldsymbol{x}'_L, x'_3)$$

signifies propagation in the actual medium. In accordance with time evolution problems for the full wave equation (3.8) or equivalent time evolution problems in quantum mechanics (Dautray and Lions, 1992; Reed and Simon, 1979), the Green's kernels $\mathrm{G}^{\pm}$ can be written as

$$\mathrm{G}^{\pm}(\boldsymbol{x}_L, x_3; \boldsymbol{x}'_L, x'_3) = \pm H(\pm(x_3 - x'_3))\mathrm{W}^{\pm}(\boldsymbol{x}_L, x_3; \boldsymbol{x}'_L, x'_3). \tag{3.21}$$

Here $H(x)$ is the Heaviside or step function defined by

$$H(x) \triangleq \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{if } x < 0, \end{cases}$$

and the kernels $\mathrm{W}^{\pm}(\boldsymbol{x}_L, x_3; \boldsymbol{x}'_L, x'_3)$, denoted as propagator kernels, are solutions of the homogeneous Cauchy problem

$$\begin{cases} \partial_3 \mathrm{W}^+ \pm j\mathcal{H}_1 \mathrm{W}^{\pm} = 0 \\ (\mathrm{W}^{\pm})_{|x_3 = x'_3} = \delta(\boldsymbol{x}_L - \boldsymbol{x}'_L). \end{cases} \tag{3.22}$$

The kernel W^+ is the propagator kernel for primary downgoing waves and the kernel W^- is the propagator kernel for primary upgoing waves. In case the medium parameters are not varying between x_3 and x_3' [11], the solution of this Cauchy problem is easily recognized to be (cf. equations 3.15 and 3.16)

$$\mathrm{W}^{\pm}(\boldsymbol{x}_L, x_3; \boldsymbol{x}_L', x_3') = \exp(\mp j(x_3 - x_3')\mathcal{H}_1)\,\delta(\boldsymbol{x}_L - \boldsymbol{x}_L'), \tag{3.23}$$

which can be written, using the symbolic notation of equation (3.14), as

$$\mathrm{W}^{\pm}(\boldsymbol{x}_L, x_3; \boldsymbol{x}_L', x_3') = \mathcal{W}^{\pm}(x_3; x_3')\delta(\boldsymbol{x}_L - \boldsymbol{x}_L'), \tag{3.24}$$

and, consequently,

$$\mathcal{W}^{\pm}(x_3; x_3') = \exp(\mp j(x_3 - x_3')\mathcal{H}_1). \tag{3.25}$$

Note that $\mathcal{W}^{\pm}$ is referred to as the propagator, and $\mathrm{W}^{\pm}$ as the propagator kernel.

Am I ready? No, (1) I have to give a meaning to the rather symbolic expression of equation (3.25) containing functions of operators, and (2) I still have to solve the full one-way wave equation (3.19). After all, the goal is to infer information on the medium parameters from the seismic data acquired at the surface. Let me start with the latter. The expression for the Green's kernel has been found (but it has not been computed yet). The second step in the solution of the one-way wave equation consists of the incorporation of the scattering process in the direction of preference through the secondary source term $\boldsymbol{\Theta P}$ of equation (3.19). The Green's kernel of equation (3.21) is a solution for propagation in the real medium; it solves the scattering process only in the plane perpendicular to the direction of preference. Given the Green's kernel and given a source distribution at one depth level, the implicit solution for the wave field $\boldsymbol{P}$ can be derived from equations (3.19) and (3.20) as

$$\boldsymbol{P}(\boldsymbol{x}_L, x_3) = \int \mathbf{G}\boldsymbol{S}\,\mathrm{d}^2\boldsymbol{x}_L' + \int \mathbf{G}\boldsymbol{\Theta}\boldsymbol{P}\,\mathrm{d}^2\boldsymbol{x}_L'\,\mathrm{d}x_3', \tag{3.26}$$

[11] For the more general situation life is not that easy. The most general solution for the propagator kernel can be derived via a path integral representation to be

$$\mathrm{W}^{\pm}(\boldsymbol{x}_L, x_3; \boldsymbol{x}_L', x_3') = \lim_{N\to\infty}\left(\frac{1}{2\pi}\right)^{2N} \int \prod_{l=1}^{N-1} \mathrm{d}^2\boldsymbol{x}_L^l \prod_{l=0}^{N-1} \mathrm{d}^2\boldsymbol{k}_L^l$$
$$\exp\left(-j\left[\sum_{l=0}^{N-1} k_\mu^l(x_\mu^l - x_\mu^{l+1}) \pm h_1(x_\mu^l, k_\mu^l, x_3 + l\Delta\zeta)\Delta\zeta\right]\right),$$

where h_1 is the left symbol of the square-root operator (Fishman and McCoy, 1984b; Shubin, 1987; de Hoop, 1992), where $\Delta\zeta = |x_3 - x_3'|/N$, and where $\boldsymbol{k}_L = (k_1, k_2)$ is the spatial Fourier pair of $\boldsymbol{x}_L$. The contribution in $(\boldsymbol{x}_L, x_3)$ from $(\boldsymbol{x}_L', x_3')$ can be seen as the summation of the contributions of all possible paths in phase space between the two points. The concept of path integrals has been introduced by Feynman in the 1940s, see, for example, Feynman and Hibbs (1965) or DeWitt-Morette et al. (1979).

which has the same structure as a Lippmann-Schwinger equation (Stolt and Weglein, 1985; Fokkema and van den Berg, 1993). The implicit equation (3.26) can be solved by a Neumann series expansion, also referred to as a Bremmer coupling series (Bremmer, 1951; Corones, 1975). For the application in seismic migration, a linearized version of the Bremmer coupling series will be utilized. The linearized version is obtained if only the first term with the scattering operator $\mathcal{R}$ is taken into account. Consider a configuration in which the upper half-space $x_3 < x_{3,0}$ is homogeneous and choose sources and receivers in this upper half-space. In particular, we consider the primary upgoing response $P^-(\boldsymbol{x}_L^{\mathrm{r}}, x_3^{\mathrm{r}})$ for $x_3^{\mathrm{r}} < x_{3,0}$, related to the source function for downgoing waves $S^+(\boldsymbol{x}_L^{\mathrm{s}}, x_3^{\mathrm{s}})$ for $x_3^{\mathrm{s}} < x_{3,0}$. Then the model for the primary upgoing response or the primary reflection data $P^-(\boldsymbol{x}_L^{\mathrm{r}}, x_3^{\mathrm{r}})$ can be easily derived from the Bremmer series

$$P^-(\boldsymbol{x}_L^{\mathrm{r}}, x_3^{\mathrm{r}}) = \int_{x_3 > x_{3,0}} \int_{\boldsymbol{x}_L'} \int_{\boldsymbol{x}_L''} \int_{\boldsymbol{x}_L^{\mathrm{s}}} \mathrm{W}^-(\boldsymbol{x}_L^{\mathrm{r}}, x_3^{\mathrm{r}}; \boldsymbol{x}_L', x_3)\, \mathrm{R}(\boldsymbol{x}_L', x_3; \boldsymbol{x}_L'', x_3)\, \mathrm{W}^+(\boldsymbol{x}_L'', x_3; \boldsymbol{x}_L^{\mathrm{s}}, x_3^{\mathrm{s}})\, S^+(\boldsymbol{x}_L^{\mathrm{s}}, x_3^{\mathrm{s}})\, \mathrm{d}^2\boldsymbol{x}_L^{\mathrm{s}}\, \mathrm{d}^2\boldsymbol{x}_L''\, \mathrm{d}^2\boldsymbol{x}_L'\, \mathrm{d}x_3, \quad (3.27)$$

which allows for a more compact notation if the kernels $\mathrm{W}^{\pm}$ and R are replaced by their operators $\mathcal{W}^{\pm}$ and $\mathcal{R}$, respectively. The primary data representation can thus be written as

$$P^-(\boldsymbol{x}_L^{\mathrm{r}}, x_3^{\mathrm{r}}) = \int_{x_3 > x_{3,0}} \mathcal{W}^-(x_3^{\mathrm{r}}; x_3)\, \mathcal{R}(x_3)\, \mathcal{W}^+(x_3; x_3^{\mathrm{s}})\, S^+(x_3^{\mathrm{s}})\, \mathrm{d}x_3. \quad (3.28)$$

It is the monochromatic one-way representation of the primary reflection data. It is an adequate model of the seismic experiment after decomposition in up- and downgoing waves and after surface related multiple elimination (Berkhout, 1982). Refinements of the model on the right-hand side of equation (3.27) can be obtained by including fine-layering effects in the propagators, yielding the so-called generalized primary data representation (Wapenaar, 1996b). The fine-layering effects influence the amplitude and the travel time. In practice, the travel time delay is automatically dealt with while looking for an optimum macro model in which the propagators have to be defined.

The representation of a seismic experiment of equation (3.27) does not only lead to a more compact notation but it allows also for an easy realization of the operators in other mathematical domains, as has been discussed in chapter 2 in considerable detail. Representations of the primary reflection data in other domains will be discussed in chapter 4.

The second of the remaining two issues raised below equation (3.25) on the preceding page has been addressed. The first issue concerns the interpretation of the

Helmholtz operator and especially functions of the Helmholtz operator. Functions of operators can be given a meaning by a modal or spectral decomposition, which will be the subject of the next section.

3.3 Operators and modal decomposition

The previous section ended with the linearized model for a seismic experiment. Of key importance here is the propagator, given by

$$\mathcal{W}^{\perp}(x_3; x_3') = \exp(\mp j(x_3 - x_3')\mathcal{H}_1). \tag{3.29}$$

The propagator is a function of an operator and we are faced with the question how to give a meaning to it. An expansion of the Helmholtz operator in eigenfunctions will be utilized in this section. The advantage of the decomposition in eigenfunctions over competing methods, which have been briefly discussed in the introduction of this chapter, is related to the accuracy and stability and to the physical insight the method provides. The advantages and disadvantages will be worked out in more detail in section 3.4.3. The material presented in this section is closely related to the general theory presented in section 2.2.

Example 3.1 *Consider the electron of a hydrogen atom*[12]*. A discrete set of bound states exists, corresponding to the situation where the electron is in the influence region of the nucleus, and a continuous set of free states in which the electron can move if it does not feel the influence of the nucleus anymore. The bound states can in terms of waves be seen as states in which the wavelength of the electron fits an integer number of times. Free electrons, on the other hand, do not feel any restrictions in their movements. A modal decomposition of the Hamiltonian of the hydrogen atom exactly shows which states can occur.*

A similar situation can occur for acoustic waves. A modal decomposition kills two (or more) birds with one stone:

1. A modal decomposition of the Helmholtz operator shows the states, the eigenfunctions, that are "allowed" to exist in the plane perpendicular to the direction of preference. Each physical wave field should be a linear combination of the allowed states according to the superposition principle.

[12]The non-relativistic Hamiltonian of the electron is given by (Messiah, 1958)

$$\hat{\mathrm{H}} = -\frac{h}{2m}\partial_\ell\partial_\ell - \frac{e^2}{r},$$

where h is Planck's constant, m is the mass of the electron, e is the electron's charge, and r is the radial distance to the center of the proton. Note that the proton is considered to be infinitely heavy.

2. The action of an operator on one of its eigenfunctions reduces the action of the operator to a simple multiplication with the eigenvalue corresponding to the eigenfunction. Moreover, if an adequate function space is chosen, the set of eigenfunctions forms an orthogonal basis. In the orthogonal basis the action of a function of an operator reduces to a multiplication as well. Hence, the following two important relations turn out to be at our disposal (cf. section 2.2)

$$\begin{aligned} \mathcal{A}\phi_\lambda(x) &= \lambda\phi_\lambda(x), \\ f(\mathcal{A})\phi_\lambda(x) &= f(\lambda)\phi_\lambda(x), \end{aligned}$$

where λ and $\phi_\lambda(x)$ belong to the set of eigenvalues and the corresponding set of eigenfunctions, respectively. If the operator $\mathcal{A}$ has the aforementioned properties, the kernel A of $\mathcal{A}$ or the kernel K of the operator $f(\mathcal{A})$ can be easily obtained.

3. The set of eigenvalues forms the spectrum of the operator, a notion to be refined in section 3.3.2. A close analysis of the spectrum reveals useful information on the character of the problem at hand. For example, the type of fields to be expected can be forecasted. But also, the well-posedness can be examined: certain parts of the spectrum are especially sensitive to slight changes in the operator. The absence of such a part of the spectrum ensures the stability of the problem.

Though trivially stated, the real world is unfortunately a little bit more complex. On which space of functions should the Helmholtz operator $\mathcal{H}_2$ work? In what sense and on what grounds can the spectrum be subdivided? Can every function of the Helmholtz operator easily be computed? These questions will be addressed here. While answering, the theory of the Hamiltonian in non-relativistic quantum mechanics will be heavily leaned upon, because the similarity (equality up to a factor) of this operator and the Helmholtz operator is striking. Operators in quantum physics are subject of a vast amount of literature. With respect to operator theory I made significant use of Reed and Simon (1972, 1975, 1979, 1978), Dunford and Schwartz (1963), Dautray and Lions (1992), and Souillard (1986). The analysis applies to the Helmholtz operator for a 3-D medium configuration. The examples in section 3.4 are carried out in 2-D configurations.

The flow of the remainder of the section will be as follows. In section 3.3.1 a proper subspace of $L^2(\mathbb{R}^2)$ will be defined on which the Helmholtz operator has the required properties. Section 3.3.2 is devoted to theoretical considerations concerning the spectrum of a self-adjoint operator and to the theory behind the modal decomposition. In sections 3.3.3 and 3.3.4 the obtained knowledge will be applied to the Helmholtz operator. In section 3.3.3 the spectrum of the Helmholtz operator will be analyzed. In section 3.3.4 the attention is focused on the decomposition of

an arbitrary wave field in terms of the eigenfunctions and on the computation of the kernels of the square-root operator and the kernel of the propagator.

3.3.1 Self-adjoint operators

Berezin and Shubin (1991) rephrased one of Von Neumann's postulates (1932) according to

> *"The states of a quantum mechanical system are described by non-zero vectors of a complex separable Hilbert space H, two vectors describing the same state if and only if they differ only by a non-zero complex factor. Each observable corresponds to a certain (unique) linear self-adjoint operator."*

This is one of the basic postulates of quantum mechanics and it shows the importance of self-adjoint forms of operators. As soon as it is shown that an operator is self-adjoint, it is known that the eigenvalues (in quantum mechanics the observables) are real, that the eigenfunctions form an orthogonal basis, and consequently, that functions of the self-adjoint operator can be computed. This is only true for self-adjoint operators (Reed and Simon, 1972).

In order to find out whether an operator is self-adjoint, the following notions are of vital importance.

1. *Adjoint operator*
   The adjoint operator $\mathcal{A}^\dagger$ of $\mathcal{A}$ can be introduced by

$$\langle \mathcal{A}\phi, \psi \rangle = \langle \phi, \mathcal{A}^\dagger \psi \rangle \qquad \forall\, \phi \in \mathtt{dom}(\mathcal{A}) \quad \text{and} \quad \forall\, \psi \in \mathtt{dom}(\mathcal{A}^\dagger). \tag{3.30}$$

   Generally the domains of $\mathcal{A}$ and $\mathcal{A}^\dagger$, $\mathtt{dom}(\mathcal{A})$ and $\mathtt{dom}(\mathcal{A}^\dagger)$, do not coincide. They can even be quite different (Reed and Simon, 1972, p. 252).

2. *Symmetric operator*
   An operator $\mathcal{A}$ is symmetric if

$$\langle \mathcal{A}\phi, \psi \rangle = \langle \phi, \mathcal{A}\psi \rangle \qquad \forall\, \phi, \psi \in \mathtt{dom}(\mathcal{A}), \tag{3.31}$$

   which is equivalent to stating that

$$\mathcal{A}\phi = \mathcal{A}^\dagger \phi \qquad \forall\, \phi \in \mathtt{dom}(\mathcal{A}), \text{ and}$$
$$\mathtt{dom}(\mathcal{A}) \subseteq \mathtt{dom}(\mathcal{A}^\dagger).$$

3. *Self-adjoint operator*
   An operator $\mathcal{A}$ is self-adjoint iff it is symmetric and

$$\mathtt{dom}(\mathcal{A}) = \mathtt{dom}(\mathcal{A}^\dagger).$$

This point makes clear that the definition of the domain is an essential part of the definition of an operator. Without a domain, an operator does not have a meaning. For different domains an operator can have different effects.

Before I turn to the Helmholtz operator and check whether or not it is symmetric or self-adjoint, let me briefly discuss as an example the second-order derivative operator.

Example 3.2 *Consider the Hilbert space $L^2(\mathbb{R})$. The operator $\mathcal{A} = \mathrm{d}^2/\mathrm{d}x^2$ is not well defined on the whole space; there certainly exist functions which second derivative is not square-integrable any more. A proper subspace of $L^2(\mathbb{R})$ has to be found on which the action is well defined. This subspace turns out to be the second Sobolev space. The second Sobolev space $H^2(\mathbb{R})$ consists of functions the second derivatives of which are square-integrable*[13].

The symmetry of the Helmholtz operator is easily established. Let me take as domain of $\mathcal{H}_2$ the space $C_0^\infty(\mathbb{R}^2)$, i.e. the space of infinitely many times differential functions having compact support[14]. On this domain $\mathcal{H}_2$ is well-defined and it can be easily shown that it is symmetric on this domain, using the fact that the stock terms vanish, i.e.

$$\langle \mathcal{H}_2\phi, \psi\rangle \;=\; \langle \phi, \mathcal{H}_2\psi\rangle\,,$$

with $\phi \in C_0^\infty(\mathbb{R}^2)$. Due to the fact that strong conditions are laid upon the choice of ϕ the choice for ψ is less restricted. Because if $\phi \in C_0^\infty(\mathbb{R}^2)$ then $\mathcal{H}_2\phi \in C_0^\infty(\mathbb{R}^2)$. It can be concluded that $\mathtt{dom}(\mathcal{H}_2) \subseteq \mathtt{dom}(\mathcal{H}_2^\dagger)$. Intuitively it can be understood that in order to get a self-adjoint operator, $\mathtt{dom}(\mathcal{H}_2)$ has to be extended, such that $\mathtt{dom}(\mathcal{H}_2^\dagger)$ gets smaller, up to the moment that both domains "meet each other". If the extension of $\mathtt{dom}(\mathcal{H}_2) = C_0^\infty(\mathbb{R}^2)$ can be done in a unique way, the operator is called essentially self-adjoint (Reed and Simon, 1972, p. 256; 1975, Ch. X). The domain of the Helmholtz operator can be enlarged from infinitely smooth compact support functions to functions of which the second derivative (in distributional sense) is square-integrable. This space is called the second Sobolev space and it is denoted by $H^2(\mathbb{R}^2)$. More technically stated, if $\phi \in \mathtt{dom}(\mathcal{H}_2) = H^2(\mathbb{R}^2)$ then $\mathcal{H}_2\phi \in L^2(\mathbb{R}^2)$. The self-adjointness of the Helmholtz operator has been established, now we are able to investigate the spectral properties.

3.3.2 The spectrum of a self-adjoint operator

The resolvent operator of the operator $\mathcal{A}$ is given by

$$\mathcal{R}_\lambda^{(\mathcal{A})} \triangleq \frac{1}{\lambda\mathcal{I} - \mathcal{A}}, \qquad \lambda \in \rho(\mathcal{A}), \tag{3.32}$$

[13]For a clear and concise introduction on Sobolev spaces the reader is referred to Dautray and Lions (1988).

[14]A function with compact support vanishes outside a finite domain (Reed and Simon, 1972, p. 111).

where the resolvent set $\rho(\mathcal{A})$ defines the domain where the resolvent exists. The spectrum $\sigma(\mathcal{A})$ is the complement in $\mathbb{C}$ of the resolvent set $\rho(\mathcal{A})$. Hence,

$$\sigma(\mathcal{A}) = \mathbb{C} \backslash \rho(\mathcal{A}).$$

The resolvent operator is of course more than only a tool to find the spectrum of an operator. The resolvent operator can be used to define, via a Dunford contour integral, functions of the operator, in particular the spectral projection, which can be seen as an operator generalization of Cauchy's residue theorem. The role of the resolvent will be discussed again in relation to the spectral theorem.

Hereinafter the operator $\mathcal{A}$ is assumed to be self-adjoint, which means that the spectrum is a subset of the real line $\mathbb{R}$. The spectrum of a self-adjoint operator can be subdivided in different parts in various ways, providing as such knowledge of the physical underlying system. The first subdivision has two attractive properties. Firstly, the subspaces associated with the different parts of the spectrum of the operator $\mathcal{A}$, are invariant under the action of $\mathcal{A}$. Secondly, the functions in the different subspaces clearly have a distinct physical nature. The second subdivision provides less physical insight, but it can be easier obtained. It is a subdivision in isolated and non-isolated eigenvalues.

Besides the fact that the spectrum of a self-adjoint operator is a subset of the real line, another property of a self-adjoint operator is the fact that it is possible to associate with its spectrum a continuous equivalent of an orthogonal basis. In the sequel, the two possible subdivisions of the spectrum will be discussed first, then the orthogonal decomposition associated with a self-adjoint operator will be dealt with.

Subdivision of spectrum I —Let me first turn to a part of the spectrum which is well-known and which is denoted as the pure point spectrum $\sigma_{\mathrm{pp}}(\mathcal{A})$. It is defined by those values λ for which the following formula is valid

$$\mathcal{A}\phi_\lambda(x) = \lambda\phi_\lambda(x) \quad \text{for} \quad \phi_\lambda(x) \in \mathtt{dom}(\mathcal{A}). \tag{3.33}$$

Functions $\phi_\lambda(x)$ fulfilling equation (3.33) are called eigenfunctions; the corresponding λ are called eigenvalues. For unbounded operators, such as the Helmholtz operator or such as the second derivative $\mathrm{d}^2/\,\mathrm{d}x^2$, the pure point spectrum generally does not cover the whole spectrum or does not exist at all. The essential element in equation (3.33) is "$\phi_\lambda \in \mathtt{dom}(\mathcal{A})$". There may be points λ in $\sigma(\mathcal{A})$ for which the corresponding "eigenfunctions" do not belong to the domain of the operator or which are even not in the space of square-integrable functions. These functions will be called generalized eigenfunctions. The proper definition of generalized eigenfunctions requires the concept of the "equipped Hilbert space", which allows the generalized eigenfunctions to lay in a space just a little bit larger than the domain of the operator (Vilenkin, 1972, p. 89).

Example 3.3 *Consider again the second-order derivative* $\mathcal{A} = \mathrm{d}^2/\mathrm{d}x^2$ *(see Example 3.2). By applying a Fourier transformation it can easily be shown that the spectrum* $\sigma(\mathcal{A}) = (-\infty, 0]$. *However, the spectrum does not contain eigenvalues, because the functions related to any* $\lambda \in \sigma(\mathcal{A})$, *do not belong to the domain of the operator. The generalized eigenfunctions are given by the Fourier components*

$$\phi_{-\xi^2}(x) = \exp(jx\xi) \qquad \forall\, \xi \in \mathbb{R},$$

which are clearly not square-integrable. In distributional sense they can be given a meaning.

Example 3.4 *Consider the multiplication operator* $\mathcal{A} = x$ *on functions* $\phi \in L^2(a, b)$, *i.e. the space of square-integrable functions on the interval* (a, b). *Hence,*

$$\mathcal{A}\phi(x) = x\phi(x) \qquad \phi(x) \in L^2(a, b).$$

The eigenvalue problem is given by

$$x\phi_\lambda(x) = \lambda\phi_\lambda(x),$$

which is satisfied by functions $\phi_\lambda(x)$ *equal to zero except for* $x = \lambda$ *where it can be different from zero. But in the domain there is no non-zero element having this property. Eigenfunctions, however, exist, namely the* δ*-distributions* $\delta(x - \lambda)$ *which are not living in* $L^2(a, b)$. *The spectrum equals* $\sigma(\mathcal{A}) = (-\infty, \infty)$. *See also section 2.2.*

Two possible parts of the spectrum that have been discussed so far, can be discriminated on the basis of their mathematical structure. The pure point spectrum $\sigma_{\mathrm{pp}}(\mathcal{A})$ consists of the part of the spectrum with which the true eigenvalues can be associated. Another part of the spectrum cannot be associated with functions in the domain of the operator. This part has been discussed in the two examples. It is called the absolute continuous spectrum $\sigma_{\mathrm{ac}}(\mathcal{A})$.

The above described division can be given a firm basis and can be refined with the help of Lebesgue integral theory and the associated abstract measure theory (Reed and Simon, 1972). This abstract theory results in the following refined subdivision:

$$\sigma(\mathcal{A}) = \overline{\sigma_{\mathrm{pp}}(\mathcal{A})} \cup \sigma_{\mathrm{ac}}(\mathcal{A}) \cup \sigma_{\mathrm{sc}}(\mathcal{A}), \tag{3.34}$$

where the $^-$ denotes the closure of $\sigma_{\mathrm{pp}}(\mathcal{A})$, which consists of $\sigma_{\mathrm{pp}}(\mathcal{A})$ and its limiting values; $\sigma_{\mathrm{sc}}(\mathcal{A})$ is the singular continuous spectrum and will be discussed below. Although the above subdivision might seemingly appear out of the blue, the associated subspaces of eigenfunctions have a relatively simple physical meaning. Moreover, the associated subspaces are stable under the action of the operator.

Let us define $H_{\rm pp}$, $H_{\rm ac}$ and $H_{\rm sc}$ as the subspaces associated with $\sigma_{\rm pp}(\mathcal{A})$, $\sigma_{\rm ac}(\mathcal{A})$ and $\sigma_{\rm sc}(\mathcal{A})$ respectively[15]. The subspaces are invariant under the action of $\mathcal{A}$:

$$\mathcal{A}: H_{\rm pp} \rightarrow H_{\rm pp} \tag{3.35}$$

$$\mathcal{A}: H_{\rm ac} \rightarrow H_{\rm ac} \tag{3.36}$$

$$\mathcal{A}: H_{\rm sc} \rightarrow H_{\rm sc}, \tag{3.37}$$

which shows that a useful subdivision has been chosen. The usefulness is emphasized if the associated (generalized) eigenfunctions are studied:

1. Functions $\phi \in H_{\rm pp}$ denote physical states which, in most cases[16], are exponentially decaying outside a finite domain. They are called localized or bound states. For the hydrogen atom these are the states with positive eigenvalues. In acoustics these are the waves trapped in a wave guide. The so-called guided wave modes.

2. Functions $\phi \in H_{\rm ac}$ denote extended physical states showing almost no decaying behavior. Free electrons have a probability to be anywhere. In acoustics a typical example is a plane wave component, having an infinite extension. In acoustics, these states are called the radiating wave modes.

3. Functions $\phi \in H_{\rm sc}$ do not have a univocal interpretation (Reed and Simon, 1972, p. 23). Souillard (1986), although refusing to commit himself, writes that these states "tend to present some kind of self-invariant structure, vanishing over extremely large regions, and taking over again much farther." He calls these states exotic.

Some of the three subspaces can be empty, but not all. It is clear now that the different parts of the spectrum are relevant to unravel the physical problem. A final reason to make the aforementioned breakup is due to the fact that different parts behave differently under perturbations of the parameters. Especially $\sigma_{\rm sc}(\mathcal{A})$ and the corresponding subspace are known to be unstable under perturbations (Reed and Simon, 1978).

Subdivision of spectrum II —Another breakup which turns out to be useful, will be given now. The usefulness of the alternative stems from the fact that an estimation of the bounds of the spectrum, i.e. which values can possibly be attained by the eigenvalues, can be more easily given than in the case of the subdivision

[15]The formal definition (Reed and Simon, 1972, p. 230) is the other way around. Abstract measure theory provides a subdivision of $\mathrm{dom}(\mathcal{A})$, the domain of the operator $\mathcal{A}$, into $H_{\rm pp}$, $H_{\rm ac}$ and $H_{\rm sc}$. Theorem VII.4 in Reed and Simon (1972) states that each of these subspaces is invariant under the action of $\mathcal{A}$ and that each of these subspaces have a complete set of eigenfunctions.

[16]Exceptions can be found in Reed and Simon (1978), p. 222 and further, but also in the examples in section 3.4.

described before. The breakup is solely based upon the structure of the spectrum and the dimension of the associated eigenfunctions and it results in an essential spectrum $\sigma_{\rm ess}(\mathcal{A})$ and a discrete spectrum $\sigma_{\rm disc}(\mathcal{A})$.

An element $\lambda \in \sigma(\mathcal{A})$ is an element of the essential spectrum $\sigma_{\rm ess}(\mathcal{A})$ of $\mathcal{A}$ if the spectral projection in the region $(\lambda - \epsilon, \lambda + \epsilon)$ $(\epsilon > 0)$ is infinite dimensional[17]. On the other hand an element $\lambda \in \sigma(\mathcal{A})$ is an element of the discrete spectrum $\sigma_{\rm disc}(\mathcal{A})$ of $\mathcal{A}$ if the spectral projection in the region $(\lambda - \epsilon, \lambda + \epsilon)$ is finite dimensional. Unlike $\sigma_{\rm pp}(\mathcal{A})$, $\sigma_{\rm ac}(\mathcal{A})$, and $\sigma_{\rm sc}(\mathcal{A})$, the newly defined spectra $\sigma_{\rm ess}(\mathcal{A})$ and $\sigma_{\rm disc}(\mathcal{A})$ are necessarily disjoint. Although the definitions of the two alternative subdivisions come from quite different directions, the subdivisions are certainly not completely independent. Since the real physical insight is provided by the first subdivision it is of considerable importance to relate the two subdivisions. Clearly an element $\lambda \in (\sigma_{\rm ac}(\mathcal{A}) \cup \sigma_{\rm sc}(\mathcal{A}))$ is also an element of the essential part of the spectrum. Also, most eigenvalues –which form the pure point spectrum– are generally part of the discrete spectrum (Reed and Simon, 1972; 1978, Ch. XIII).

Example 3.5 *Consider again the second-order derivative operator* $\mathcal{A} = \mathrm{d}^2/\mathrm{d}x^2$ *(see also Examples 3.2 and 3.3). The spectrum only consists of an absolute continuous part. The pure point spectrum and the singular continuous spectra are empty. The essential spectrum* $\sigma_{\rm ess}(\mathcal{A})$ *coincides with the absolute continuous spectrum.*

Example 3.6 *Consider again the hydrogen atom (see also Example 3.1). The spectrum consists of a pure point spectrum and an absolute continuous spectrum. The discrete spectrum coincides with the pure point spectrum and the essential spectrum coincides with the absolute continuous spectrum.*

Questions to be answered —A quantitative analysis of the full spectrum would ideally answer which kinds of spectra are present and would ideally result in an exact localization of all different kinds of spectra. Due to the fact that we are dealing with an earth that is even in the laterally direction quite irregular, by far not as regular as for example a single hydrogen atom, it is not of interest to answer all questions for the Helmholtz operator in the next section in full quantitative detail. For particular medium configurations, however, a quantitative analysis can be useful for completely understanding the nature of the observed phenomena. A number of interesting rather general questions remain:

- What are the bounds of the essential spectrum?
- What are the bounds of the discrete spectrum?

[17]The formal definition can be found in Reed and Simon (1972, p. 236).

- Can the singular continuous spectrum be excluded? This is an important question, because this part of the spectrum (if present) is relatively unstable under perturbations (Reed and Simon, 1978; Faris, 1995). Here, it is assumed that the singular continuous spectrum does not exist.

The first two questions will be addressed while studying the spectrum of the Helmholtz operator in section 3.3.3.

Spectral theorem and spectral representation — In this section a brief outline will be given of two mutually closely related consequences of self-adjointness. The first one goes under the name of "spectral representation"[18] and it says that to a self-adjoint operator an orthonormal basis is associated in which every function in the domain of the operator can be expanded (Dunford and Schwartz, 1963, Th.5, p. 1209). The second one goes under the name "spectral theorem" and it comes in different forms. It says that the action of a self-adjoint operator to a function in the domain of the operator reduces to a simple multiplication and that the action of a function of a self-adjoint operator reduces to a multiplication with the function of the eigenvalues. The form of the spectral theorem I like most, that is also the most abstract one, is called the "spectral theorem – projection valued measure" (Reed and Simon, 1972, Th.VIII.6, p. 263) given by

$$\mathcal{A}\psi = \int \lambda \, \mathrm{d}\mathcal{P}_\lambda \psi, \tag{3.38}$$

where, loosely speaking, $\mathrm{d}\mathcal{P}_\lambda$ is the infinitesimal projection operator which projects the function ψ on a space spanned by the (generalized) eigenfunctions corresponding to the part of the spectrum $(\lambda, \lambda + \mathrm{d}\lambda)$. The projection operator is defined via the resolvent operator[19]. Seemingly abstract, equation (3.38) combines in one simple formula the action of the operator $\mathcal{A}$ for all physical parts of the spectrum: the pure point spectrum, the absolute continuous spectrum and the singular continuous spectrum. For example for a non-degenerate[20] point spectrum the integral in equation (3.38) takes the form of a summation according to

$$(\mathcal{A}\psi)(x) = \sum_i \tilde{\psi}(\lambda_i)\, \lambda_i \, \phi_{\lambda_i}(x), \tag{3.39}$$

where $\tilde{\psi}(\lambda_i) = \langle \psi, \phi_{\lambda_i} \rangle$ is the representation of ψ in the modal domain. In the case of a non-degenerate absolute continuous spectrum, the abstract integral (3.38) takes

[18] The spectral representation is actually a part of the cited spectral theorem.

[19] See equation (3.32) and the remark concerning the resolvent on page 82. The evaluation of the projection operator goes under the name "Stone's formula" (Reed and Simon, 1972, Th.VII.13), which can be recognized as a Dunford contour integral around the eigenvalue. See also Dunford and Schwartz (1963), de Hoop (1992), and Blok (1995).

[20] The order of degeneracy is defined by the number of independent (generalized) eigenfunctions corresponding to the same eigenvalue. An eigenvalue is called non-degenerate if the degeneracy is of order 1.

the following form

$$(\mathcal{A}\psi)(x) = \int \tilde{\psi}(\lambda)\,\lambda\,\phi_\lambda(x)\,\mathrm{d}\lambda. \tag{3.40}$$

where $\tilde{\psi}(\lambda) = \langle \psi, \phi_\lambda \rangle$ is the representation of ψ in the modal domain. Whereas a division into the pure point, the absolute continuous and the singular continuous spectrum provides the most physical insight, it has been argued that a more practical road uses a division in the essential part of the spectrum and the discrete part of the spectrum. Provided that the spectrum of $\mathcal{A}$ is completely non-degenerate, the action of the operator $\mathcal{A}$ on the function ψ in the domain of $\mathcal{A}$ takes the following form

$$(\mathcal{A}\psi)(x) = \int\limits_{\lambda\in\sigma_{\mathrm{ess}}(\mathcal{A})} \tilde{\psi}(\lambda)\,\lambda\,\phi_\lambda(x)\,\mathrm{d}\lambda + \sum_{\lambda_i\in\sigma_{\mathrm{disc}}(\mathcal{A})} \tilde{\psi}(\lambda_i)\,\lambda_i\,\phi_{\lambda_i}(x), \tag{3.41}$$

which will be written in shorthand symbolic notation as

$$(\mathcal{A}\psi)(x) = \sum\!\!\!\!\!\!\int\limits_{\substack{\lambda\in\sigma_{\mathrm{ess}}(\mathcal{A})\\ \lambda\in\sigma_{\mathrm{disc}}(\mathcal{A})}} \tilde{\psi}(\lambda)\,\lambda\,\phi_\lambda(x)\,\mathrm{d}\lambda. \tag{3.42}$$

Substitution of $\tilde{\psi}(\lambda)$ and $\tilde{\psi}(\lambda_i)$, and changing the order of integration yield in extended format

$$(\mathcal{A}\psi)(x) = \int \left(\int \phi^*_\lambda(x')\,\lambda\,\phi_\lambda(x)\,\mathrm{d}\lambda \right) \psi(x')\,\mathrm{d}x' + \int \left(\sum \phi^*_{\lambda_i}(x')\,\lambda_i\,\phi_{\lambda_i}(x) \right) \psi(x')\,\mathrm{d}x', \tag{3.43}$$

which upon comparison with (3.13) gives an expression for the kernel $\mathrm{A}(x,x')$ of $\mathcal{A}$ according to

$$\mathrm{A}(x,x') = \sum\!\!\!\!\!\!\int\limits_{\substack{\lambda\in\sigma_{\mathrm{ess}}(\mathcal{A})\\ \lambda\in\sigma_{\mathrm{disc}}(\mathcal{A})}} \phi^*_\lambda(x')\,\lambda\,\phi_\lambda(x)\,\mathrm{d}\lambda. \tag{3.44}$$

Note that this expression has to be interpreted in the sense of generalized functions. For operators $\mathcal{K}$ given by $\mathcal{K} = f(\mathcal{A})$ similar equations can be derived with the only difference that λ has to be replaced by $f(\lambda)$. The action of $\mathcal{K} = f(\mathcal{A})$ on $\psi(x)$ thus yields (cf. equation 3.42)

$$(f(\mathcal{A})\psi)(x) = \sum\!\!\!\!\!\!\int\limits_{\substack{\lambda\in\sigma_{\mathrm{ess}}(\mathcal{A})\\ \lambda\in\sigma_{\mathrm{disc}}(\mathcal{A})}} \tilde{\psi}(\lambda)\,f(\lambda)\,\phi_\lambda(x)\,\mathrm{d}\lambda, \tag{3.45}$$

and an expression for the kernel $\mathrm{K}(x, x')$ of $\mathcal{K}$ can be derived as (cf. equation 3.44)

$$\mathrm{K}(x, x') = \sum\!\!\!\!\!\!\int_{\substack{\lambda \in \sigma_{\mathrm{css}}(\mathcal{A}) \\ \lambda \in \sigma_{\mathrm{disc}}(\mathcal{A})}} \phi_\lambda^*(x')\, f(\lambda)\, \phi_\lambda(x)\, \mathrm{d}\lambda. \tag{3.46}$$

Remark 3.2 *In the case that one or all of the spectral parts are degenerate, expressions* (3.41)-(3.46) *should be adjusted accordingly (Messiah, 1958, Sec. 7.13). Let me give a specific example for a spectrum consisting of an absolute continuous part (coinciding with the essential spectrum) and a pure point spectrum (coinciding with the discrete spectrum). Suppose that the absolute continuous part of the spectrum is infinitely degenerate. Assume now that the set of eigenfunctions belonging to one eigenvalue in the absolute continuous spectrum can be labeled with a continuous parameter. If it is possible to write the eigenvalue λ in the absolute continuous spectrum as $\lambda = \lambda(\boldsymbol{\kappa})$ with $\boldsymbol{\kappa} = (\kappa_1, \kappa_2)$, then one can rewrite for example equation* (3.41) *according to*

$$(\mathcal{A}\psi)(x) = \int_{\boldsymbol{\kappa} | \lambda(\boldsymbol{\kappa}) \in \sigma_{\mathrm{css}}(\mathcal{A})} \tilde{\psi}(\boldsymbol{\kappa})\, \lambda(\boldsymbol{\kappa})\, \phi_{\lambda(\boldsymbol{\kappa})}(x)\, \mathrm{d}^2\boldsymbol{\kappa} + \sum_{\lambda_i \in \sigma_{\mathrm{disc}}(\mathcal{A})} \tilde{\psi}(\lambda_i)\, \lambda_i\, \phi_{\lambda_i}(x), \tag{3.47}$$

where $\tilde{\psi}(\boldsymbol{\kappa})$ should be interpreted as $\tilde{\psi}(\boldsymbol{\kappa}) = \langle \psi, \phi_{\lambda(\boldsymbol{\kappa})} \rangle$. The other equations have to be adjusted accordingly. In the case of the Helmholtz operator a degeneracy as discussed in this remark will occur.

3.3.3 Spectrum of the Helmholtz operator

Now that the tools have been developed, the attention will be focused on the Helmholtz operator[21]. A general quantitative analysis will be applied to the spectrum of the Helmholtz operator. It has been assumed that the medium is constant in the plane perpendicular to the direction of preference outside some finite circle (see section 3.2). In other words, the medium is embedded in some homogeneous background. In addition, in this section the depth level is considered to be fixed. Therefore the dependency on x_3 is suppressed in this subsection. As a result of these premises, the operator $\mathcal{H}_2$ (cf. equation 3.12) can be written as

$$\mathcal{H}_2 = k^2(\boldsymbol{x}_L) + \partial_\mu \partial_\mu = \mathcal{H}_2^0 + k_0^2, \tag{3.48}$$

where k_0 is the wavenumber of the homogeneous embedding, and where the perturbed Helmholtz operator $\mathcal{H}_2^0$ is given by

$$\mathcal{H}_2^0 = V(\boldsymbol{x}_L) + \partial_\mu \partial_\mu. \tag{3.49}$$

[21] The theory presented in section 3.3.2 is equally well applicable to any self-adjoint operator, for example to the reduced wave operator arising in wave time evolution problems (Reed and Simon, 1979; Wilcox, 1984).

The function $V(\boldsymbol{x}_L) = k^2(\boldsymbol{x}_L) - k_0^2$, is the potential and it is a function of compact support. It can be seen as a perturbation of the homogeneous Helmholtz operator. The operator $\mathcal{H}_2^0$ bears a lot of resemblance with the Hamiltonian operator in non-relativistic quantum mechanics. The latter is for a general potential V_h in a two dimensional[22] setting given by

$$\hat{\mathrm{H}}(\boldsymbol{x}_L) = -\partial_\mu \partial_\mu + V_h(\boldsymbol{x}_L). \tag{3.50}$$

The Hamiltonian equals (with $V_h = -V$ and with an opposite sign) the operator $\mathcal{H}_2^0$. In Reed and Simon (1978) the spectrum of the Hamiltonian is quantified. Based upon their work it is possible to express upon the bounds of the complete spectrum of $\mathcal{H}_2^0$, the bounds of the essential spectrum of the perturbed Helmholtz operator $\mathcal{H}_2^0$, the bounds of the discrete spectrum of $\mathcal{H}_2^0$, and the bounds of the spectrum of the Helmholtz operator $\mathcal{H}_2$ itself.

Bounded or unbounded operator —The first thing to establish is whether the operator under consideration is a bounded, unbounded or semi-bounded operator. The bounds are the spectral bounds. Consider a function ψ in the domain of $\mathcal{H}_2^0$ [23]. The bounds are defined by the minimum and maximum value that $\langle \mathcal{H}_2^0 \psi, \psi \rangle$ can reach:

$$\langle \mathcal{H}_2^0 \psi, \psi \rangle = \langle \partial_1^2 \psi, \psi \rangle + \langle \partial_2^2 \psi, \psi \rangle + \langle V\psi, \psi \rangle .$$

The first two terms in the right-hand side can be partially integrated to obtain

$$\langle \mathcal{H}_2^0 \psi, \psi \rangle = \underbrace{-\int_{\mathbb{R}^2} \partial_1 \psi \, (\partial_1 \psi)^* \, \mathrm{d}^2 \boldsymbol{x}_L}_{\leq 0} \underbrace{-\int_{\mathbb{R}^2} \partial_2 \psi \, (\partial_2 \psi)^* \, \mathrm{d}^2 \boldsymbol{x}_L}_{\leq 0} + \underbrace{\int_{\mathbb{R}^2} V(\boldsymbol{x}_L) \psi \, \psi^* \, \mathrm{d}^2 \boldsymbol{x}_L}_{\leq \max(V) \, \langle \psi, \psi \rangle} .$$

It can be concluded that the spectrum is semi-bounded from above, according to

$$\sigma(\mathcal{H}_2^0) \in (-\infty, \max(V)]. \tag{3.51}$$

Weyl's essential spectrum theorem —Bounds on the essential spectrum are relatively easy obtained. It does not tell much about the physical nature of the problem, because, generally, σ_{ess} consists of the absolute continuous spectrum, the

[22] In the general theory a Hamiltonian in $\mathbb{R}^n$ is considered.

[23] The domain of $\mathcal{H}_2^0$ equals the domain of $\mathcal{H}_2$. It is discussed in section 3.3.1.

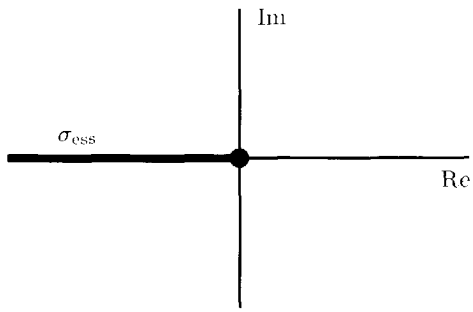


Fig. 3.1 *The essential spectrum of the perturbed Helmholtz operator $\mathcal{H}_2^0$. The essential spectrum equals the essential spectrum of the Laplacian in one or more dimensions.*

singular continuous spectrum and the end points of the pure point spectrum (see section 3.3.2). Since it has been assumed that the singular continuous spectrum is empty, the significance of the estimates of the bounds of the essential spectrum increases. The keynote of Weyl's essential spectrum theorem (Reed and Simon, 1978, Th.XIII.14, Cor.2c, p. 112 and 113) is that if $\mathcal{A}$ is a self-adjoint operator and $\mathcal{C}$ is a relatively compact perturbation of $\mathcal{A}$, the following relation holds

$$\sigma_{\text{ess}}(\mathcal{A} + \mathcal{C}) = \sigma_{\text{ess}}(\mathcal{A}).$$

This theorem can be applied to $\mathcal{H}_2^0$ as defined in equation (3.49): $\mathcal{A}$ should be replaced by the two-dimensional Laplacian $\partial_\mu \partial_\mu$ and $V(\boldsymbol{x}_L)$ should replace the relatively compact[24] perturbation $\mathcal{C}$ of the Laplacian. On that account the essential spectrum of $\mathcal{H}_2^0$ equals the essential spectrum of the two-dimensional Laplacian (see also Example 3.5 on page 86):

$$\sigma_{\text{ess}}(\mathcal{H}_2^0) = \sigma_{\text{ess}}(\partial_\mu \partial_\mu) = (-\infty, 0]. \tag{3.52}$$

The essential spectrum of the perturbed Helmholtz operator covers the negative real axis; it is shown in Figure 3.1. The essential spectrum of $\mathcal{H}_2^0$ has an infinite degeneracy. The set of eigenfunctions corresponding to a single eigenvalue can be labeled with a continuous parameter. According to Remark 3.2 on page 89 this can be achieved by choosing $\lambda = \lambda(\boldsymbol{\kappa})$. Another consequence of Weyl's essential spectrum theorem is the fact that the degeneracy of the essential spectrum is unaltered by the perturbation V.

[24] $V(\boldsymbol{x}_L)$ itself is compact and hence automatically relatively compact. Consequently, the classical Weyl theorem is actually dealt with (Reed and Simon, 1978, Example 3, p. 117).

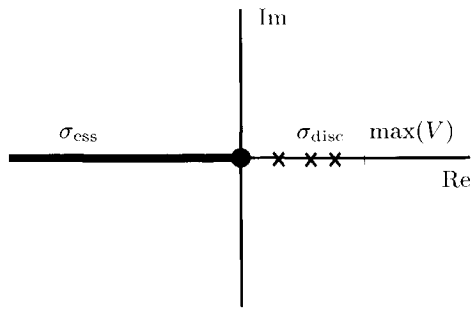


Fig. 3.2 *The spectrum of the perturbed Helmholtz operator $\mathcal{H}_2^0$ is divided into an essential part and a discrete part. The discrete part consists of a finite number of isolated eigenvalues in the interval marked out by the bound of the essential spectrum and the bound of the full spectrum.*

The discrete spectrum – The definition of the upper bound of the spectrum in (3.51) and the location of the essential spectrum in equation (3.52) determine the interval where the elements of the discrete spectrum might lie:

$$\sigma_{\text{disc}}(\mathcal{H}_2^0) \subset (0, \max(V)].$$

It can be shown (Reed and Simon (1978), Th.XIII.6, p. 87; Dunford and Schwartz (1963), Cor.57)[25] that within this interval a finite number of eigenvalues exist, if $V(\boldsymbol{x}_t)$ is decaying sufficiently fast at the borders of its support. Figure 3.2 shows the schematic structure of the spectrum of $\mathcal{H}_2^0$.

Spectrum of $\mathcal{H}_2$ —So far, the discussion concerned the spectrum of $\mathcal{H}_2^0$. The spectrum of $\mathcal{H}_2$ is related to the spectrum of $\mathcal{H}_2^0$ through a simple shift. Heuristically, this can be understood by considering the resolvent defined in equation (3.32). The resolvent of $\mathcal{H}_2^0$ is given by

$$\mathcal{R}_{\lambda_1}^{(\mathcal{H}_2^0)} = \frac{1}{\lambda_1 - \mathcal{H}_2^0}.$$

For λ_1 outside the resolvent set of $\mathcal{H}_2^0$, i.e. for $\lambda_1 \in \sigma(\mathcal{H}_2^0)$, the resolvent is not well-defined. On account of equation (3.48) the resolvent of the Helmholtz operator $\mathcal{H}_2$ can be written as

$$\mathcal{R}_{\lambda}^{(\mathcal{H}_2)} = \frac{1}{\lambda - \mathcal{H}_2} = \frac{1}{(\lambda - k_0^2) - \mathcal{H}_2^0}.$$

[25] Reed and Simon (1978) actually consider a 3-D Hamiltonian, while the operator $\mathcal{H}_2^0$ is defined in two dimensions. On a later page it is stated that the theorem can be extended to other dimensions as well (Reed and Simon, 1978, p. 119).

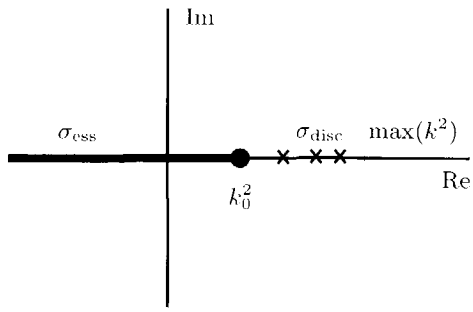


Fig. 3.3 *The spectrum of the Helmholtz operator $\mathcal{H}_2$ is obtained by shifting the spectrum of the perturbed Helmholtz operator $\mathcal{H}_2^0$, as shown in Figure 3.2, over a distance k_0^2.*

The resolvent of $\mathcal{H}_2$ is not well-defined if $\lambda - k_0^2 \in \sigma(\mathcal{H}_2^0)$. Consequently, $\lambda = \lambda_1 + k_0^2 \in \sigma(\mathcal{H}_2)$, which is the expected shift by k_0^2. The spectrum of $\mathcal{H}_2$ is shown in Figure 3.3.

The shift applies equally well to the essential and discrete spectrum: $\sigma_{\text{ess}}(\mathcal{H}_2) = (-\infty, k_0^2]$ and $\sigma_{\text{disc}}(\mathcal{H}_2) \subset (k_0^2, \max(k^2)]$. The generalized eigenfunctions of $\mathcal{H}_2^0$ and $\mathcal{H}_2$ are the same, but note that they have different phase velocities, since they belong to different eigenvalues. The degeneracy of the essential spectrum of the Helmholtz operator is, just as the perturbed Helmholtz operator, infinite. The degeneracy can be taken into account by choosing $\lambda = \lambda(\boldsymbol{\kappa})$ as discussed in Remark 3.2 on page 89.

3.3.4 Expanding the Helmholtz operator and functions of the Helmholtz operator

In this section the spectral theory of the last part of section 3.3.2 will be applied. The theory in that section has been worked out for a non-degenerate spectrum in one dimension. The Helmholtz operator is working in the 2-D plane perpendicular to the direction of preference. Due to the completeness of the basis of eigenfunctions, any function $\psi(\boldsymbol{x}_L)$ from the domain of $\mathcal{H}_2$ can be expanded in terms of the eigenfunctions of $\mathcal{H}_2$. Taking into account the degeneracy of the essential spectrum and assuming that the discrete spectrum is non-degenerate, the expansion can be written as

$$\psi(\boldsymbol{x}_L) = \int_{\mathbb{R}^2} \tilde{\psi}(\boldsymbol{\kappa})\, \phi_{\lambda(\boldsymbol{\kappa})}(\boldsymbol{x}_L)\, \mathrm{d}^2\boldsymbol{\kappa} + \sum_{\lambda_i \in \sigma_{\text{disc}}} \tilde{\psi}(\lambda_i)\, \phi_{\lambda_i}(\boldsymbol{x}_L). \tag{3.53}$$

On account of the results of the previous subsection with respect to the essential part of the spectrum, the relation between the eigenvalue λ and the parameters $\boldsymbol{\kappa}$ is

$$\lambda(\boldsymbol{\kappa}) = k_0^2 - \kappa_\mu \kappa_\mu, \tag{3.54}$$

whence the integral in equation (3.53) over $\mathbb{R}^2$. The expansion coefficients $\tilde{\psi}(\lambda_i)$ in the second term on the right-hand side of (3.53) correspond to the discrete eigenvalues λ_i. The first term on the right-hand side contains the expansion coefficients $\tilde{\psi}(\boldsymbol{\kappa})$, which correspond to the essential part of the spectrum. Together, $\tilde{\psi}(\lambda_i)$ and $\tilde{\psi}(\boldsymbol{\kappa})$ form the representation of ψ in the domain defined by the operator $\mathcal{H}_2$. Equation (3.53) can also be interpreted as an inverse transformation from the modal domain to the space domain. Using the orthogonality of the eigenfunctions and a proper normalization, the related forward transform, i.e. the modal decomposition, can thus be written as

$$\begin{aligned} \tilde{\psi}(\boldsymbol{\kappa}) &= \langle \psi, \phi_{\lambda(\boldsymbol{\kappa})} \rangle \\ &= \int_{\mathbb{R}^2} \psi(\boldsymbol{x}_L) \phi^*_{\lambda(\boldsymbol{\kappa})}(\boldsymbol{x}_L)\, \mathrm{d}^2 \boldsymbol{x}_L, \end{aligned} \tag{3.55}$$

and

$$\begin{aligned} \tilde{\psi}(\lambda_i) &= \langle \psi, \phi_{\lambda_i} \rangle \\ &= \int_{\mathbb{R}^2} \psi(\boldsymbol{x}_L) \phi^*_{\lambda_i}(\boldsymbol{x}_L)\, \mathrm{d}^2 \boldsymbol{x}_L. \end{aligned} \tag{3.56}$$

Example 3.7 *As an example, consider the laterally invariant case. In this case, the discrete spectrum disappears. It is easily seen then, that the following complex exponential functions are the eigenfunctions of* $\mathcal{H}_2$*:*

$$\phi_{k_0^2 - \kappa_\mu \kappa_\mu}(\boldsymbol{x}_L) = \frac{1}{2\pi} \exp\{-j\, \kappa_\mu x_\mu\}. \tag{3.57}$$

In the right-hand side of equation (3.57) *a plane wave can be recognized for each* $\boldsymbol{\kappa}$*. The parameters* $\boldsymbol{\kappa}$ *are recognized as the horizontal wavenumbers, and* λ *is recognized as the square of the vertical wavenumber. For the laterally invariant case, an infinite number of plane waves have the same eigenvalue. This property corresponds to an infinite degeneracy. Consequently, at a single depth level an observer cannot pronounce upon the propagation direction of two plane waves for which the norms of* $\boldsymbol{\kappa}$ *are equal. Since* $\mathcal{H}_2$ *is a real-valued self-adjoint operator having a real spectrum, we can alternatively choose the eigenfunctions given in* (3.57) *to be real-valued:*

$$\phi_{k_0^2 - \kappa_\mu \kappa_\mu}(\boldsymbol{x}_L) = \frac{1}{\pi\sqrt{2}} \cos\{\kappa_\mu x_\mu - \pi/4\}. \tag{3.58}$$

Note that in this equation, the $\pi/4$ phase-shift is essential for the construction of both odd- and even functions $\psi(\boldsymbol{x}_L)$. Substitution of (3.57) in (3.55) yields the 2-D spatial Fourier transformation, and substitution of (3.58) in (3.55) yields the 2-D spatial Hartley transformation (Bracewell, 1986).

In the domain which is constituted by its eigenfunctions, $\mathcal{H}_2$ becomes a multiplication operator. Therefore, according to (3.41) we may write

$$(\mathcal{H}_2\psi)(\boldsymbol{x}_L) = \int_{\mathbb{R}^2} \tilde{\psi}(\boldsymbol{\kappa})\,\lambda(\boldsymbol{\kappa})\,\phi_{\lambda(\boldsymbol{\kappa})}(\boldsymbol{x}_L)\,\mathrm{d}^2\boldsymbol{\kappa} + \sum_{\lambda_i\in\sigma_{\mathrm{disc}}} \tilde{\psi}(\lambda_i)\,\lambda_i\,\phi_{\lambda_i}(\boldsymbol{x}_L). \tag{3.59}$$

On account of equations (3.13) and (3.44) the kernel $\mathrm{H}_2(\boldsymbol{x}_L, \boldsymbol{x}'_L)$ can be written as

$$\mathrm{H}_2(\boldsymbol{x}_L, \boldsymbol{x}'_L) = \int_{\mathbb{R}^2} \phi_{\lambda(\boldsymbol{\kappa})}(\boldsymbol{x}_L)\,\lambda(\boldsymbol{\kappa})\,\phi^*_{\lambda(\boldsymbol{\kappa})}(\boldsymbol{x}'_L)\,\mathrm{d}^2\boldsymbol{\kappa} + \sum_{\lambda_i\in\sigma_{\mathrm{disc}}(\mathcal{H}_2)} \phi_{\lambda_i}(\boldsymbol{x}_L)\,\lambda_i\,\phi^*_{\lambda_i}(\boldsymbol{x}'_L), \tag{3.60}$$

which has to be interpreted in the sense of generalized functions.

Expanding the one-way propagator —Using equations (3.18) and (3.46), it is easily seen that the kernel of the square-root operator can be written as

$$\mathrm{H}_1(\boldsymbol{x}_L, \boldsymbol{x}'_L) = \int_{\mathbb{R}^2} \phi_{\lambda(\boldsymbol{\kappa})}(\boldsymbol{x}_L)\,\lambda^{1/2}(\boldsymbol{\kappa})\,\phi^*_{\lambda(\boldsymbol{\kappa})}(\boldsymbol{x}'_L)\,\mathrm{d}^2\boldsymbol{\kappa} + \sum_{\lambda_i\in\sigma_{\mathrm{disc}}} \phi_{\lambda_i}(\boldsymbol{x}_L)\,\lambda_i^{1/2}\,\phi^*_{\lambda_i}(\boldsymbol{x}'_L), \tag{3.61}$$

where for later convenience, the signs of the square-root are chosen according to

$$\mathrm{Re}(\lambda^{\frac{1}{2}}) \geq 0 \quad \text{for } \lambda \geq 0, \tag{3.62}$$

$$\mathrm{Im}(\lambda^{\frac{1}{2}}) < 0 \quad \text{for } \lambda < 0. \tag{3.63}$$

Figure 3.4 shows schematically the spectrum of the square-root operator.

Assume now that the constitutive parameters are independent of the direction of preference between x'_3 and x_3. In that case, the kernel of the propagator as defined in equation (3.25) can be derived in a similar way, on account of equation (3.46),

$$\mathrm{W}^{\pm}(\boldsymbol{x}_L, x_3; \boldsymbol{x}'_L, x'_3) = \int_{\mathbb{R}^2} \phi_{\lambda(\boldsymbol{\kappa})}(\boldsymbol{x}_L)\,\exp\{\mp j(x_3 - x'_3)\,\lambda^{1/2}(\boldsymbol{\kappa})\}\,\phi^*_{\lambda(\boldsymbol{\kappa})}(\boldsymbol{x}'_L)\,\mathrm{d}^2\boldsymbol{\kappa} + \sum_{\lambda_i\subset\sigma_{\mathrm{disc}}} \phi_{\lambda_i}(\boldsymbol{x}_L)\,\exp\{\mp j(x_3 - x'_3)\,\lambda_i^{1/2}\}\,\phi^*_{\lambda_i}(\boldsymbol{x}'_L). \tag{3.64}$$

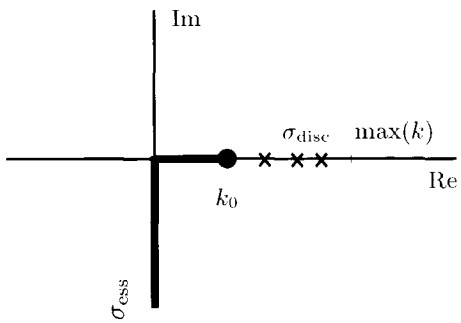


Fig. 3.4 *The spectrum of the square-root operator.*

Equation (3.64) can be interpreted as a transformation from the modal domain to the spatial domain representation. The modal representation of the propagator $\mathcal{W}^{\pm}(x_3; x_3')$ reads

$$\tilde{\mathrm{W}}^{\pm}(\boldsymbol{\kappa}, x_3; \boldsymbol{\kappa}, x_3') = \exp\{\mp j(x_3 - x_3')\lambda^{1/2}(\boldsymbol{\kappa})\}, \tag{3.65}$$

where λ can be in the essential part of the spectrum or in the discrete part of the spectrum.

For the laterally invariant situation, the eigenfunctions of the Helmholtz operator become plane waves (see Example 3.7) and the discrete spectrum vanishes. Since the parameters $\boldsymbol{\kappa}$ can be interpreted as horizontal wavenumbers in this situation, the square-root of the eigenvalues λ can be recognized as the vertical wavenumber. The vertical wavenumber gives the vertical phase velocity of the plane wave via the relation $\lambda^{1/2} = \omega/c_3$, where c_3 is the phase velocity in the direction of preference. The representation $\tilde{\mathrm{W}}$ reduces to the well-known phase-shift operator (Stoit, 1978; Gazdag, 1978) describing the propagation of plane waves from x_3' to x_3. For $\kappa_\mu \kappa_\mu$ larger than k_0^2 the plane waves become vertically evanescent[26]. The propagating plane waves propagate changelessly through the medium, except for a phase change.

For the laterally variant situation, the square-root of the eigenvalue can be interpreted as a generalized vertical wavenumber. Hence, $\lambda^{1/2}$ gives the effective vertical phase velocity of the eigenfunction, via the relation $\lambda^{1/2} = \omega/c_3$. The propagation of the eigenfunctions from depth level x_3' to x_3 is described by the generalized phase shift operator $\tilde{\mathrm{W}}$. For $\lambda > 0$, propagating eigenfunctions are considered. The propagating eigenfunctions propagate changelessly through the medium, except for a phase change. For $\lambda < 0$, the eigenfunctions become vertically evanescent, just as in the laterally invariant situation.

[26]Note, however, that "evanescent waves" propagate perpendicular to the direction of preference.

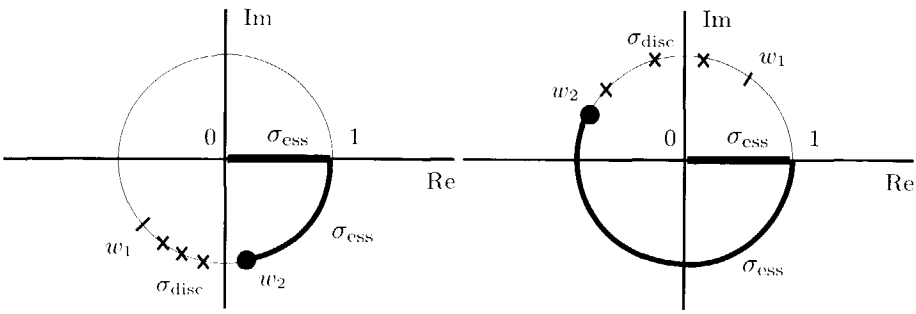


Fig. 3.5 *The spectrum of the propagator $\mathcal{W}^+$ for two depth steps $(x_3 - x_3')$. In this figure $w_1 = \exp\{-j \max(k)\,(x_3 - x_3')\}$ and $w_2 = \exp\{-jk_0(x_3 - x_3')\}$*

Figure 3.5 shows schematically the spectrum of the propagator for two depth steps $(x_3 - x_3')$. Elements of the spectrum corresponding to non-evanescent wave modes lie on the unit circle. The vertically evanescent waves lie on the real axis between 0 and 1.

3.4 Examples

The theory developed in the previous two sections requires some examples to be fully appreciated. For that purpose the numerical equivalent of the theory has been implemented. The numerical examples apply to the Helmholtz operator, but are equally valid (with minor modifications) for the wave operator in (wave) time evolution problems. The examples are primarily chosen

1. to illustrate the physical meaning and usefulness of (a) the eigenfunctions and (b) the distinction between different kind of eigenfunctions related to different parts of the spectra (see section 3.4.1),

2. to illustrate the accuracy of a modal decomposition if applied in migration (see section 3.4.2) .

The numerical computations are carried out in a 2-D configuration. For details concerning the implementation of the second-order derivative and the numerical eigenvalue decomposition, the reader is referred to Golub and van Loan (1984), Grimbergen (1995) or Grimbergen et al. (1998).

	c_0 [m/s]	k_0 [rad/m]	k_0^2 [rad^2/m^2]	max(k) [rad/m]	max(k^2) [rad^2/m^2]
(1)	2000	0.094	0.009	0.094	0.009
(2)	2000	0.094	0.009	0.166	0.027
(3)	2000	0.094	0.009	0.094	0.009
(4)	1633	0.115	0.013	0.232	0.054

Table 3.2 *Some relevant quantities for the profiles (1) to (4) shown in Figure 3.6. The angular frequency is taken $\omega = 2\pi \times 30$ rad/s.*

3.4.1 Spectrum and eigenfunctions

In order to illustrate the use of the theoretical considerations concerning the spectrum, the eigenvalue decomposition is applied to a number of medium configurations. The attention will be focussed on the spectrum of the Helmholtz operator, the associated spectra of functions of the Helmholtz operator, being the square-root operator and the propagator, and the associated eigenfunctions.

Consider a 2-D medium configuration. In a 2-D medium the Helmholtz operator of equation (3.12) takes the form

$$\mathcal{H}_2 = k^2(x_1, x_3) + \partial_1^2, \tag{3.66}$$

with $k^2(x_1, x_3) = \omega^2/c^2(x_1, x_3)$. The density is chosen to be constant. In this section the frequency is chosen to be $\omega = 2\pi f = 2\pi \times 30\,\text{rad/s}$. The essential spectrum is degenerate with a multiplicity of two, instead of an infinite multiplicity for the 3-D medium. A left and right propagating wave mode corresponds to one eigenvalue. The medium profiles that will be studied in this section are taken at a certain fixed depth level x_3, and they are sampled with a sampling distance $\Delta x_1 = 10\,\text{m}$. The profiles are shown in Figure 3.6. The left column shows the lateral velocity profiles $c(x_1, x_3)$ and the right column shows the squared wavenumber functions $k^2(x_1, x_3)$ arising in the Helmholtz operator. The medium profiles are numbered from top to bottom (1) to (4). The profiles are discussed in the caption. Recalling the structure of the spectrum of the Helmholtz operator as discussed in section 3.3.3 and as schematically shown in Figure 3.3, one can conclude that the variables k_0^2 and max(k^2) are the ruling quantities of the discrete and essential spectrum. For the four profiles of Figure 3.6 these quantities are shown in Table 3.2 in addition to k_0 and max(k).

Profile (1) —Figure 3.7a shows the spectrum of the Helmholtz operator for the constant velocity profile. It is a degenerate spectrum with a multiplicity of two, except for the first and last eigenvalue. The spectrum of the square-root operator is shown in Figure 3.7b. If it is assumed that the medium properties do not

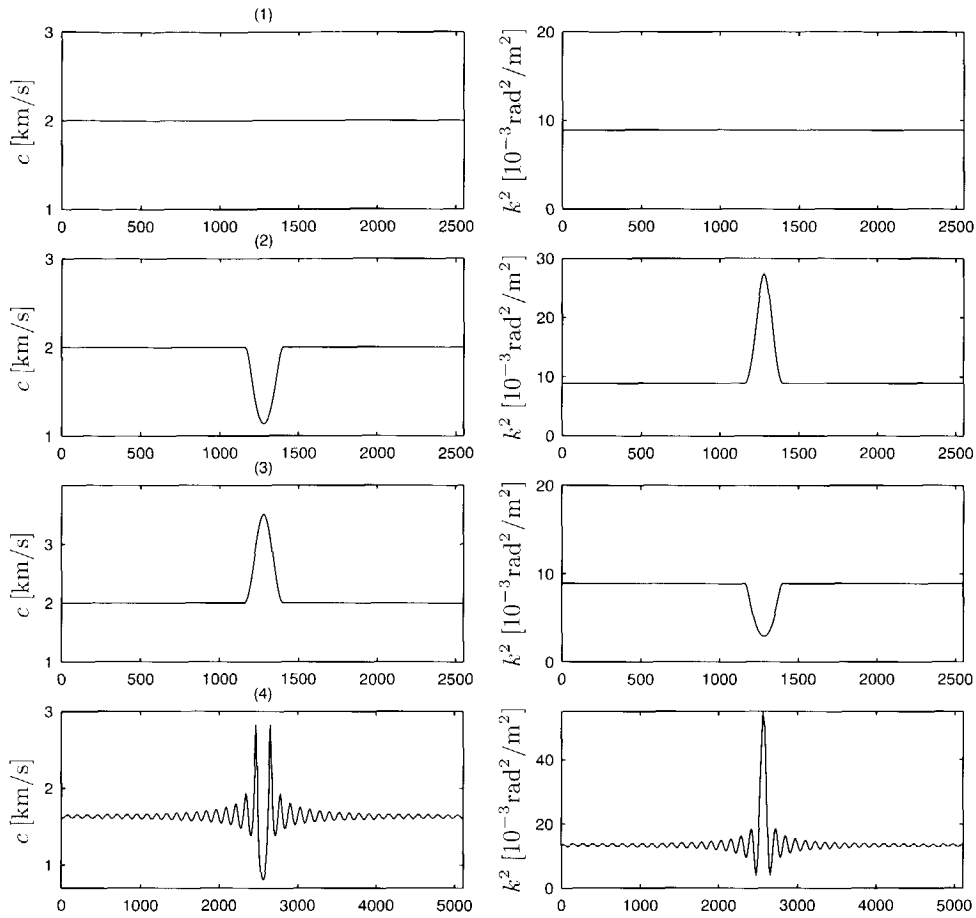


Fig. 3.6 *Velocity profiles $c(x_1, x_3)$ (left column) and the squared wavenumber profiles $k^2(x_1, x_3)$, which is the potential in the Helmholtz operator (right column). The profiles are numbered (1) to (4). Profile (1) is a constant velocity medium. Profile (2) has a local low velocity region which acts as a wave guide. Profile (3) has a local high velocity region, which acts as a anti-wave guide. Profile (4) is a slowly decaying oscillatory velocity profile.*

change within a small depth range around x_3, it makes sense to compute the spectrum of the propagator. For two different extrapolation steps $(x_3 - x_3') = 10\,\mathrm{m}$ and $(x_3 - x_3') = 20\,\mathrm{m}$, the spectrum of $\mathcal{W}^+$ is shown in Figures 3.7c-d. As expected the discrete spectrum $\sigma_{\mathrm{disc}}(\mathcal{H}_2)$ is empty, and there are only radiating wave modes corresponding to the absolute continuous spectrum. The absolute continuous spectrum equals the essential spectrum in this case. The essential spectrum covers $(-\infty, k_0^2]$.

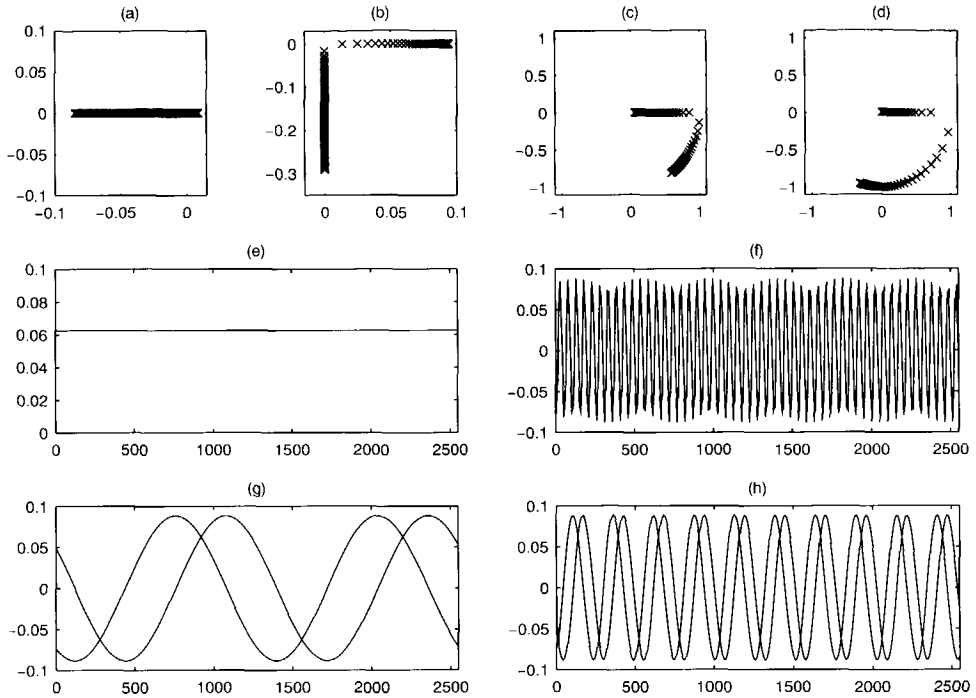


Fig. 3.7 *The spectrum of the Helmholtz operator $\mathcal{H}_2$, spectra of functions of the Helmholtz operator and a number of generalized eigenfunctions for Profile (1). (a) Spectrum of the Helmholtz operator. The discrete spectrum is empty. (b) Spectrum of the square-root operator obtained by applying $\lambda^{1/2}$ to the elements of the spectrum of $\mathcal{H}_2$. (c) and (d) Spectrum of the propagator $\mathcal{W}^{\dagger}$ for two different depth steps $(x_3 - x_3') = 10\,\mathrm{m}$ and $(x_3 - x_3') = 20\,\mathrm{m}$ obtained by applying $\exp\{-j(x_3 - x_3')\lambda^{1/2}\}$ to elements of the spectrum of $\mathcal{H}_2$. Note that propagating wave modes lie on the unit circle in the complex plane and evanescent wave modes on the real axis between 0 and 1. (e), (f), (g) and (h) A number of generalized eigenfunctions.*

The theoretical maximum is numerically confirmed. The limit $-\infty$ is in practical computations never reached. The lowest value reached depends on the sampling rate, in the same way as the Nyquist frequency depends on the sampling rate. Note that the continuous part of the spectrum is seemingly discrete. This is due to the limited aperture size. The other expectations are completely fulfilled, including the character of the eigenfunctions, according to Figures 3.7e-h. Figures 3.7e-h show a number of generalized eigenfunctions: (e) shows the eigenfunction corresponding to

the maximum eigenvalue; this is one of the two non-degenerate eigenfunctions. (f) shows the eigenfunction corresponding to the 100th eigenvalue. This is an evanescent wave mode. (g) and (h) show two pair of eigenfunctions corresponding to one eigenvalue each, illustrating the degeneracy of the spectrum. The two eigenfunctions have a phase difference of $\pi/2$.

Profile (2) —Figure 3.8a shows the spectrum of the Helmholtz operator corresponding to profile (2) in Figure 3.6. The spectrum of the square-root operator and the spectrum of the propagator for two different extrapolation steps are shown in Figures 3.8b-d. Looking at the values in Table 3.2 and remembering that for a sufficiently fast decaying perturbation a finite number of bound states are expected in the region $(k_0^2, \max(k^2)]$, the reader can verify the theory by observing the spectrum in Figure 3.8a or, equivalently, Figures 3.8b-d. Figures 3.8e-f show two radiating modes corresponding to elements in the essential spectrum. Figures 3.8g-l show the first six of seven guided wave modes for this configuration. The guided wave modes can clearly only exist within the perturbation. Outside the perturbation they are evanescent in the lateral direction. The single multiplicity of the guided modes and the number of zero crossings is forecasted by Theorem 55 of Dunford and Schwartz (1963). Note that the number of guided wave modes and the structure of the guided wave modes only depend on the structure of the disturbance. They do not depend on sampling issues or on the size of the medium around the perturbation. They do, however, depend on the frequency.

Profile (3) —This medium profile is opposite to the medium configuration of profile (2). It is shown in Figure 3.6. It is an anti-wave-guide structure, which means that all energy will be radiated out of the perturbation. Figure 3.9a shows the spectrum of the Helmholtz operator. Figures 3.9b-d show the spectra of functions of the Helmholtz operator. The essential spectrum is undisturbed by the perturbation and there is not a discrete spectrum, i.e. $\sigma_{\text{ess}}(\mathcal{H}_2) = \sigma(\mathcal{H}_2) = (-\infty, k_0^2]$. Consequently, only radiating wave modes are expected. The expectation is indeed fulfilled (see Figures 3.9e-h). The first part of the spectrum consists of radiating wave modes which are laterally evanescent, i.e. exponentially decaying, in the region of the perturbation.

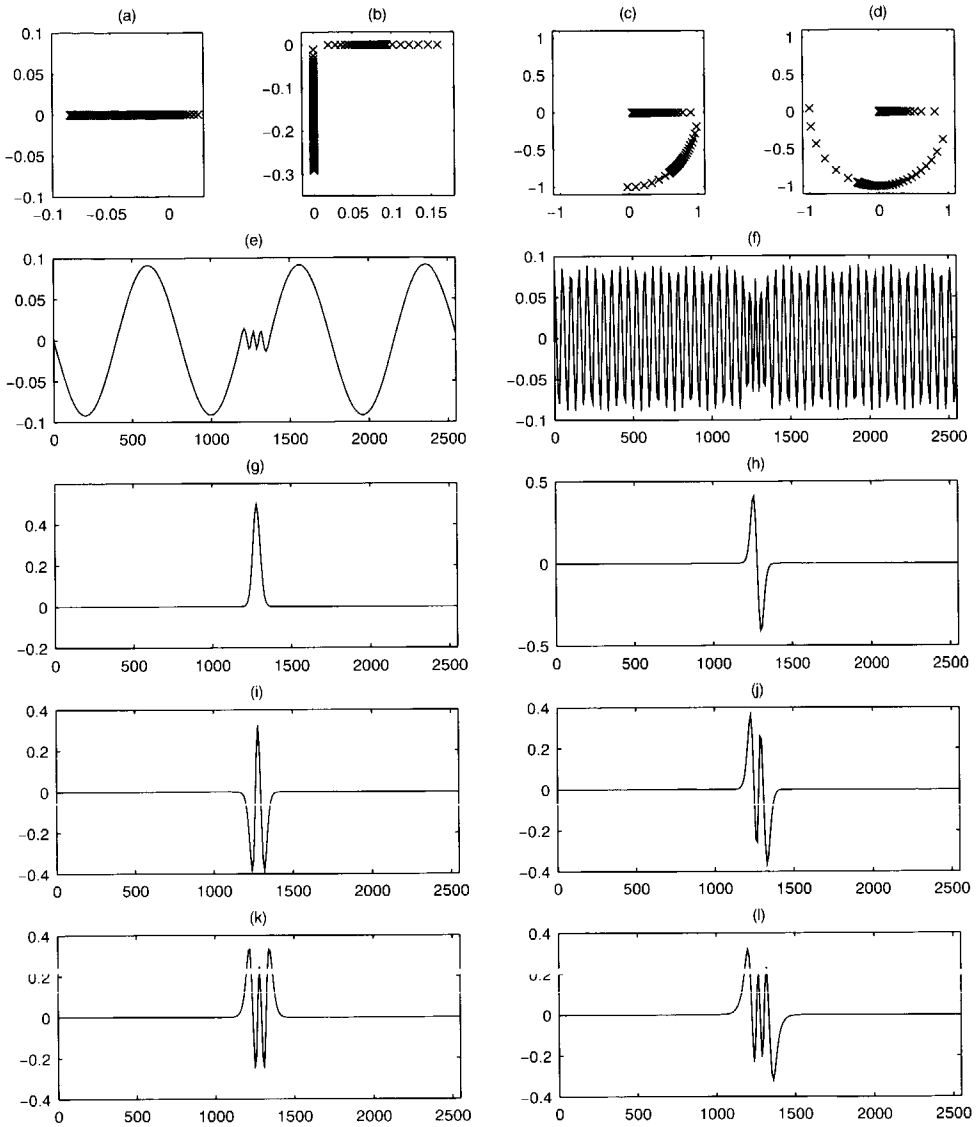


Fig. 3.8 *Spectrum of the Helmholtz operator $\mathcal{H}_2$, spectra of functions of the Helmholtz operator and a number of (generalized) eigenfunctions corresponding to Profile (2). (a)-(d) See caption of Figure 3.7 for a general description. In addition to the essential spectrum, which coincides with the spectrum of Profile (1), there are a number of eigenvalues corresponding to guided wave modes present. (e)-(f) Two radiating wave modes. (g)-(l) The first six of seven guided wave modes. The number and the structure of the guided wave modes depend on the frequency, and on the structure of the disturbance.*

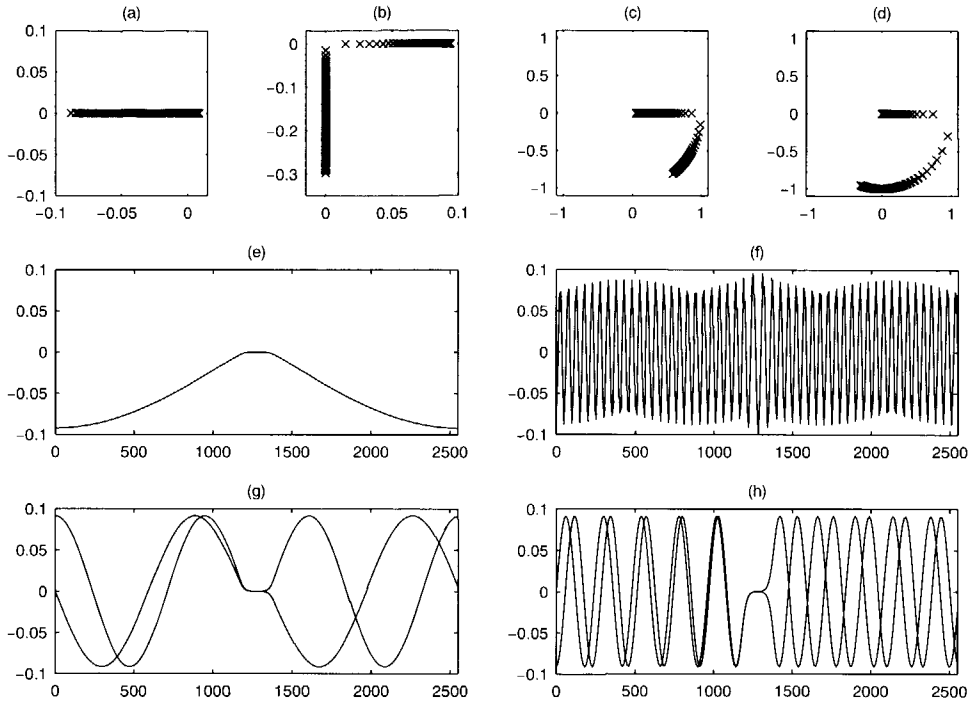


Fig. 3.9 *Spectrum of the Helmholtz operator $\mathcal{H}_2$, spectra of functions of the Helmholtz operator and a number of generalized eigenfunctions corresponding to Profile (3). (a)-(d) See caption of Figure 3.7. The essential spectrum is not changed by the disturbance, and there is not a discrete spectrum. The eigenfunctions however are disturbed as is clear in (e)-(h). The eigenvalues corresponding to the eigenfunctions shown here, are the same as the eigenvalues of the eigenfunctions shown in the homogeneous example in Figure 3.7.*

Profile (4) -It has been mentioned a number of times in section 3.3 that the discrete spectrum is not necessarily equal to the pure point spectrum. There may even be elements of the point spectrum that lie within the essential spectrum, or within the absolute continuous spectrum if the singular continuous spectrum is excluded. These strange wave constituents are guided modes with an infinite support. "Guided modes with infinite support" seems to be a contradiction in terms, but as this example shows, these objects can exist. The function $k^2(x_1, x_3)$ has been constructed according to

$$k^2(x_1, x_3) = k_0^2 + k_0^2 C \frac{\sin(2x_1/(4\Delta x_1))}{x_1/\Delta x_1} \quad \text{with} \quad \Delta x_1 = 10\,\text{m}$$

where C is an appropriately chosen constant. The expression for $k^2(x_1, x_3)$ is a free adjustment of the potential function proposed for the Hamiltonian by Reed and Simon (1978, p. 223). In Figure 3.6 the velocity profile and $k^2(x_1, x_3)$ are plotted. Figure 3.10a shows the spectrum of the Helmholtz operator. The transition from guided wave modes to other wave modes is not very clear. Such a transition corresponds to a clear cut difference between the discrete and the essential spectrum, which is absent for this profile. This observation is not that strange if one realizes that the decay properties of the potential given by the equation above is not fast enough for the theory described on page 92 and further, to be valid. There are both guided wave modes (Figures 3.10e-h), quasi-guided modes (Figure 3.10i), and radiating wave modes (Figure 3.10j and l). The strange object which was hinted at above is observed in Figure 3.10k. This object corresponds to an eigenvalue $\lambda = 0.0102\,\text{rad}^2\text{m}^{-2}$ which lies within the essential spectrum. It has a low amplitude radiating structure and a high amplitude guided structure. The eigenfunctions belonging to the eigenvalues just neighboring this eigenvalue are again radiating modes. Figure 3.10l displays one of the two. This example illustrates that strange things can happen. A computation of the guided wave modes on the basis of eigenvalues in the discrete part of the spectrum will certainly lead to an underestimation of the amount of trapped energy.

3.4.2 Wave field extrapolation and migration

The previous section showed how the modal decomposition can be used to unravel the complexity of the lateral medium structure. Another property of the modal decomposition is the fact that the kernel of the propagator can be computed accurately (see equation 3.64). The propagator plays a key role in the data representation of (3.28). In this section it will be illustrated that the computation of the propagator kernel with the help of a modal decomposition yields an accurate migration scheme. The algorithms are compared with schemes in which the propagator kernel is computed with the local explicit method (Holberg, 1988; Blacquière et al., 1989; Thorbecke, 1997). For an example of an accurate well-to-well wave field extrapola-

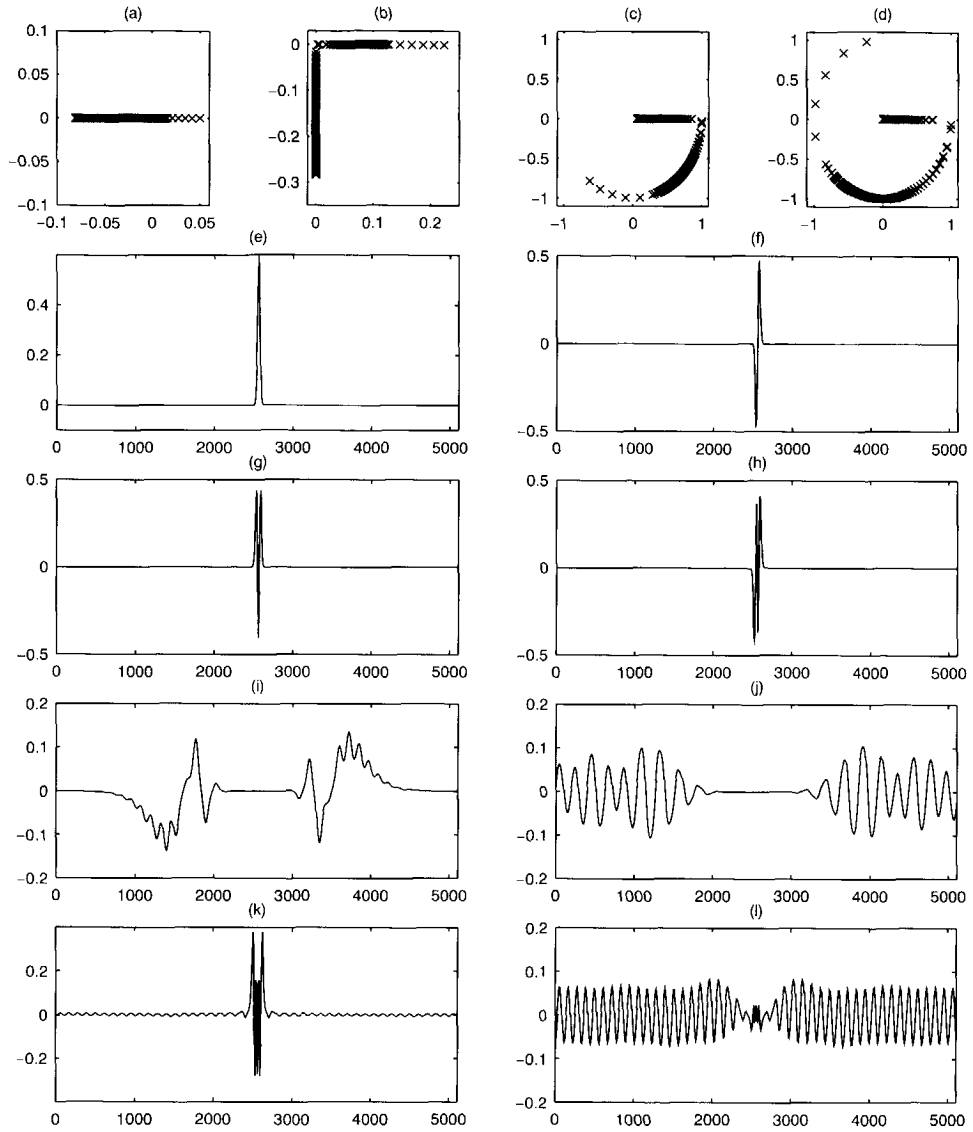


Fig. 3.10 *Spectrum of the Helmholtz operator $\mathcal{H}_2$ and spectra of functions of the Helmholtz operator and a number of (generalized) eigenfunctions corresponding to Profile (4). (a)-(d) See caption of Figure 3.7. (e)-(h) Some guided wave modes. (i) An intermediate wave mode. (j)-(l) Three generalized eigenfunctions corresponding to eigenvalues in the essential part of the spectrum. Note that the eigenfunction shown in (k) has a clear guided structure, despite the fact that its eigenvalue lies in the essential part of the spectrum.*

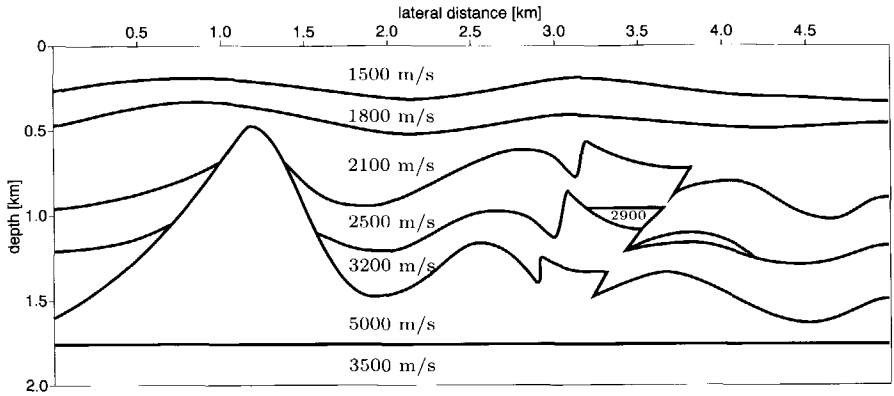


Fig. 3.11 *The velocity model for the migration. Velocities are indicated in the corresponding layers.*

tion the reader is referred to Grimbergen et al. (1998).

Migration example —A set of synthetic shot records have been migrated[27]. The shot records have been generated by finite-difference modeling using the subsurface model in Figure 3.11. In this case both lateral (i.e., horizontal) variations and depth variations are present in the model. The model contains a high velocity layer (salt), piercing through a number of layers. The salt dome is a high velocity structure acting as an anti-wave guide. To the right of this structure, the block-shaped elevation implies yet another lateral discontinuity. The acquisition parameters are summarized in Table 3.3. Because we are now dealing with depth variations, the wave field is extrapolated in small steps (i.e. recursive extrapolation), using the complex conjugate transposed of the representations of the propagators W^+. The medium within each step is assumed to be depth invariant.

The migration has been carried out with two spatial representations of the propagator. One is computed via the modal domain according to equation (3.64). The other is computed with the local explicit operators. The local explicit operator method assumes that the medium parameters are constant within the size of the operator. A local explicit operator has two important free parameters: its length and the optimization angle. The optimization angle is the angle up to which the extrapolation has to fulfill a certain accuracy.

The stacked result of the separate shot record migrations using modal expansion is shown in Figure 3.12. The flanks of the salt as well as the faults are clearly

[27] This example has been first presented in a Technical Report by Dessing and Grimbergen (1996). See also Grimbergen et al. (1998).

geometry	fixed spread
number of shots	11
shot spacing	500 m
number of detectors per shot	251
receiver spacing	20 m
recording time	3 s
time sampling	4 ms
frequency content wavelet	up to 35 Hz

Table 3.3 *The acquisition parameters for the migration example*

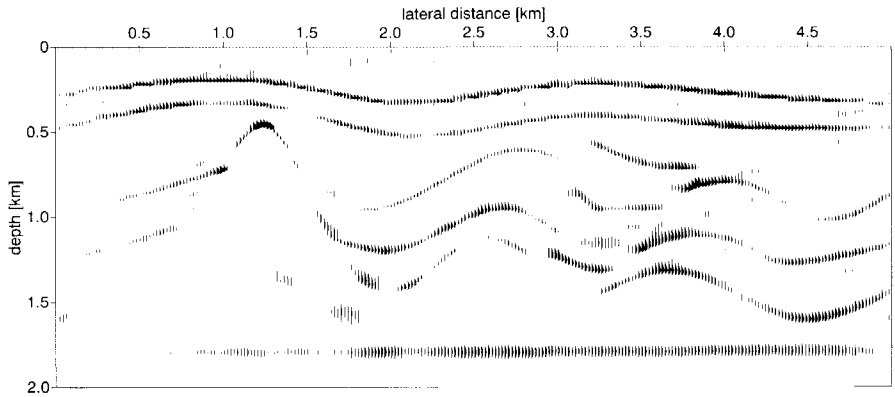


Fig. 3.12 *Migrated section using modal expansion extrapolation operators.*

imaged due to the absence of dip limitation. Note the overall crisp character of the result. An unambiguous comparison with the available local explicit method is not straightforward because both the length and the optimization angle can be varied. Choosing these parameters may lead to conflicting requirements. In case of strong lateral variations, short operators are needed in order to avoid instabilities. Imaging of steep dips however, asks for longer operators that allow for higher optimization angles. Figure 3.13 shows the results for several choices of these parameters. Note that a higher optimization angle improves the imaging of steep dips but causes stronger artifacts at the same time. These artifacts are caused by the increased spatial length of the operator. Note that all results in Figure 3.13 are inferior to the modal expansion result in Figure 3.12.

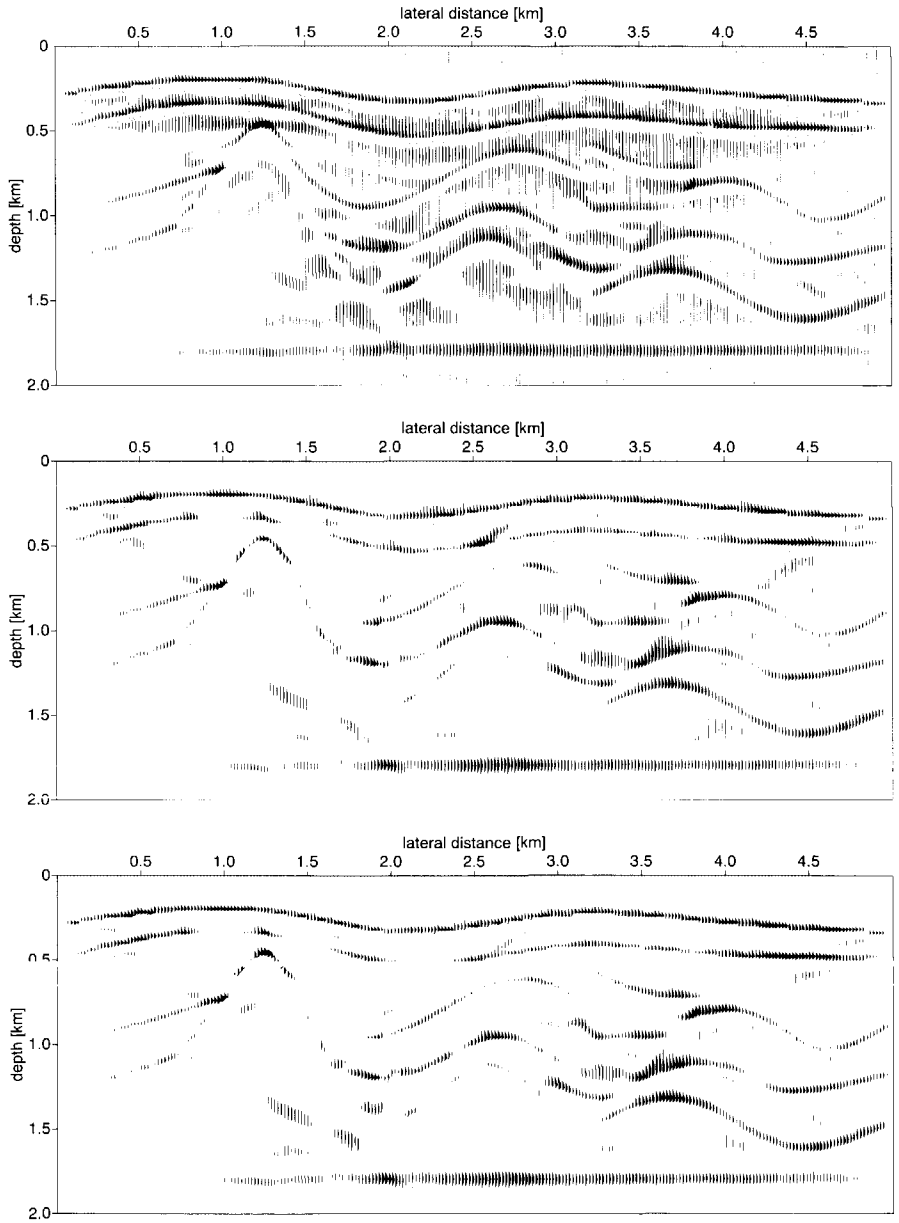


Fig. 3.13 *Migrated sections using local explicit operators. Operator length 27, optimization up to 60 degrees (top); operator length 27, optimization up to 80 degrees (middle) and operator length 37, optimization up to 80 degrees (bottom). For details with respect to the local explicit operators the reader is referred to Holberg (1988), or Thorbecke (1997).*

3.4.3 Discussion

Analysis of spectrum and eigenfunctions —In relation to the examples in section 3.4.1, one might be tempted to say that the profiles discussed so far are rather hypothetical. Profiles with exactly the same functional form as discussed will indeed never be met in the real world. However, similar structures can locally occur in the earth, for example in a well-to-well configuration, where the direction of preference is taken along a horizontal axis (Grimbergen et al., 1998). Since in well-to-well configurations the major changes are perpendicular to the direction of preference, guided and radiating wave modes are equally probable. Hence, for those configurations it is even more important to realize which quantities rule the wave field, and which are relatively unimportant, but also where the energy is expected to be trapped, etc. Finally, it is noted that a similar analysis is, mutatis mutandis, also applicable in time evolution problems in a layered earth (Wilcox, 1984; Faris, 1995; Herrmann, 1997) for which the guided wave fields with a radiating structure should be expected.

Migration —It has been shown that the proposed method to calculate one-way operators has desirable properties such as the absence of a dip limitation, the accurate handling of lateral variations and the unconditional stability of the operators. The obvious drawback of the method is the computational cost of a full eigenvalue decomposition, which is considerable compared to the construction of the local explicit operators. However, the following considerations may help to overcome this problem:

- The Helmholtz matrix operator is a sparse symmetric band matrix. For a full symmetric $M \times M$ matrix, the number of floating point operations necessary to calculate all eigenvalues and all eigenvectors will increase with the third power of M. However, in case of a matrix operator with a fixed number of non-zero diagonals independent of M, the number of floating point operations will increase only with the square of M (Golub and van Loan, 1984).

- Not all eigenvalues need to be calculated. Calculating only the positive eigenvalues (propagating modes) still leads to accurate results, as the evanescent part of the wave field decays exponentially with the extrapolation distance (in inverse extrapolation, the evanescent field is suppressed anyway to obtain stable operators). This argument holds in particular for low temporal frequencies, where a large number of the eigenvalues are associated to evanescent wave modes.

The modal expansion method also provides an interesting scope for turning wave migration. In some references (Claerbout, 1985; Hale et al., 1992), the phase-shift method is applied because it has no dip limitation, which is an essential requirement

for turning wave migration. However, the phase-shift method is applicable only in laterally invariant media (which was acknowledged by the authors). The modal expansion method combines both the ability to deal with lateral variations and the ability to handle dips up to 90 degrees.

The implementation of the modal decomposition in a 3-D medium configuration is straightforward (Grimbergen et al., 1998). The advantages as discussed before are equally true in a 3-D configuration. Efficient algorithms as discussed in the previous paragraph are of vital importance if one wants to compute the 2-D eigenfunctions in a 3-D medium.

3.5 Summary

In this chapter the linearized partial differential equations, forming a hyperbolic system and governing the acoustic wave field, have been introduced. Via the two-way wave equation the one-way wave equation has been arrived at. The one-way wave equation comprises a separation between downgoing and upgoing waves. In addition, the one-way decomposition separates propagation along the direction of preference and scattering along the direction of preference. A subsequent use of the Green's function concept and the first term of a Bremmer coupling series yield the monochromatic one-way representation of primary seismic reflection data. The one-way representation is a model for the seismic reflection data in terms of a downward propagator, a reflection operator, and an upward propagator. Both the propagators and the reflection operator can be written as a function of the Helmholtz operator.

The second part of this chapter has been devoted to the interpretation of the Helmholtz operator and functions of the Helmholtz operator. To this end I have made use of functional calculus. The first step here is the derivation of a nice self-adjoint form of the Helmholtz operator. As a result of the self-adjoint form, the spectrum of the Helmholtz operator is a subset of the real axis and the eigenfunctions form a complete set. The spectrum of the Helmholtz operator can be subdivided either in a pure point spectrum, an absolute continuous spectrum and a singular continuous spectrum, or in a essential spectrum and a discrete spectrum. The first subdivision provides direct insight in the type of wave fields that can occur and the second subdivision provides clear spectral bounds. The two subdivisions are related. Besides the physical insight via an analysis of the spectral properties, the complete set of eigenfunctions enables an easy and accurate calculation of functions of the Helmholtz operator, such as the propagator.

In the third part of this chapter the examples illustrate that the spectral analysis provides physical insight on the one hand and the possibility to carry out an unconditionally stable and accurate migration scheme on the other hand.

Chapter 4

Generalized data representations and generalized migration

In the present chapter the opportunities of transformation techniques for the representation of seismic data and a subsequent migration are discussed. The emphasis is put on a wavelet domain approach to migration. The results of this chapter has been dealt with in a number of conference proceedings, most notably Dessing and Wapenaar (1994, 1995).

4.1 Introduction

Seismic migration is the most time consuming step in a seismic processing sequence. The geophysical community is continuously looking for methods and tools to improve the efficiency and quality of this step. New and promising tools are in general wholeheartedly embraced. One of those promiseful tools is the wavelet transform. The question is what the wavelet transform can do for migration?

In order to be able to answer this question one has to know what the new tool comprises. What are its properties and how is it positioned with respect to other tools? These issues have been dealt with in chapter 2, where I introduced the spatial domain representation, the Fourier domain representation, the Gabor domain representation, and the wavelet domain representation, their properties and their mutual relationships. One does not only have to know the ins and outs of the new tool, but one has to know to which system the tool has to be applied as well. I have chosen to take the one-way representation of primary seismic reflection data as a point of departure for wavelet transform based seismic migration. The one-way data representation has been introduced in chapter 3, it is given by (cf. equation 3.28)

$$P^- = \mathcal{W}^- \mathcal{R} \mathcal{W}^+ S^+. \tag{4.1}$$

It is a model in terms of propagators $\mathcal{W}^{\pm}$, and a reflection operator $\mathcal{R}$. S^{+} denotes the one-way downgoing source distribution and P^{-} represents the seismic reflection data at the surface after decomposition and surface-related multiple elimination. The conceptual physical model of equation (4.1) clearly discriminates between propagation and reflection with the help of an operator notation. Besides, the operator notation makes it trivial to derive representations in an arbitrary domain.

A good overview of the use of wavelets for representing operators is given by Beylkin (1996). He discriminates between their use for integral equations and their use for partial differential equations. An integral equation generally leads after discretization to a dense system of linear algebraic equations with a small condition number. On the other hand, a partial differential equation leads via finite-difference or finite-element methods to a sparse linear system. The cost of this sparsity is, as Beylkin (1996) points out, a large condition number. Wavelet bases can be used for both types of equations. The integral operators associated with integral equations can be rendered sparsely in wavelet bases. For the finite-difference or finite-element methods the wavelet bases provide diagonal preconditioners. Since the system of equation (4.1) contains a number of integral operators, I will focus on results related to integral operators.

The integral operators that are successfully dealt with by wavelet transform methods are associated with elliptic problems. These operators are characterized by the fact that they are smooth away from the diagonal. A well-known example is the operator of the Calderón-Zygmund type. Such an operator is given by an integral kernel $\mathrm{K}(x,y)$ that is indeed smooth away from the diagonal, according to

$$|\mathrm{K}(x,y)| \leq \frac{1}{|x-y|}$$
$$|\partial_x^M \mathrm{K}(x,y)| + |\partial_y^M \mathrm{K}(x,y)| \leq \frac{C_0}{|x-y|^{1+M}},$$

for $M \geq 1$. Sparse representations of Calderón-Zygmund operators in wavelet bases and efficient algorithms for these operators in the wavelet domain have been discussed for the first time in a consistent way by Beylkin et al. (1991). They showed how $O(N^2)$-algorithms can be reduced to $O(N)$-algorithms by utilizing wavelet bases.

The propagators $\mathcal{W}^{\pm}$ are integral operators associated with a hyperbolic problem. They are oscillatory. Hence, they do not belong to the class of operators that are smooth away from the diagonal. Dramatic savings as with integral operators of elliptic problems can not be expected. Results of the use of wavelets bases in electromagnetic and acoustic integral equations have been reported on by several authors (Steinberg and Leviatan, 1993; Goswami et al., 1995; Wagner and Chew, 1995; Kim et al., 1996). These authors discuss first-kind integral equations of the

form (cf. equations 2.1 and 3.1)

$$\int_a^b \mathrm{A}(x,x')u(x')\,\mathrm{d}x' = f(x),$$

where $\mathrm{A}(x,x')$ is a known kernel of an integral equation, $f(x)$ is a known source distribution and $u(x)$ is the unknown field or response. For the kernel, the referred authors take the 2-D monochromatic Green's function associated with electromagnetic and acoustic problems in a homogeneous medium, and project it on a set of wavelets. In this way they generalize the method of moments.

In the present chapter, the actual problem is to derive from equation (4.1) an explicit expression for the reflectivity operator $\mathcal{R}$ in terms of the data, which is of course a different problem than the determination of the unknown field $u(x)$. Since the wavelet domain does not provide dramatic sparseness for the propagators, I have concentrated on efficiently subdividing the reflectivity operator in approximations and details, and on getting explicit expressions for the different parts in terms of the measured data. This method resembles the homogenization approach by Brewster and Beylkin (1995), and the multiscale scattering approach of Steinberg (1994) and Steinberg and McCoy (1994).

The contents of the present chapter is as follows. In section 4.2 I will commence the discussion with a representation of primary seismic reflection data in general domains. In the subsections the discussion will be specified to the data representation in the space domain (section 4.2.1), the spatial Fourier domain (section 4.2.2), the Gabor domain (section 4.2.3), and the wavelet domain (section 4.2.4). The physical significance of the different representations and the relation with existing solution methods will be pointed out. In section 4.3 the properties of the propagators in the wavelet domain for a homogeneous medium will be discussed in more detail. This analysis helps in understanding how the wavelet transform decomposes a wave field. In section 4.4 a recursive depth migration scheme will be discussed, both in the space domain and in the wavelet domain. Section 4.5 is devoted to examples. In section 4.6 I will discuss a number of limitations and open questions of the presented method. Finally, in section 4.7 the findings of this chapter will be summarized.

4.2 Data representations revisited

Consider the monochromatic description of primary seismic reflection data related to the reflection at a single depth level x_3' with sources and receivers at one and the same depth level x_3 (cf. equation 3.28)

$$P^-(\boldsymbol{x}_L^{\mathrm{r}}, x_3) = \mathcal{W}^-(x_3; x_3')\,\mathcal{R}(x_3')\,\mathcal{W}^+(x_3'; x_3)\,S^+(x_3). \tag{4.2}$$

Remember that the operators $\mathcal{W}^-$, $\mathcal{R}$ and $\mathcal{W}^+$ are working in the horizontal plane. Hence, equation (4.2) is a full 3-D representation of the seismic reflection at a single

depth level. In this chapter, however, I will only consider 2-D configurations. Equation (4.2) is the point of departure for representations in the domains introduced in chapter 2: the spatial domain representation, the spatial Fourier domain representation, the Gabor domain representation and the wavelet domain representation. The latter will be most extensively dealt with. In general terms, the transformations that lead to the four representations, are medium independent. The representations themselves are not medium independent. In the previous chapter the modal decomposition has been introduced, for which both the transformation and the representation are medium dependent[1].

Before I treat the four domains separately, let me first consider the representation of the seismic reflection data in an arbitrary domain. Suppose $\{\psi_a\}$ forms a complete set, where the variable a is chosen in a suitable, possibly more dimensional, discrete or continuous set J. It can be concluded from chapter 2 that the kernel $\mathrm{K}(a;a')$ of the operator $\mathcal{K}$ can be obtained via

$$\tilde{\mathrm{K}}(a,a') = \langle \mathcal{K}\psi_{a'}, \psi_a \rangle \,, \tag{4.3}$$

and that the representation of a function f in the domain constituted by the set $\{\psi_a\}$ can be obtained via the inner product according to

$$\tilde{f}(a) = \langle f, \psi_a \rangle \,. \tag{4.4}$$

At this stage the symbol $\tilde{}$ refers to either of the four representations to be discussed in subsequent subsections. With equations (4.3) and (4.4) in mind, the data representation of equation (4.2) takes the following form in the domain constituted by the functions ψ_a

$$\tilde{P}^-(a^r,x_3) = \sum\!\!\!\!\!\!\int_{a\in J} \sum\!\!\!\!\!\!\int_{a'\in J} \sum\!\!\!\!\!\!\int_{a^s\in J} \tilde{\mathrm{W}}^-(a^r,x_3;a,x_3')\,\tilde{\mathrm{R}}(a,x_3';a',x_3')\,\tilde{\mathrm{W}}^+(a',x_3';a^s,x_3)\,\tilde{S}^+(a^s,x_3)\,\mathrm{d}a^s\,\mathrm{d}a'\,\mathrm{d}a. \tag{4.5}$$

The quantities in equation (4.5) can be interpreted in the following way. The function $\tilde{S}^+(a^s,x_3) = \langle S^+(x_3),\psi_{a^s}\rangle$ is the source distribution for downgoing waves[2]. The source distribution is expressed in terms of the basis functions ψ_{a^s}. The propagator kernel $\tilde{\mathrm{W}}^+(a',x_3';a^s,x_3)$ denotes the response of a downgoing wave field of the type

[1] It is interesting to observe that the approaches that lead to certain predefined wave fields in the subsurface can also be seen as medium independent transformations. For example, the method of controlled illumination (Rietveld, 1995) always leads to a plane wave in the subsurface, and the common focus point illumination (Thorbecke, 1997; Berkhout, 1997a,b) always leads to a point source in the subsurface.

[2] In the sequel the source distribution for downgoing waves is also referred to as the downgoing source distribution.

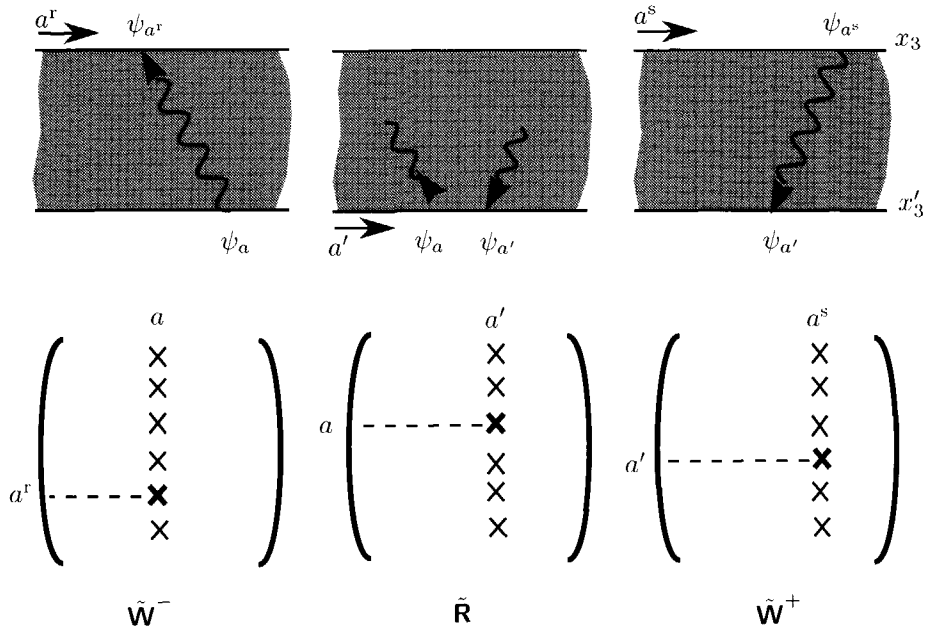


Fig. 4.1 *Representation of a monochromatic seismic experiment in the domain constituted by the complete set of functions* $\{\psi_a\}$.

ψ_{a^s} at x_3. The downgoing wave field is measured at x'_3 with a set of receivers weighted with the function $\psi_{a'}$. The kernel $\tilde{\mathrm{R}}(a, x'_3; a', x'_3)$ denotes the response due to a downgoing wave field of the type $\psi_{a'}$ at x'_3. The upgoing response is measured with a set of receivers weighted with ψ_a also at depth level x'_3. The propagator kernel $\tilde{\mathrm{W}}^-(a^r, x_3; a, x'_3)$ denotes the upgoing response at x_3 related to an upgoing wave field of the type ψ_a at x'_3. The wave field is measured with a set of receivers weighted with ψ_{a^r} at depth level x_3. The upgoing wave field related to the source distribution $\tilde{S}^+(a^s, x_3)$ is denoted by $\tilde{P}^-(a^r, x_3)$. In practice a discretization with respect to the variable a is utilized via an appropriate discretization scheme, which reduces the kernel equation (4.5) to a matrix-vector equation according to[3]

$$\tilde{\mathbf{p}}^-(x_3) = \tilde{\mathbf{W}}^-(x_3; x'_3)\, \tilde{\mathbf{R}}(x'_3)\, \tilde{\mathbf{W}}^+(x'_3; x_3)\, \tilde{\mathbf{s}}^+(x_3). \tag{4.6}$$

This is the discrete seismic data representation constituted by the set $\{\psi_a\}$. In Figure 4.1, the interpretation of a single experiment in the ψ_a-domain is schematically illustrated. Moreover, it shows how the propagator matrices $\tilde{\mathbf{W}}^{\pm}$ and the reflectivity matrix $\tilde{\mathbf{R}}$ are filled. For example, one column of the matrix $\tilde{\mathbf{W}}^+$ denotes the response at x'_3 measured in terms of the discrete set of functions ψ_a due to a downgoing source wave field at x_3 of the form ψ_{a^s}.

[3]The matrix-vector notation is introduced in appendix A.

In the next four subsections I will discuss the representations in the four domains mentioned previously and extensively dealt with in chapter 2. The representation in the wavelet domain will be discussed in more detail than the other representations.

4.2.1 Spatial domain representation

The spatial domain representation has already been discussed in chapter 2. For completeness, I repeat it here for the specific configuration considered in this chapter (cf. equation 3.27)

$$P^-(x_1^{\mathrm{r}}, x_3) = \int_{x_1}\int_{x_1'}\int_{x_1^{\mathrm{s}}} \mathrm{W}^-(x_1^{\mathrm{r}}, x_3; x_1, x_3')\, \mathrm{R}(x_1, x_3'; x_1', x_3')\, \mathrm{W}^+(x_1', x_3'; x_1^{\mathrm{s}}, x_3)\, S^+(x_1^{\mathrm{s}}, x_3)\, \mathrm{d}x_1^{\mathrm{s}}\, \mathrm{d}x_1'\, \mathrm{d}x_1. \quad (4.7)$$

The function $S^+(x_1^{\mathrm{s}}, x_3)$ denotes the spatial downgoing source distribution at depth level x_3. The propagator kernel for downgoing waves $\mathrm{W}^+(x_1', x_3'; x_1^{\mathrm{s}}, x_3)$ denotes the spatial response at x_3' as a function of x_1' due to a downgoing point source[4] at (x_1^{s}, x_3). The reflection kernel $\mathrm{R}(x_1, x_3'; x_1', x_3')$ denotes the upgoing response measured at (x_1, x_3') due to a downgoing point source at (x_1', x_3'). The nature of the reflection kernel will be discussed in more detail in section 4.4. The propagator kernel for upgoing waves $\mathrm{W}^-(x_1^{\mathrm{r}}, x_3; x_1, x_3')$ denotes the response measured at depth level x_3 as a function of x_1^{r} due to a upgoing point source at (x_1, x_3'). The function $P^-(x_1^{\mathrm{r}}, x_3)$ contains the upgoing wave field at x_3 as a function of x_1^{r}.

In the case that the lateral spatial coordinates are discretized equation (4.7) reduces to a matrix-vector equation according to (cf. equation 4.6)

$$\mathbf{P}^-(x_3) = \mathbf{W}^-(x_3, x_3')\, \mathbf{R}(x_3')\, \mathbf{W}^+(x_3', x_3)\, \mathbf{S}^+(x_3). \quad (4.8)$$

A more general notation can be obtained if we put the responses due to a series of seismic experiments together in this system, according to (Berkhout, 1982)

$$\mathbf{P}^-(x_3) = \mathbf{W}^-(x_3; x_3')\, \mathbf{R}(x_3')\, \mathbf{W}^+(x_3'; x_3)\, \mathbf{S}^+(x_3), \quad (4.9)$$

where for a series of N_s shots the source matrix $\mathbf{S}^+$ and the data matrix $\mathbf{P}^-$ are given by

$$\mathbf{P}^- = \left(\mathbf{p}_1^-, \mathbf{p}_2^-, \ldots, \mathbf{p}_{N_s}^-\right) \quad \text{and} \quad \mathbf{S}^+ = \left(\mathbf{s}_1^+, \mathbf{s}_2^+, \ldots, \mathbf{s}_{N_s}^+\right).$$

[4]The flux-normalized decomposition that has been utilized in chapter 3 does not allow for an interpretation in terms of dipole sources and monopole receivers, which is the case if an acoustic pressure normalization would have been used. Without specifying the exact nature of the source and receiver characteristic, I refer to them as one-way, downgoing or upgoing, point sources.

One element of the data matrix $\mathbf{P}^-$ denotes one seismic experiment in the space domain. The interpretation of such an experiment in terms of the propagator matrix $\mathbf{W}^+$, reflectivity matrix $\mathbf{R}$, and propagator matrix $\mathbf{W}^-$ is shown in Figure 4.2a.

In section 4.3 I will study the behavior of the propagator kernel for a homogeneous medium under the wavelet transform. I will give now the expressions for the propagator kernel in the homogeneous situation. The kernels $\mathrm{W}^{\pm}$ are oscillatory integral kernels. For a 2-D medium, homogeneous between x_3 and x_3', the kernels take the following form

$$\mathrm{W}^{\pm}(x_1, x_3; x_1', x_3') = -\frac{jk}{2} \cos\varphi \, H_1^{(2)}(kr), \tag{4.10}$$

where $H_1^{(2)}$ is the first order Hankel function of the second kind, the wavenumber k is given by $k = \omega/c$, the distance r between (x_1, x_3) and (x_1', x_3') is given by $r = \sqrt{(x_1 - x_1')^2 + (x_3 - x_3')^2}$, and the angle φ denotes the angle between the vector $\boldsymbol{i}_3$ and the vector pointing from (x_1', x_3') to (x_1, x_3), hence $\cos\varphi = \frac{|x_3 - x_3'|}{r}$. In the far-field, where $kr \gg 1$, the propagator kernel is given by

$$\mathrm{W}^{\pm}(x_1, x_3; x_1', x_3') = \sqrt{\frac{jk}{2\pi}} \cos\varphi \frac{\mathrm{e}^{-jkr}}{\sqrt{r}}. \tag{4.11}$$

In the recursive depth migration, to be discussed in this chapter, the propagator matrices are filled with the help of local explicit operators (Holberg, 1988; Thorbecke, 1997). Because in recursive depth migration the depth step is small, propagator matrices constructed in this way are here referred to as near-field matrices[5]. Figure 4.3 shows a near-field and far-field propagator matrix for $(x_3 - x_3') = 10\,\mathrm{m}$ and $(x_3 - x_3') = 500\,\mathrm{m}$, respectively.

4.2.2 Spatial Fourier domain representation

In chapter 2 I have dealt in some detail with the Fourier transform and its properties. For the application to the lateral spatial coordinate the Fourier transform takes the following form[6]

$$\tilde{F}(k_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} F(x_1) \exp(jk_1 x_1) \, \mathrm{d}x_1, \tag{4.12}$$

and its inverse by

$$F(x_1) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \tilde{F}(k_1) \exp(-jk_1 x_1) \, \mathrm{d}k_1, \tag{4.13}$$

[5]Note, however, that a near-field approximation has not been carried out.

[6]Note that the spatial Fourier transforms have an opposite sign in the complex exponential with respect to the temporal Fourier transform. With this choice, one can associate with positive k_1 and ω, waves propagating in the positive $\boldsymbol{i}_1$-direction (Aki and Richards, 1980, p. 129).

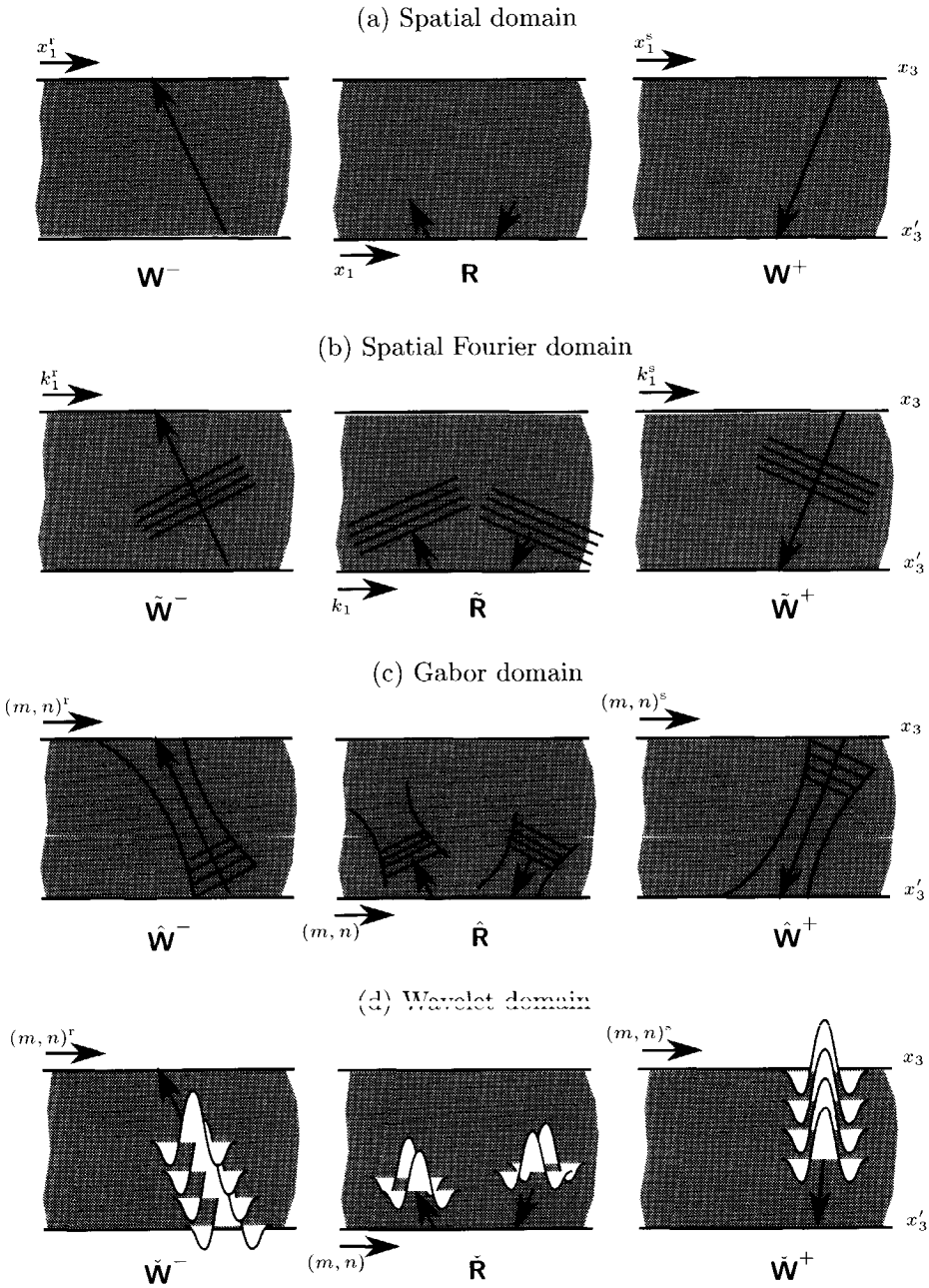


Fig. 4.2 *An artist impression of the representation of a monochromatic seismic experiment in different domains. (a) A single experiment in the space domain (section 4.2.1, eq. 4.9). (b) A single plane-wave experiment (section 4.2.2, eq. 4.17). (c) A single beam experiment (section 4.2.3, eq. 4.22). (d) A single multiscale experiment (section 4.2.4, eq. 4.28).*

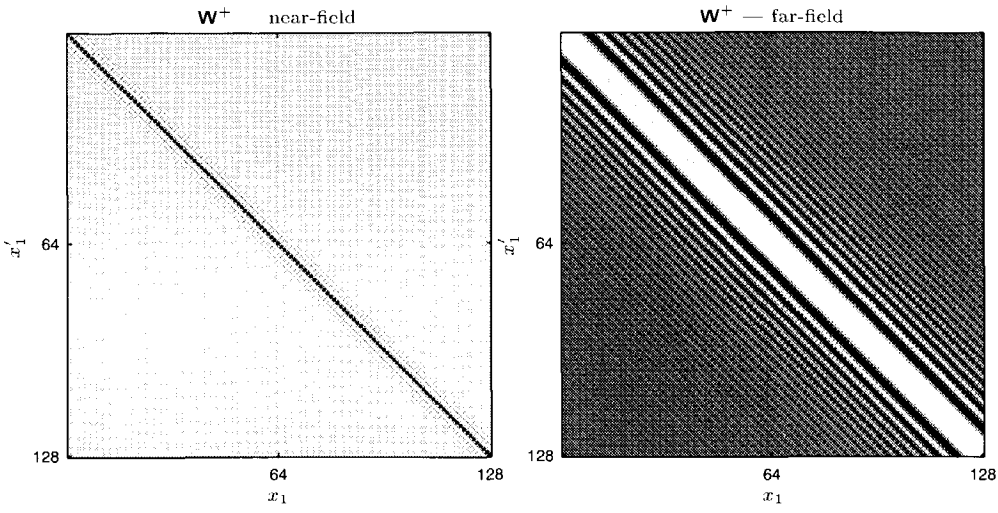


Fig. 4.3 *Propagator kernel in the near-field and far-field. The frequency is chosen to be 30 Hz; the velocity is taken* $c = 2\,km/s$. *(a) The propagator kernel* $\mathrm{W}^+(x_1, x_3; x_1', x_3')$ *for* $(x_3 - x_3') = 10\,m$ *constructed with local explicit operators of 19 points. (b) The propagator kernel for* $(x_3 - x_3') = 500\,m$ *obtained with the far-field expression of equation* (4.11). *Note that* x_1 *and* x_1' *are given in sample numbers. The sampling distance is* $25\,m$.

where k_1 is the lateral spatial Fourier variable associated with x_1. The function $\tilde{F}(k_1)$ is the realization of F in the spatial Fourier domain. The spatial Fourier transform decomposes a wave field at a certain depth level in plane wave components. In the spatial Fourier domain the data representation of equation (4.5) can be written as

$$\tilde{P}^-(k_1^{\mathrm{r}}, x_3) = \int_{k_1} \int_{k_1'} \int_{k_1^{\mathrm{s}}} \tilde{\mathrm{W}}^-(k_1^{\mathrm{r}}, x_3; k_1, x_3')\, \tilde{\mathrm{R}}(k_1, x_3'; k_1', x_3')\, \tilde{\mathrm{W}}^+(k_1', x_3'; k_1^{\mathrm{s}}, x_3)\, \tilde{S}^+(k_1^{\mathrm{s}}, x_3)\, \mathrm{d}k_1^{\mathrm{s}}\, \mathrm{d}k_1'\, \mathrm{d}k_1. \quad (4.14)$$

Equation (4.14) is the monochromatic one-way representation of primary reflection data in the spatial Fourier domain. On the right-hand side, one encounters, from right to left the downgoing source distribution, downward propagation from depth level x_3 to x_3', reflection at depth level x_3', upward propagation from depth level x_3' to x_3, all in terms of plane wave components. The left-hand side of equation (4.14) denotes the seismic reflection data decomposed in plane wave components. The representation of the kernels in the spatial Fourier domain can be obtained via equation (4.3). Alternatively, a kernel in the spatial Fourier domain can be obtained by applying a double Fourier transform to the kernel representation in the spatial

domain. For example, the propagator kernel $\tilde{\mathrm{W}}^+$ can be obtained via

$$\tilde{\mathrm{W}}^+(k_1', x_3'; k_1, x_3) = \frac{1}{2\pi} \int\int \exp(jk_1' x_1')\, \mathrm{W}^+(x_1', x_3'; x_1, x_3)\, \exp(-jk_1 x_1)\, \mathrm{d}x_1\, \mathrm{d}x_1'. \quad (4.15)$$

In the case that the spatial Fourier coordinates are discretized equation (4.14) reduces to a matrix-vector equation according to (cf. equation 4.8)

$$\tilde{\mathbf{p}}^-(x_3) = \tilde{\mathbf{W}}^-(x_3; x_3')\, \tilde{\mathbf{R}}(x_3')\, \tilde{\mathbf{W}}^+(x_3'; x_3)\, \tilde{\mathbf{s}}^+(x_3). \quad (4.16)$$

A more general notation in the spatial Fourier domain can be obtained if we put the responses due to a series of seismic plane-wave experiments together in this system, according to (cf. equation 4.9)

$$\tilde{\mathbf{P}}^-(x_3) = \tilde{\mathbf{W}}^-(x_3; x_3')\, \tilde{\mathbf{R}}(x_3')\, \tilde{\mathbf{W}}^+(x_3'; x_3)\, \tilde{\mathbf{S}}^+(x_3). \quad (4.17)$$

Note that this procedure requires an additional spatial Fourier transform along the source coordinate. One element of the double Fourier transformed data matrix $\tilde{\mathbf{P}}^-$ represents one experiment in the spatial Fourier domain. The structure of such an experiment is illustrated in Figure 4.2b on page 118.

Let me discuss the structure of the propagator $\mathcal{W}^+(x_3'; x_3)$ in the Fourier domain again. The kernel $\tilde{\mathrm{W}}^+(k_1', x_3'; k_1, x_3)$ denotes the downgoing response of a downgoing plane wave at x_3 measured with plane wave receivers at depth level x_3'. For a homogeneous medium or a laterally invariant medium the propagator in the spatial Fourier domain gets a diagonal structure and reduces to the so-called phase-shift operator. This property can be used advantageously in seismic migration as has been shown by Gazdag (1978) and by Stolt (1978). In chapter 3, section 3.3.4, it has been derived that in this situation the spatial Fourier transform is equal to the modal decomposition. For laterally varying medium configurations plane wave interaction takes place, which means that a plane wave gets distorted while propagating in the vertical direction[7]. The numerical computation of the kernel in that situation is approached in various ways, for example via the phase-shift plus interpolation method or via a so-called split-step Fourier method (Gazdag and Sguazzero, 1984; Pai, 1988; Stoffa et al., 1990; Pai, 1991; Lee et al., 1991).

4.2.3 Gabor domain representation

In equation (4.7) a pure spatial description of a seismic experiment is given. On the other hand, in equation (4.14) a description of a seismic experiment is given in terms of plane waves of infinite lateral size. A natural intermediate domain is provided

[7]The modal decomposition of chapter 3 still leads to a diagonal kernel, under the condition that the medium is not varying in the vertical direction between x_3 and x_3'.

by the Gabor transform which has been introduced in full detail in chapter 2. The spatial discrete Gabor transform of a wave field is given by (cf. equation 2.99)

$$\hat{F}(m,n) = \langle F, g_{mn} \rangle = \int_{-\infty}^{\infty} F(x_1) \exp(jma_0x_1) g^*(x_1 - nb_0)\, \mathrm{d}x_1, \tag{4.18}$$

which can be interpreted as a local spatial Fourier transform for each (m,n)-pair with a spatial concentration around $x_1 = nb_0$ and a wavenumber concentration around $k_1 = ma_0$. Hence, a wave field is locally decomposed into its plane wave components. The window function $g(x_1)$ is generally chosen to be a Gaussian function as described in chapter 2. For a stable reconstruction of $F(x_1)$ the parameters a_0 and b_0 have to be chosen such that $a_0b_0 < 2\pi$. The reconstruction of $F(x_1)$ from its Gabor coefficients reads

$$F(x_1) = \sum_{m,n} \hat{F}(m,n)\, \underline{g}_{mn}(x_1), \tag{4.19}$$

where $\underline{g}_{mn}$ is given by

$$\underline{g}_{mn}(x_1) = \exp(-jma_0x_1)\underline{g}(x_1 - nb_0).$$

The function $\underline{g}(x_1)$ is referred to as the dual function of $g(x_1)$. Its actual form depends on the function g and on the values a_0 and b_0. A number of dual functions of a Gaussian function g have been shown in Figure 2.9. Note that the Gabor transform used here necessarily leads to a discrete function in the transformed domain.

Due to the discrete nature of the Gabor transform the seismic data representation of equation (4.5) automatically takes a matrix-vector form . The Gabor domain representation reads

$$\hat{\mathbf{p}}^-(x_3) = \hat{\mathbf{W}}^-(x_3;x_3')\, \hat{\mathbf{R}}(x_3')\, \hat{\mathbf{W}}^+(x_3';x_3)\, \hat{\mathbf{s}}^+(x_3). \tag{4.20}$$

Equation (4.20) is the monochromatic one-way representation of the primary reflection data related to reflecting depth level x_3' in the Gabor domain. On the right-hand side, one encounters, from right to left, the downgoing source distribution, downward propagation from x_3 to x_3', reflection at depth level x_3', and upward propagation from x_3' to x_3, all expressed in Gabor coefficients. More specifically, each of the matrices $\hat{\mathbf{W}}^-(x_3;x_3')$, $\hat{\mathbf{R}}(x_3')$, and $\hat{\mathbf{W}}^+(x_3';x_3)$ consists of elements denoting the interaction between different Gabor functions. For example, the propagator matrix $\hat{\mathbf{W}}^+(x_3';x_3)$ consists of elements $\hat{\mathrm{W}}^+_{mnm'n'}(x_3';x_3)$ which are given by

$$\hat{\mathrm{W}}^+_{mnm'n'}(x_3';x_3) = \langle \mathcal{W}^+(x_3';x_3)\ \underline{g}_{m'n'}, g_{mn} \rangle. \tag{4.21}$$

The other matrices are defined accordingly. The element $\hat{\mathrm{W}}^+_{mnm'n'}$ denotes a response of a beam departing at depth level x_3 around position $x_1 = n'b_0$ with a

central wavenumber $k_1 = m'a_0$. The response is measured at depth level x_3' around location $x_1 = nb_0$ and around a wavenumber $k_1 = ma_0$.

The significance of the Gabor domain representation comes across if one realizes that the Gabor transform leads to a division of the data in Gaussian beam experiments (Einziger et al., 1986; Bastiaans, 1980). The Gaussian beam is an advantageous mathematical entity. First, because the Gaussian beam is a solution of the parabolic wave equation with a source point in the complex plane (Deschamps, 1971). Secondly, the central ray of a Gaussian beam adheres to the eikonal and transport equation of geometrical optics (Červený et al., 1982). These two aspects allow for a relatively easy computation of Gaussian beams even in complex media. That is why Gaussian beams can be advantageously used in migration (Raz, 1987; Hill, 1990). However, generally Gaussian beam migration is carried out with a redundant set of beams. The added value of the Gabor domain approach lies in the flexibility with respect to window type, window size and window distance, giving a good mathematical control over the amount of redundancy.

Let me finally discuss the multi-source beam version. For a series of seismic experiments the Gabor transform can also be applied to the source coordinate, yielding a double Gabor transformed source matrix and a double Gabor transformed data matrix. The data representation takes the following form (cf. equation 4.9 and 4.17)

$$\hat{\mathbf{P}}^{-}(x_3) = \hat{\mathbf{W}}^{-}(x_3; x_3')\,\hat{\mathbf{R}}(x_3')\,\hat{\mathbf{W}}^{+}(x_3'; x_3)\,\hat{\mathbf{S}}^{+}(x_3). \tag{4.22}$$

The physical meaning of the Gabor domain approach is illustrated in Figure 4.2c on page 118. One element of the data matrix $\hat{\mathbf{P}}^{-}$ in the Gabor domain denotes one Gaussian beam experiment.

4.2.4 Wavelet domain representation

Whereas the previous three data representations have obtained directly or indirectly significant attention, the wavelet transform[8] is a relatively new tool and its usefulness for the data representation of equation (4.2) has not been studied in great detail yet. In chapter 2 two main assets of the wavelet transform have been identified:

1. a wavelet is blind to something smoothly varying;
2. a wavelet transform allows for a natural decomposition in approximations and details.

Before I address these assets in relation to the seismic data representation, the use of the discrete wavelet transform in two dimensions has to be treated. Following the results of chapter 2, especially section 2.5.4, the discrete wavelet transform within a multiresolution approximation starts with a representation of a function in a certain

[8]Since I am aiming for reasonably efficient schemes I have to use the discrete wavelet transform.

fine approximation space V_0. The discrete approximation can be obtained by taking the inner product of the function with a set of smoothing functions $\{\phi_{mn}\}$ with $m = 0$ and $n \in \mathbb{Z}$. The smoothing functions are given by (cf. equation 2.109)

$$\phi_{mn}(x) = 2^{-m/2}\phi(2^{-m}x - n).$$

Accordingly, a representation of an operator $\mathcal{K}$ (or a two-dimensional function) on a certain fine scale amounts to taking inner products with smoothing functions in $\boldsymbol{V}_0 = V_0 \otimes V_0$, according to

$$\mathrm{K}_{nn'} = \langle \mathcal{K}\phi_{0n'}, \phi_{0n} \rangle . \tag{4.23}$$

Together, the elements $\mathrm{K}_{nn'}$, with $n, n' \in \{0, \ldots, N-1\}$ form a discrete approximation matrix $\mathbf{K}$[9]. The value of N depends on the aperture size and on the resolution m. For larger resolutions less functions ϕ_{mn} "fit" within the aperture. The decomposition of the matrix $\mathbf{K}$ in a coarser approximation and details can be done in different ways. Here, I distinguish between the standard and non-standard 2-D decomposition (Beylkin et al., 1991; Beylkin, 1992; Daubechies, 1992). The standard 2-D decomposition is obtained by applying the discrete decomposition to the two directions in the matrix independently, which means that the two-dimensional subspace $\boldsymbol{V}_0$ is decomposed according to (cf. equations 2.116 and 2.117)

$$\begin{aligned}\boldsymbol{V}_0 &= V_0 \otimes V_0 \\ &= (V_M \oplus O_M \oplus O_{M-1} \oplus \cdots \oplus O_1) \otimes (V_M \oplus O_M \oplus O_{M-1} \oplus \cdots \oplus O_1) .\end{aligned}$$

Alternatively, the non-standard 2-D decomposition is such that it uncouples scales, which means that $\boldsymbol{V}_0$ is decomposed step by step. The first step takes the following form

$$\begin{aligned}\boldsymbol{V}_0 &= V_0 \otimes V_0 \\ &= (V_1 \oplus O_1) \otimes (V_1 \oplus O_1) \\ &= (V_1 \otimes V_1) \oplus (V_1 \otimes O_1) \oplus (O_1 \otimes V_1) \oplus (O_1 \otimes O_1) \\ &= \boldsymbol{V}_1 \oplus \boldsymbol{O}_1,\end{aligned}$$

where $\boldsymbol{V}_1 = V_1 \otimes V_1$, and $\boldsymbol{O}_1 = (V_1 \otimes O_1) \oplus (O_1 \otimes V_1) \oplus (O_1 \otimes O_1)$. Subsequent steps can be used to further decompose $\boldsymbol{V}_1$. The decomposition of the space $\boldsymbol{V}_0$ up to a certain coarse resolution M can thus be written as

$$\boldsymbol{V}_0 = \boldsymbol{V}_M \oplus \boldsymbol{O}_M \oplus \cdots \oplus \boldsymbol{O}_2 \oplus \boldsymbol{O}_1.$$

[9]The matrix $\mathbf{K}$ can be seen as the two-dimensional version of the discrete approximation $\mathcal{A}_0^d f$ of a function f introduced in equation (2.111).

The uncoupling of scales in the non-standard wavelet transform turns out to be advantageous for the seismic data representation in the wavelet domain. Hence, I will further consider the non-standard 2-D wavelet transform[10].

Since the set $\{\phi_{1n}\}$ forms an orthonormal basis of V_1 and the set $\{\psi_{1n}\}$ forms an orthonormal basis of O_1, one step of the non-standard decomposition of $\boldsymbol{V}_0$ yields for the matrix $\mathbf{K}$ the following structure

$$\check{\mathbf{K}}_1 = \begin{pmatrix} \mathbf{K}_{T(1)} & \mathbf{K}_{B(1)} \\ \mathbf{K}_{C(1)} & \mathbf{K}_{A(1)} \end{pmatrix}, \tag{4.24}$$

where, in general, the submatrices $\mathbf{K}_{T(m)}$, $\mathbf{K}_{B(m)}$, $\mathbf{K}_{C(m)}$, and $\mathbf{K}_{A(m)}$ are given by

$$\mathbf{K}_{T(m)} = \langle \mathcal{K}\phi_{mn'}, \phi_{mn} \rangle \qquad \mathbf{K}_{B(m)} = \langle \mathcal{K}\psi_{mn'}, \phi_{mn} \rangle \tag{4.25}$$

$$\mathbf{K}_{C(m)} = \langle \mathcal{K}\phi_{mn'}, \psi_{mn} \rangle \qquad \mathbf{K}_{A(m)} = \langle \mathcal{K}\psi_{mn'}, \psi_{mn} \rangle, \tag{4.26}$$

where $m \in \{1, 2, \ldots, \log N\}$, and $n, n' \in \{0, \ldots, N/2^m - 1\}$. The submatrix $\mathbf{K}_{T(1)}$ can be further decomposed yielding the matrix $\check{\mathbf{K}}_2$ consisting of the submatrices $\mathbf{K}_{T(2)}$, $\mathbf{K}_{B(2)}$, $\mathbf{K}_{C(2)}$, and $\mathbf{K}_{A(2)}$, etc. It is interesting to observe that $\mathbf{K}_{T(0)}$ equals the matrix $\mathbf{K}$ defined via equation (4.23). Since $\mathbf{K}_{T(0)}$ is referred to as the approximation at resolution $m = 0$, the matrices $\mathbf{K}_{T(m)}$ are referred as the approximations at resolution m. The submatrices $\mathbf{K}_{B(m)}$, $\mathbf{K}_{C(m)}$, $\mathbf{K}_{A(m)}$ are referred to as the detail parts at the resolution m. The decomposition from an approximation at a particular resolution to approximations and details at coarser resolutions is carried out with the fast wavelet transform in $O(N^2)$ operations for an $(N \times N)$-matrix.

The application of one step of the non-standard wavelet transform to the data representation of equation (4.2), according to the scheme of the previous paragraph, yields

$$\check{\mathbf{p}}^-(x_3) = \check{\mathbf{W}}_1^-(x_3; x_3')\, \check{\mathbf{R}}_1(x_3')\, \check{\mathbf{W}}_1^+(x_3'; x_3)\, \check{\mathbf{s}}_1^+(x_3), \tag{4.27}$$

or in extended form, using equation (4.24),

$$\check{\mathbf{p}}^-(x_3) = \begin{pmatrix} \mathbf{W}_{T(1)}^- & \mathbf{W}_{B(1)}^- \\ \mathbf{W}_{C(1)}^- & \mathbf{W}_{A(1)}^- \end{pmatrix} \begin{pmatrix} \mathbf{R}_{T(1)} & \mathbf{R}_{B(1)} \\ \mathbf{R}_{C(1)} & \mathbf{R}_{A(1)} \end{pmatrix} \begin{pmatrix} \mathbf{W}_{T(1)}^+ & \mathbf{W}_{B(1)}^+ \\ \mathbf{W}_{C(1)}^+ & \mathbf{W}_{A(1)}^+ \end{pmatrix} \check{\mathbf{s}}^+(x_3). \tag{4.28}$$

Equation (4.28) is the monochromatic one-way representation of the primary reflection related to reflecting depth level x_3' in the wavelet domain. On the right-hand side, one encounters, from right to left, the downgoing source distribution, downward propagation from x_3 to x_3', reflection at depth level x_3', and upward propagation from x_3' to x_3, all expressed in wavelet coefficients at resolution 1. More specifically, the

[10]For more details with respect to difference between the standard and non-standard form the reader is referred to Beylkin et al. (1991) or Beylkin (1992). There it is also described how the standard form can be computed from the non-standard form.

submatrices of the propagator matrix $\check{\mathbf{W}}_1^+$ in the wavelet domain can be interpreted in the following way: $\mathbf{W}^+_{T(1)}$ contains the response of (laterally) smooth downgoing source wave fields at resolution 1, measured with (laterally) smooth receivers at resolution 1; the submatrix $\mathbf{W}^+_{B(1)}$ represents the response of downgoing source wave fields laterally weighted with a wavelet at resolution 1, also referred to as downgoing detail source wave fields, measured by smooth receivers at the same resolution; $\mathbf{W}^+_{C(1)}$ is opposite to $\mathbf{W}^+_{B(1)}$: $\mathbf{W}^+_{C(1)}$ is a submatrix representing the response of a number of downgoing laterally smooth wave fields at x_3, measured with a set of receivers weighted with a wavelet at resolution 1 at x_3', also referred to as detail receivers at x_3'; finally, the submatrix $\mathbf{W}^+_{A(1)}$ contains the response of a number of detail sources at x_3, measured with a number of detail receivers at x_3'.

In section 4.5 I will make use of the multi-source version introduced in equation (4.9). For an appropriately chosen source and receiver geometry the data representation in the wavelet transform domain takes the form

$$\check{\mathbf{P}}_1^-(x_3) = \check{\mathbf{W}}_1^-(x_3; x_3')\,\check{\mathbf{R}}_1(x_3')\,\check{\mathbf{W}}_1^+(x_3'; x_3)\,\check{\mathbf{S}}_1^+(x_3), \tag{4.29}$$

or in extended form, using equation (4.24),

$$\begin{pmatrix} \mathbf{P}^-_{T(1)} & \mathbf{P}^-_{B(1)} \\ \mathbf{P}^-_{C(1)} & \mathbf{P}^-_{A(1)} \end{pmatrix} = \begin{pmatrix} \mathbf{W}^-_{T(1)} & \mathbf{W}^-_{B(1)} \\ \mathbf{W}^-_{C(1)} & \mathbf{W}^-_{A(1)} \end{pmatrix} \begin{pmatrix} \mathbf{R}_{T(1)} & \mathbf{R}_{B(1)} \\ \mathbf{R}_{C(1)} & \mathbf{R}_{A(1)} \end{pmatrix} \begin{pmatrix} \mathbf{W}^+_{T(1)} & \mathbf{W}^+_{B(1)} \\ \mathbf{W}^+_{C(1)} & \mathbf{W}^+_{A(1)} \end{pmatrix} \begin{pmatrix} \mathbf{S}^+_{T(1)} & \mathbf{S}^+_{B(1)} \\ \mathbf{S}^+_{C(1)} & \mathbf{S}^+_{A(1)} \end{pmatrix}. \tag{4.30}$$

Each element of the data matrix $\check{\mathbf{P}}_1^-$ denotes a single scale experiment in the wavelet transform domain. Figure 4.2d on page 118 schematically illustrates the structure of a single seismic experiment in the wavelet domain. Equation (4.30) will be the point of departure for seismic migration in the wavelet transform domain. Note that the submatrices denoted with the subscript $T(1)$ can be further decomposed in details and approximations.

Although I have been able to obtain a seismic data representation in the wavelet domain, I did not explicitly address the assets of the wavelet transform as given on page 122. In order to do so, the structure of the propagator in the wavelet domain has to be studied in more detail. This issue will be discussed in the next section.

4.3 Properties of the propagator in the wavelet domain

In this section the non-standard wavelet transform of the propagator matrices $\mathbf{W}^{\pm}$ will be studied. The emphasis will be put on the general structure of the propagator in the wavelet domain and whether one of the assets listed in the previous section

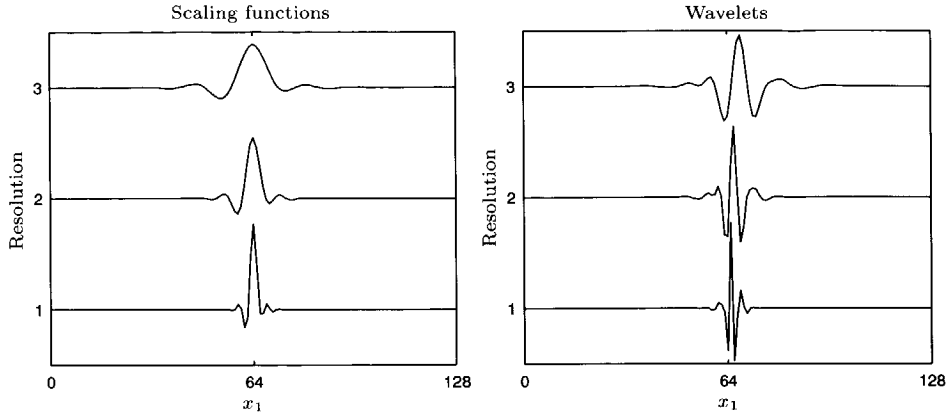


Fig. 4.4 *Quadrature mirror filters* **h** *(left) and* **g** *(right) at different scales for nearly symmetric Daubechies scaling functions and wavelets for 10 vanishing moments. See Daubechies (1992) for details concerning this type of wavelet and scaling function.*

can be used. The study will be limited to homogeneous medium configurations. Since analytic solutions of the propagator in the wavelet domain are not available, all of the wavelet domain propagators are numerically transformed from their space domain realizations.

Let me start to fix the choice of the wavelet filter pair **h** and **g**. In chapter 2 it has been explained how the filter pair is related to the wavelet and scaling function. Since the propagator matrix is symmetric, the wavelet and scaling function should be symmetric as well. Since the propagator matrix has a certain (oscillatory) smoothness the wavelet should have a number of vanishing moments to capture at least partially the structure of the propagator. Since the wavelet transforms should be carried out efficiently, short filter pairs are required. Since the sources and receivers should be interpreted in the same way, orthogonality is required (in contrast to bi-orthogonal wavelets). This is a set of conflicting requirements. Without claiming that I have made the absolute best choice, I decided to use nearly symmetric Daubechies wavelets and scaling functions with 10 vanishing moments. Wagner and Chew (1995) report reasonably good results with similar wavelets. Alternative wavelet choices are discussed in Alpert et al. (1993), and Goswami et al. (1995). The filter pairs are shown in Figure 4.4 at three different resolutions m.

Consider the propagator kernel for the far-field given in equation (4.11). The propagator matrix in the space domain and its non-standard wavelet transform for one, two and three steps are shown in Figure 4.5. The frequency is here 30 Hz. The lateral sampling is $\Delta x_1 = 25\,\mathrm{m}$, and the compressional wave velocity is $2500\,\mathrm{m/s}$.

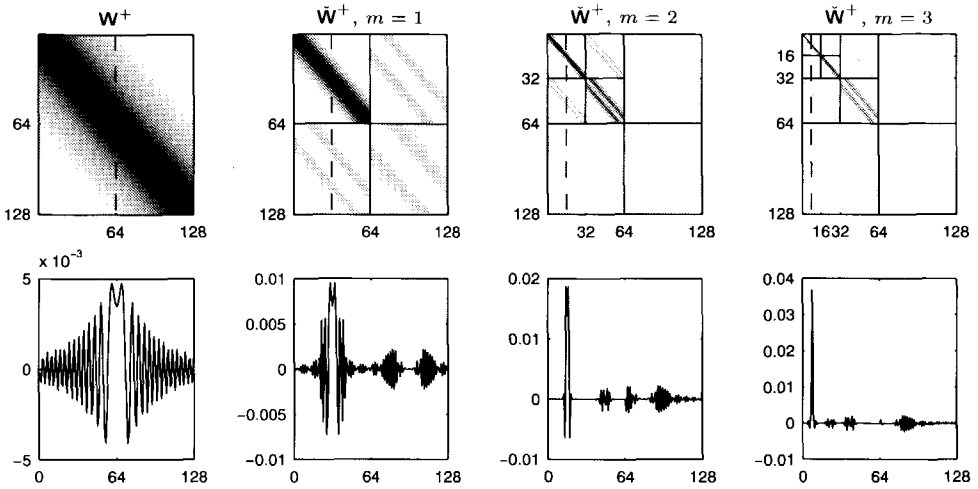


Fig. 4.5 *Propagator matrix for the far-field in the space domain and in the wavelet domain for one, two and three steps of the non-standard wavelet transform. The axis labels are in sample numbers. The spatial propagator matrix has been generated for $f = 30$ Hz, a depth step of 500 m, and a lateral sampling of $\Delta x_1 = 25$ m. The top row shows the absolute values of the matrices and the bottom row shows one column of each of the matrices.*

The depth step is taken $|x_3 - x_3'| = 500\,\mathrm{m}$. Due to the oscillatory nature of the propagator kernel, the realizations in the wavelet domain are not much sparser. It is also clear that interaction between scaling functions and wavelets takes place. Wagner and Chew (1995) and Goswami et al. (1995) carry out more detailed analyses with respect to the sparseness. I will not repeat their results here.

Consider the near-field realization of the propagator in the space domain, obtained with the local explicit operator method. The frequency, the compressional wave velocity and the lateral sampling are taken the same as for the far-field matrix. The depth step is taken $|x_3 - x_3'| = 10\,\mathrm{m}$. The propagator matrix is shown in Figure 4.6. For one, two and three steps of the non-standard wavelet transform, the propagator matrices are shown in the same figure. The top row shows the absolute values of the matrices, the middle row the sparsity structure, and the bottom row shows one column of each of the matrices. For small depth steps the off-diagonal matrices, $\mathbf{W}^+_{B(m)}$ and $\mathbf{W}^+_{C(m)}$ are small with respect to the diagonal submatrices $\mathbf{W}^+_{T(m)}$ and $\mathbf{W}^+_{A(m)}$, depending on the velocity and frequency. In the remainder of this chapter I will neglect the off-diagonal submatrices. From the structure of the propagators in the far-field I conclude that the neglect will certainly introduce some distortions. I decided to accept this error for the sake of a tractable algorithm.

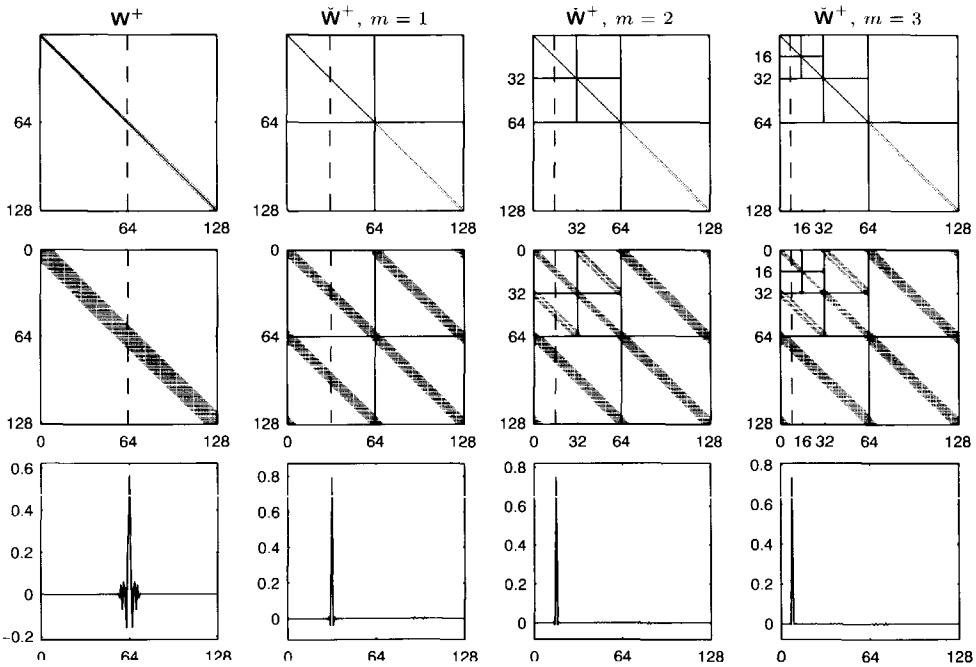


Fig. 4.6 *Near-field expression for the propagator matrix in the space domain and in the wavelet domain for one, two and three steps of the non-standard wavelet transform. The depth step is* 10 m. *The other parameters are equal as for the far-field matrix. The top row shows the absolute values of the matrix. The middle row shows the sparsity structure and the bottom row shows a single column of each of the matrices.*

4.4 Migration

The result of a migration procedure is a so-called structural image of the subsurface. A structural image consists of the locations of the reflecting boundaries. The reflecting boundaries can be derived by extracting the reflectivity operator $\mathcal{R}$ in one of the representations discussed before. In the space domain, the reflectivity kernel $\mathrm{R}(x_1, x_3'; x_1', x_3')$ of the reflectivity operator $\mathcal{R}$ expresses the upgoing response at the point (x_1, x_3') due to a downgoing source at (x_1', x_3'). It is a symmetric kernel containing the full angle-dependent reflectivity. In structural imaging, however, one is interested in the angle-averaged reflectivity, which is given by the diagonal $\mathrm{R}(x_1, x_3'; x_1, x_3')$ of the reflectivity kernel. For a locally-reacting boundary the diagonal of the reflectivity kernel does completely determine the reflection behavior. Such a situation does only occur if the medium parameters are not changing as a function of the lateral coordinates and if the velocity is constant, i.e. only the density is changing in the vertical direction. In all other situations the boundary is not locally-reacting, hence the reflectivity is angle dependent (Aki and Richards, 1980). The diagonal of the spatial reflectivity kernel is also referred to as the zero-offset or angle-averaged reflectivity. In the subsequent subsections the extraction of the reflectivity operator in the space domain and the wavelet domain will be discussed.

4.4.1 Migration in the space domain

Consider the data representation of equation (4.9) in a fixed-spread configuration with the same number of sources as receivers on a regular grid[11]. Hence, in equation (4.9) the matrices are square. Migration comes to the extraction of the reflectivity matrix $\mathbf{R}$. This can be accomplished by removing the propagation effects through the overburden and a subsequent imaging step (Claerbout, 1971; Berkhout, 1982).

Using Berkhout's formulation, the removal of the propagation effects through the overburden is realized through

$$\mathbf{R}(x_3') = \mathbf{F}^-(x_3'; x_3)\mathbf{P}^-(x_3)\mathbf{F}^+(x_3; x_3'), \tag{4.31}$$

which is also referred to as the redatuming step. The matrix $\mathbf{F}^+$ is the representation of an operator which inverts for the propagation from the surface at depth level x_3 to the reflecting boundary at depth level x_3', and for the source geometry and source signature. The matrix $\mathbf{F}^-$ is a representation of an operator which inverts for the propagation from the reflecting boundary at depth level x_3' to the surface at depth level x_3, and for the receiver geometry and receiver signature. Since a fixed-spread geometry is assumed here, one can take for $\mathbf{F}^\pm$ the band-limited inverses

[11] For the subsequent derivations in the space domain these requirements are not necessary. In the wavelet domain, these properties are advantageously used.

of the propagator matrices $\mathbf{W}^{\pm}$ [12]. The inverse of the propagator matrices can be approximated via a least-squares inversion, according to (Berkhout, 1982)

$$\mathbf{F}^{\pm} = \left[\left(\mathbf{W}^{\pm} \right)^{H} \mathbf{W}^{\pm} + \epsilon^2 \mathbf{I} \right]^{-1} \left(\mathbf{W}^{\pm} \right)^{H}, \tag{4.32}$$

where $\epsilon^2 \mathbf{I}$ is a stabilization term and where the superscript H denotes complex conjugation and transposition. The kinematic effects are taken into account by the last factor in the right-hand side of equation (4.32). For the regularly sampled flux-normalized propagators derived in chapter 3 it can be shown that $\left[\left(\mathbf{W}^{\pm} \right)^{H} \mathbf{W}^{\pm} \right] \simeq \mathbf{I}$ (Wapenaar, 1997b), therefore the following approximation is utilized[13]

$$\mathbf{F}^{-}(x_3'; x_3) = \left[\mathbf{W}^{-}(x_3; x_3') \right]^{H} \qquad \mathbf{F}^{+}(x_3; x_3') = \left[\mathbf{W}^{+}(x_3'; x_3) \right]^{H}, \tag{4.33}$$

which is referred to as the matched-filter approach.

The imaging step consists of an estimate of the angle-averaged reflectivity. At depth x_3' it is obtained by an averaging over all frequencies according to

$$\langle \mathbf{R}(x_3') \rangle_{\mathrm{diag}} = \frac{1}{N_\omega} \mathrm{Re} \left(\sum_{\omega} \mathbf{R}(x_3')_{\mathrm{diag}} \right), \tag{4.34}$$

where N_ω is the number of frequencies, and where the subscript 'diag' refers to the fact that the summation over all frequencies is carried out for the diagonal elements only. Henceforth, the subscript diagonal will be omitted for notational convenience.

A few remarks apply to the above derivations. First, the summation over all frequencies is only meaningful for the diagonal of the reflectivity matrix, which is exactly where I am interested in for structural imaging. Berkhout and Wapenaar (1993) have shown that with the help of the Radon transform the off-diagonal elements, hence the angle-dependent effects, can be extracted in a similar way. Secondly, if the migration is carried out per gather, especially the imaging step has to be slightly adjusted (Claerbout, 1971; Rietveld, 1995). Finally, note that $(x-\omega)$-migration is generally carried out recursively with small depth steps. At each depth level the reflectivity is extracted.

[12] Note that the source and receiver signature has been left out. It is assumed that they have been removed during the preprocessing of the seismic reflection data.

[13] If an acoustic-pressure normalization would have been used in the decomposition from the two-way wave equation to the one-way wave equation in chapter 3, the matched-filter approach would take the following form

$$\mathbf{F}^{-}(x_3'; x_3) = \left[\mathbf{W}^{+}(x_3'; x_3) \right]^{*} \qquad \mathbf{F}^{+}(x_3; x_3') = \left[\mathbf{W}^{-}(x_3; x_3') \right]^{*}.$$

4.4.2 Migration in the wavelet domain

The point of departure for migration in the wavelet domain is equation (4.29). Just as in the space domain, migration in the wavelet domain consists of a redatuming step and a subsequent imaging step. For one step of the non-standard wavelet transform the redatuming step can be written as (cf. equations 4.30 and 4.31)

$$\begin{pmatrix} \mathbf{R}_{T(1)} & \mathbf{R}_{B(1)} \\ \mathbf{R}_{C(1)} & \mathbf{R}_{A(1)} \end{pmatrix} = \begin{pmatrix} \mathbf{F}^{-}_{T(1)} & \mathbf{F}^{-}_{B(1)} \\ \mathbf{F}^{-}_{C(1)} & \mathbf{F}^{-}_{A(1)} \end{pmatrix} \begin{pmatrix} \mathbf{P}^{-}_{T(1)} & \mathbf{P}^{-}_{B(1)} \\ \mathbf{P}^{-}_{C(1)} & \mathbf{P}^{-}_{A(1)} \end{pmatrix} \begin{pmatrix} \mathbf{F}^{+}_{T(1)} & \mathbf{F}^{+}_{B(1)} \\ \mathbf{F}^{+}_{C(1)} & \mathbf{F}^{+}_{A(1)} \end{pmatrix}. \tag{4.35}$$

The matrices $\check{\mathbf{F}}^{\pm}_{1}$, given by

$$\check{\mathbf{F}}^{\pm}_{1} = \begin{pmatrix} \mathbf{F}^{\pm}_{T(1)} & \mathbf{F}^{\pm}_{B(1)} \\ \mathbf{F}^{\pm}_{C(1)} & \mathbf{F}^{\pm}_{A(1)} \end{pmatrix}, \tag{4.36}$$

are the wavelet transforms of the matched-filter matrices $\mathbf{F}^{\pm}$. Due to the orthogonality of the wavelet transform they can also be derived as the matched filter of the wavelet transform of $\mathbf{W}^{\pm}$. An efficient redatuming scheme is achieved either if the inverse propagator matrices are sparser than in the space domain, or if it is possible to neglect certain parts of the propagator matrices at the cost of some accuracy. The former is unfortunately not true; the latter is possible as can be concluded from section 4.3. The neglect of the off-diagonal submatrices $\mathbf{W}^{\pm}_{B(1)}$ and $\mathbf{W}^{\pm}_{C(1)}$ or, equivalently, the off-diagonal submatrices $\mathbf{F}^{\pm}_{B(1)}$ and $\mathbf{F}^{\pm}_{C(1)}$, yields for equation (4.35) four independent redatuming equations for $\mathbf{R}_{T(1)}$, $\mathbf{R}_{B(1)}$, $\mathbf{R}_{C(1)}$, and $\mathbf{R}_{A(1)}$ according to

$$\mathbf{R}_{T(1)} = \mathbf{F}^{-}_{T(1)} \mathbf{P}^{-}_{T(1)} \mathbf{F}^{+}_{T(1)} \tag{4.37}$$

$$\mathbf{R}_{B(1)} = \mathbf{F}^{-}_{T(1)} \mathbf{P}^{-}_{B(1)} \mathbf{F}^{+}_{A(1)} \tag{4.38}$$

$$\mathbf{R}_{C(1)} = \mathbf{F}^{-}_{A(1)} \mathbf{P}^{-}_{C(1)} \mathbf{F}^{+}_{T(1)} \tag{4.39}$$

$$\mathbf{R}_{A(1)} = \mathbf{F}^{-}_{A(1)} \mathbf{P}^{-}_{A(1)} \mathbf{F}^{+}_{A(1)}. \tag{4.40}$$

Since the elements of equation (4.37) can be further decomposed into approximations and details, the redatuming equations in the wavelet domain can be written more generally at resolution m as

$$\mathbf{R}_{T(m)} = \mathbf{F}^{-}_{T(m)} \mathbf{P}^{-}_{T(m)} \mathbf{F}^{+}_{T(m)} \tag{4.41}$$

$$\mathbf{R}_{B(m)} = \mathbf{F}^{-}_{T(m)} \mathbf{P}^{-}_{B(m)} \mathbf{F}^{+}_{A(m)} \tag{4.42}$$

$$\mathbf{R}_{C(m)} = \mathbf{F}^{-}_{A(m)} \mathbf{P}^{-}_{C(m)} \mathbf{F}^{+}_{T(m)} \tag{4.43}$$

$$\mathbf{R}_{A(m)} = \mathbf{F}^{-}_{A(m)} \mathbf{P}^{-}_{A(m)} \mathbf{F}^{+}_{A(m)}. \tag{4.44}$$

The imaging step is a little bit more subtle. In the space domain it is given by equation (4.34). Below equation (4.26) it has been noted that there is no conceptual difference between an approximation at resolution $m = 0$, which is seen as the

normal spatial representation, and any of the representations at coarser resolutions. Therefore, the imaging step for the diagonal of the matrices $\mathbf{R}_{T(m)}$ is equal to the spatial imaging of equation (4.34), i.e.

$$\langle \mathbf{R}_{T(m)} \rangle_{\text{diag}} = \frac{1}{N_\omega} \text{Re} \left(\sum_\omega \mathbf{R}_{T(m)\text{diag}} \right), \tag{4.45}$$

where the subscript 'diag' will again be omitted in the sequel. Before the imaging step for the other parts of the reflectivity matrix can be carried out, their contributions to the specific resolution of interest has to be computed. Nevertheless, I represent the imaging for the other submatrices in a similar way as equation (4.45), according to

$$\langle \mathbf{R}_{B(m)} \rangle = \frac{1}{N_\omega} \text{Re} \left(\sum_\omega \mathbf{R}_{B(m)} \right) \tag{4.46}$$

$$\langle \mathbf{R}_{C(m)} \rangle = \frac{1}{N_\omega} \text{Re} \left(\sum_\omega \mathbf{R}_{C(m)} \right) \tag{4.47}$$

$$\langle \mathbf{R}_{A(m)} \rangle = \frac{1}{N_\omega} \text{Re} \left(\sum_\omega \mathbf{R}_{A(m)} \right). \tag{4.48}$$

For example, for the contribution of $\langle \mathbf{R}_{B(2)} \rangle$ to the structural image at resolution $m = 1$, an inverse wavelet transform has to be carried out first and then the diagonals can be summed. From equations (4.41)-(4.48) it is clear that migration can be carried out at different lateral scales and for different parts of the reflectivity matrix independently. Preferably migration starts at a certain coarse scale M with the approximation $\langle \mathbf{R}_{T(M)} \rangle$. The structural image at that particular scale is given by the diagonal of $\langle \mathbf{R}_{T(M)} \rangle$ at all depth levels. If more detail in the resulting image is required, it can be added according to the process in Figure 4.7[14]. This process is an inverse 2-D wavelet transform which can be carried out efficiently. It should be noted that at a low resolution the matrices involved in the migration are much smaller than the original reflectivity matrix in the subspace $\boldsymbol{V}_0$. For each lower resolution step the size of the matrices reduces with a factor of four. Hence, the matrices involved in computing $\langle \mathbf{R}_{T(M-1)} \rangle$ are 4^M times as small.

[14]Since in the derivation of this procedure, the scale interaction quantities have been neglected one can not expect to be able to carry out the reconstruction process from very coarse scale all the way to very fine scales.

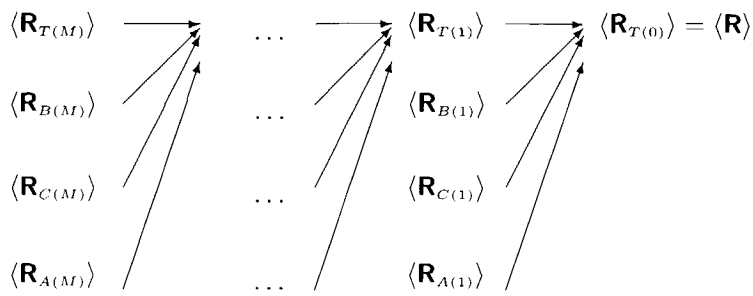


Fig. 4.7 *Schematic representation of a migration scheme in the wavelet domain, starting at a low resolution. Via an inverse wavelet transform the different parts of the reflectivity operator can be combined to the full reflectivity.*

4.5 Examples

This section serves as an illustration for the derivations presented in the previous sections. As examples I will use an "unrealistic" synthetic model and a marine data set. I will illustrate the method with two steps of the wavelet transform approach. The results of a step by step approach will be compared with a full prestack $(x-\omega)$-migration. Both data sets consist of a single 2-D line.

4.5.1 Synthetic data set

The model used for the synthetic experiment is given in Figure 4.8a. It is a constant density model, the compressional wave velocity ranges from $c = 1500\,\mathrm{m/s}$ up to $c = 3400\,\mathrm{m/s}$. 128 shot records have been created with a finite-difference program, second order in space and second order in time, in a fixed-spread configuration. The source and receiver positions were ranging from $x_1 = 0$ m up to $x_1 = 3175$ m with $\Delta x_1 = 25$ m. The sampling in the model for the finite-difference modeling was 5 m. The macro model for the recursive prestack depth migration is sampled with $\Delta x_1 = 25$ m and $\Delta x_3 = 10$ m. The shot records have been sampled with $\Delta t = 4$ ms; the total amount of samples per trace is 512. The seismic wavelet has its peak at a frequency of 20 Hz. The maximum frequency is 70 Hz. A number of shot records after surface related multiple elimination for source positions $x_1^s = 200$ m, $x_1^s = 1575$ m and $x_1^s = 2950$ m are shown in Figure 4.8b.

For the redatuming procedure a smoothed version of the exact model has been used. The full prestack migration result, carried out recursively in the space domain according to equations (4.31) and (4.34), is shown in Figure 4.8c. Migration in the wavelet domain has been carried out according to the scheme of Figure 4.7 where

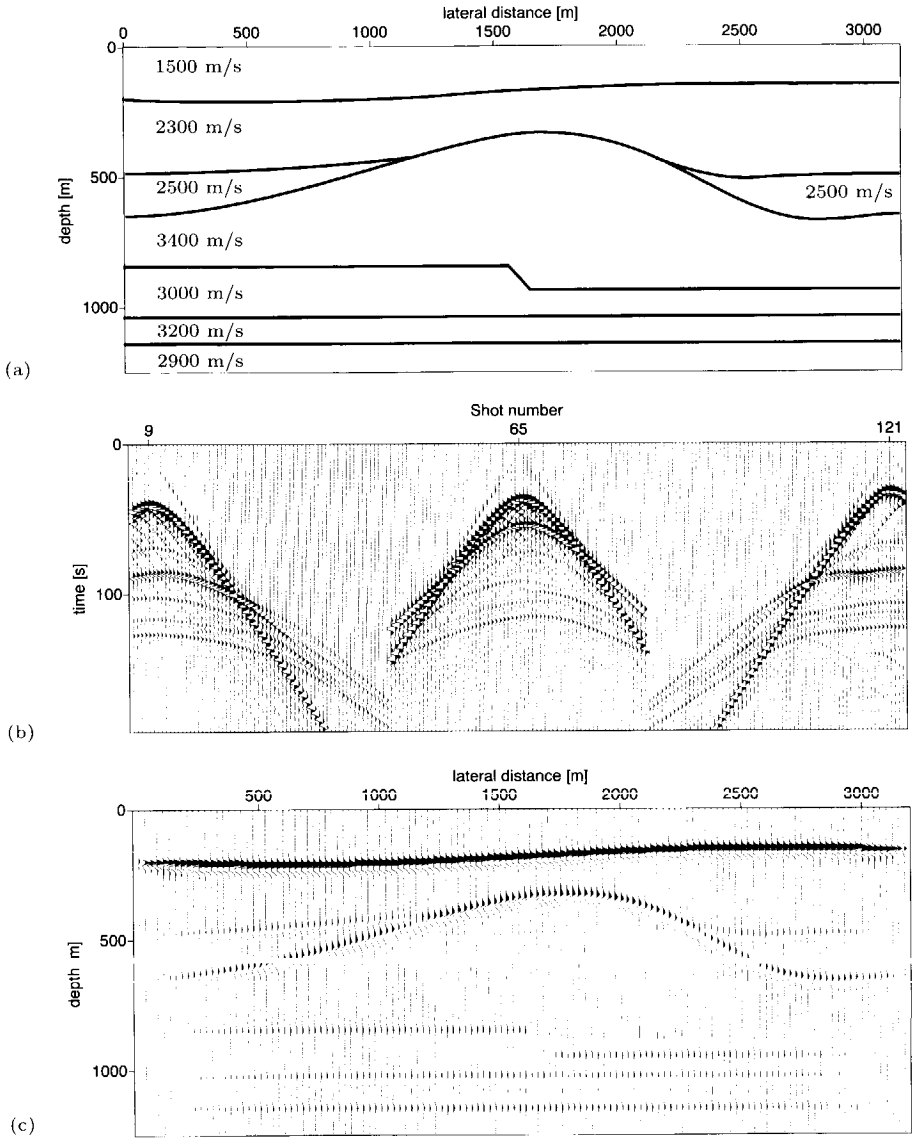


Fig. 4.8 *(a) Acoustical model with velocities ranging from 1500 m/s to 3400 m/s. (b) Three shot records at source positions* $x_1^s = 200\,m$, $x_1^s = 1575\,m$ *and* $x_1^s = 2950\,m$. *(c) The full prestack migration result. The migration has been carried out in the space domain.*

the submatrices are computed according to equations (4.41)-(4.48). The coarsest resolution used is $m = 3$. At this resolution the propagator matrices and the reflection matrices are 64 times as small as those at the original resolution $m = 0$. The zero-offset reflectivity resulting from the submatrices $\langle \mathbf{R}_{T(3)} \rangle$, $\langle \mathbf{R}_{B(3)} \rangle$, $\langle \mathbf{R}_{C(3)} \rangle$, and $\langle \mathbf{R}_{A(3)} \rangle$ at all depth levels is shown in Figure 4.9a-d. In Figure 4.9e, the sum of these results is depicted, which equals $\langle \mathbf{R}_{T(2)} \rangle$. Note that the summation of the submatrices improves the image of the dipping flanks. Grey circles denote these regions.

Alternatively, one can start the procedure at resolution $m = 2$. In this step the submatrices are 16 times as small as the matrices at the original resolution. The zero-offset reflectivity resulting from the submatrices $\langle \mathbf{R}_{T(2)} \rangle$, $\langle \mathbf{R}_{B(2)} \rangle$, $\langle \mathbf{R}_{C(2)} \rangle$, and $\langle \mathbf{R}_{A(2)} \rangle$ is shown in Figure 4.10a-d. The sum of the images yields the image related to $\langle \mathbf{R}_{T(1)} \rangle$ at all depth levels, which is shown in Figure 4.10e.

The final optional step can be used to further increase the lateral resolution of the subsurface image by computing the zero-offset reflectivity related to $\langle \mathbf{R}_{B(1)} \rangle$, $\langle \mathbf{R}_{C(1)} \rangle$, and $\langle \mathbf{R}_{A(1)} \rangle$, and adding these images to that of $\langle \mathbf{R}_{T(1)} \rangle$. For this example, these parts do no add much new information. Hence, I have omitted the results here.

Looking at the results of this section it is allowed to conclude that at least for this data set the proposed scheme works. The wavelet transform allows to start with a coarse approximation of the reflectivity. The image can be improved by adding the contributions to the structural image of the detail parts of the reflectivity matrix. The basic process up to the finest possible accuracy is two times as fast as the migration process in the space domain. The amount of computations related to the wavelet transform and the sparseness of the propagator matrices in the space domain are not accounted for. The efficiency increases if one is satisfied with a lower lateral resolution.

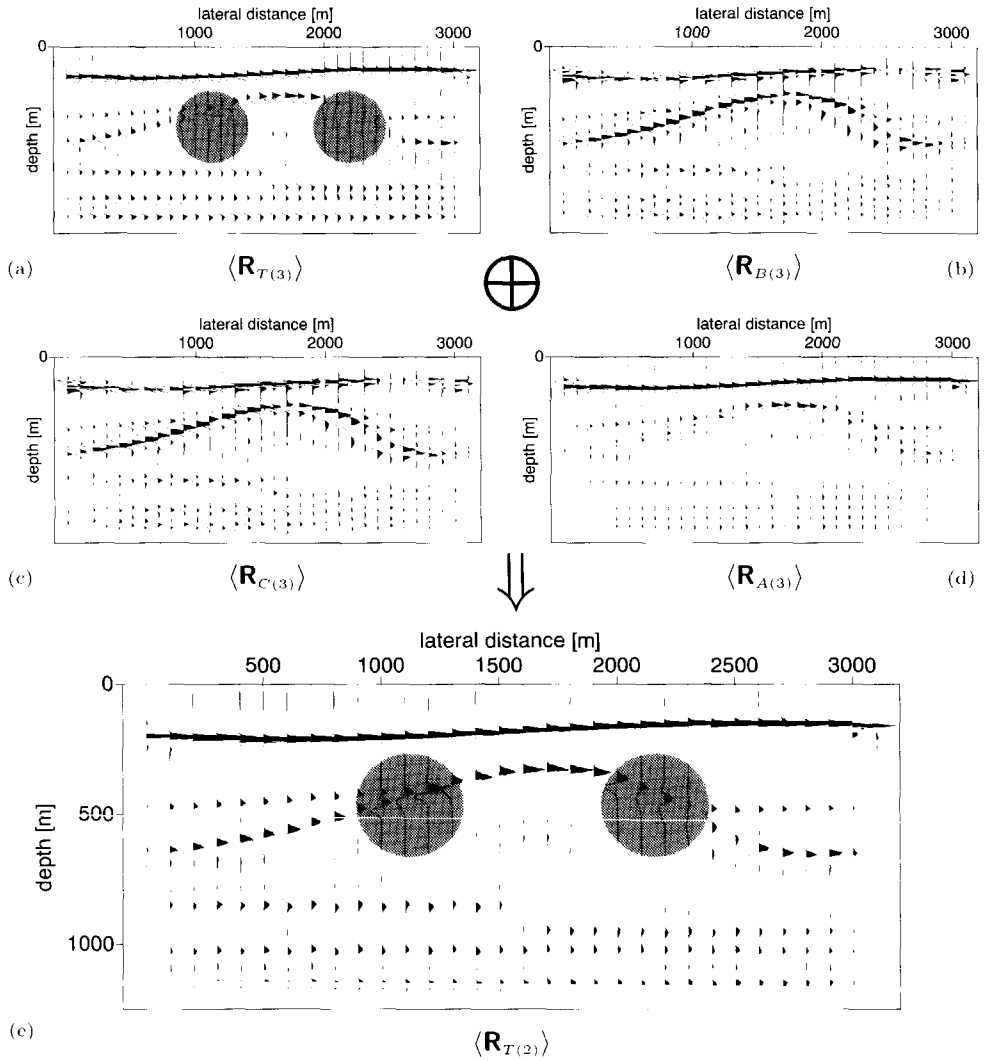


Fig. 4.9 *Zero-offset reflectivity for a full prestack migration in the wavelet domain. The migration is carried out at resolution $m = 3$ (the original data is at resolution $m = 0$), i.e. matrices are 64 times as small. Before the imaging step is carried out the submatrices are transformed back to resolution $m = 2$, in order to be able to compute their structural contribution at that resolution. (a) The migrated result according to equations* (4.41) *and* (4.45). *(b) The migrated result according to equations* (4.42) *and* (4.46). *(c) The migrated result according to equations* (4.43) *and* (4.47). *The migrated result according to equations* (4.44) *and* (4.48). *(e) The sum of the migrated results of (a)-(d). The grey circles denote regions, where the image particularly improves in going from resolution 3 to resolution 2.*

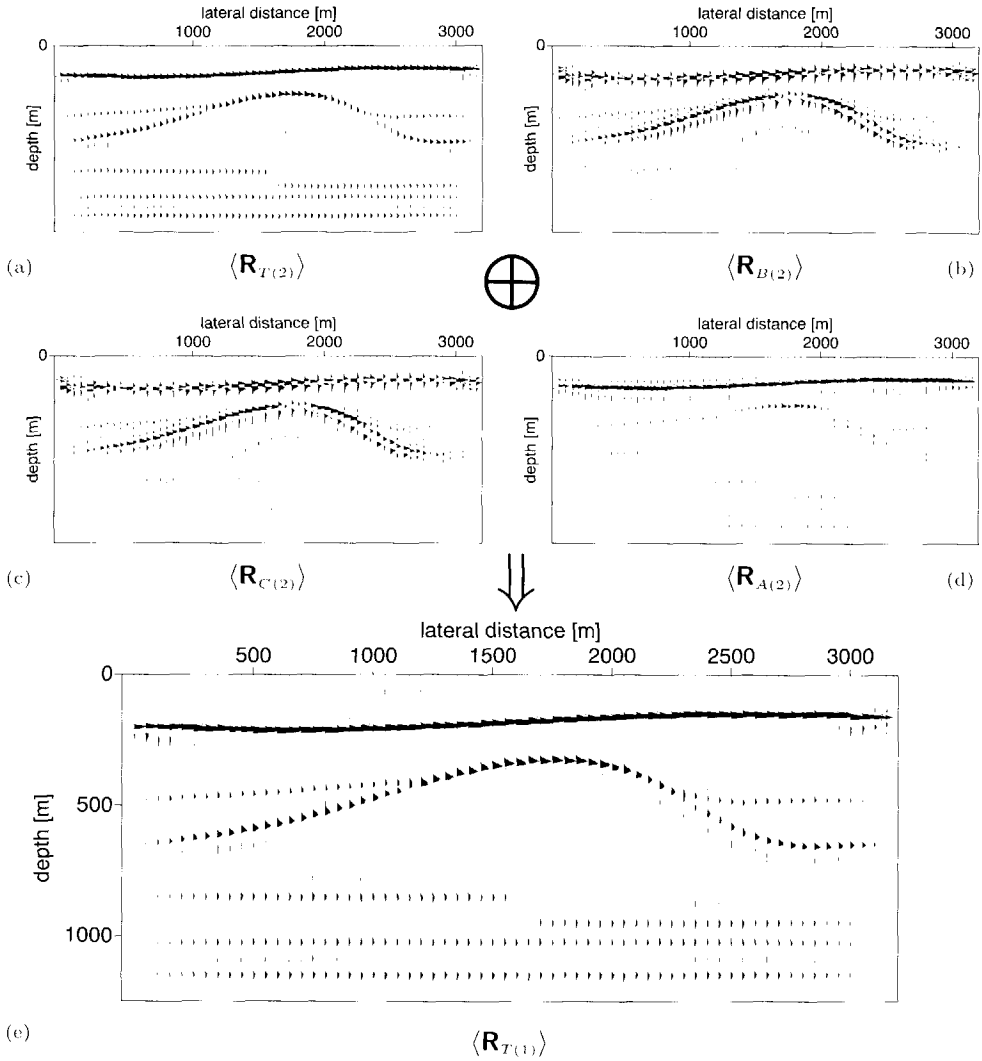


Fig. 4.10 *Zero-offset reflectivity for a full prestack migration in the wavelet domain. The migration is carried out at resolution $m = 2$, i.e. matrices are 16 times as small. The results are transformed back to resolution $m = 1$ before the imaging step is carried out. (a) The migrated result according to equations (4.41) and (4.45). (b) The migrated result according to equations (4.42) and (4.46). (c) The migrated result according to equations (4.43) and (4.47). The migrated result according to equations (4.44) and (4.48). (e) The sum of the migrated results of (a)-(d).*

4.5.2 Marine data set

The same procedure as described in the previous section has been applied to a marine data set on top of the Midgard field in the Haltenbanken area, offshore Norway (Data Courtesy SAGA). For a general description of the Midgard field the reader is referred to Ekern (1987). The main acquisition parameters are shown in Table 4.1. The near-offset data have been interpolated. Surface related multiples have been eliminated and the data have been deconvolved for the source wavelet. For the application under study, the data have been reordered in a number of fixed-spread data sets covering a part of the full line. A macro model has been estimated using the common focus point velocity estimation (Kabir, 1997). The estimated macro model for a part of the illuminated subsurface is shown in Figure 4.11a. In the model no steep dips are present. Hence, one can expect that already at coarse scales a relatively good estimation of the structure can be obtained.

The full prestack migration result carried out recursively according to the migration scheme of equations (4.31) and (4.34) is shown in Figure 4.11b. Migration in the wavelet domain has been carried out according to the scheme described in Figure 4.7. More specifically, I have followed the procedure which yielded Figure 4.9 to obtain a structural image for the SAGA data set at resolution $m = 2$. To create this image, matrices 64 times as small as the original size have been used. The result is shown in Figure 4.12a. The main structural information is already present. For example, the reflector at depth 2.5 km with the fault around the lateral distance of 16 km can be identified. On the other hand, the structure between depth 1.5 km and 2 km is not very clear. Their is not enough lateral coherency.

The same procedure has been carried out at resolution $m = 2$ just as has been done in the synthetic data example. This step yields the structural image at resolution $m = 1$. At this resolution the matrices are 16 times as small as at the finest resolution $m = 0$. The result is shown in Figure 4.12b. Upon comparison with the migration result at the finest resolution, which is shown again in Figure 4.12c

geometry	fixed spread
number of shots	387
shot spacing	25 m
number of detectors per shot	126
receiver spacing	25 m
recording time	4 s
time sampling	4 ms

Table 4.1 *The acquisition parameters for the marine data set.*

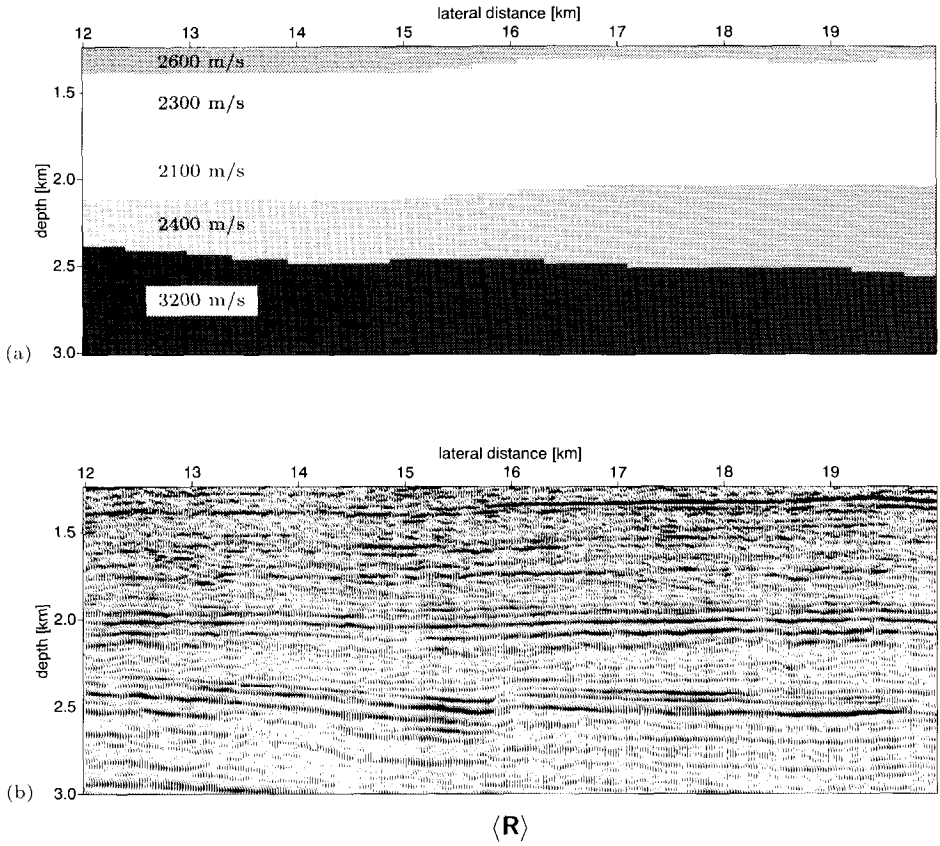


Fig. 4.11 *(a) A part of the estimated macro model for the marine data set (Data Courtesy SAGA). One of the target areas is at a depth larger than 2km at a lateral position of 17.7km. (b) The full prestack migration result of the same part. The migration has been carried out recursively in the space domain.*

it can be concluded that at this resolution all structural information seems to be present, which is quite remarkable if one keeps in mind the crude estimates of the propagator matrices that have been used. In practice one will stop at this resolution, which I did.

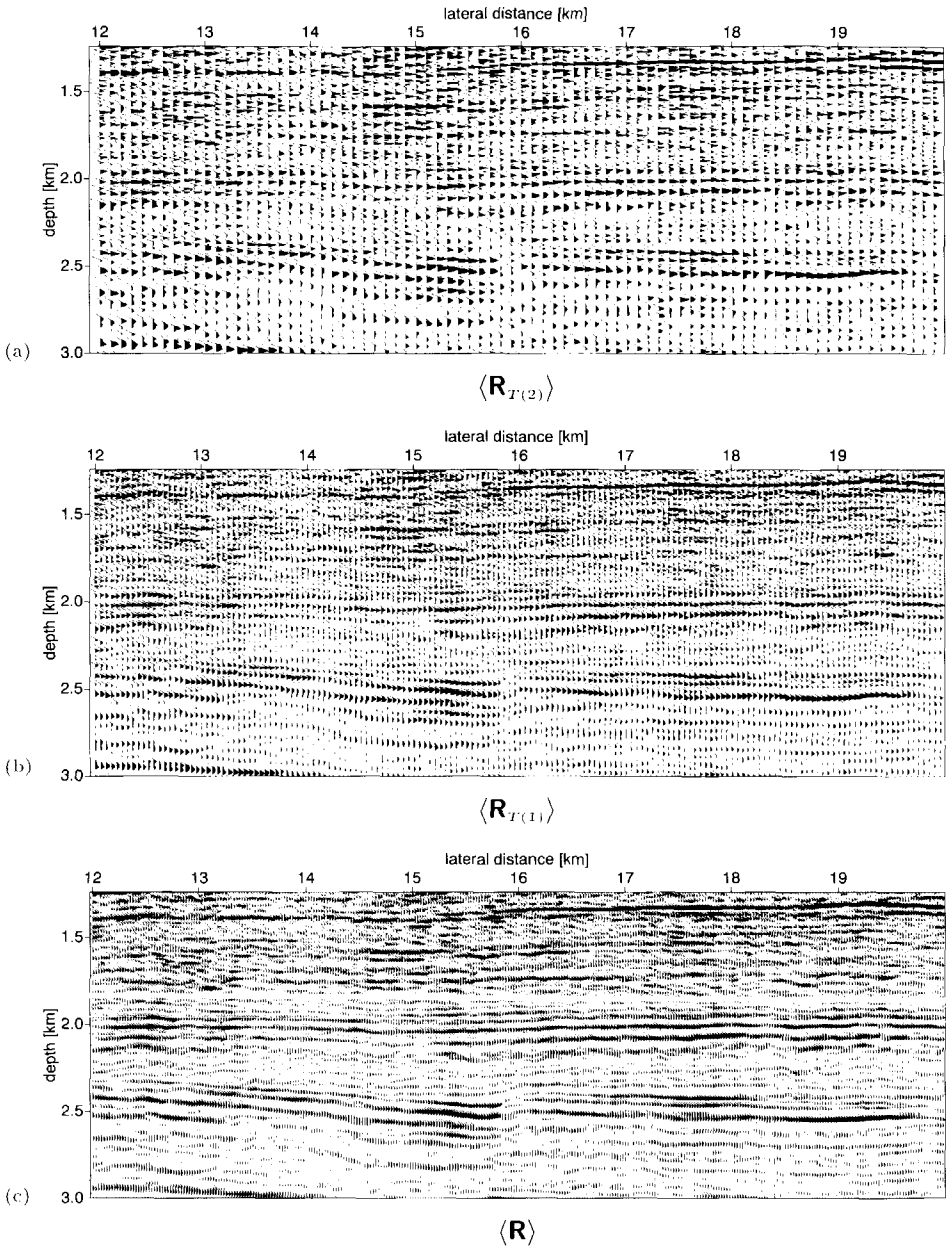


Fig. 4.12 *Structural migration results at three different lateral resolutions. (a) Structural image at resolution $m = 2$, obtained via equations (4.41)-(4.48), and carried out at resolution $m = 3$ in the wavelet domain. (b) Structural image at resolution $m = 1$, obtained via equations (4.41)-(4.48), and carried out at resolution $m = 2$ in the wavelet domain. (c) The full prestack migration result carried out in the space domain.*

4.6 Discussion

In this section I will discuss some open questions and drawbacks of the presented wavelet transform approach to migration and I will consider alternatives that I have come aware of.

Drawbacks and open questions —A first drawback is related to the fact that the propagators are not directly computed in the wavelet domain. The propagators are obtained in the wavelet domain by transforming their realizations in the space domain. Their are no efficient numerical schemes to directly compute the response of a wavelet or scaling function as source. This is of course due to the fact that there are no analytic solutions. Such solutions are known in homogeneous media for point sources, for plane waves and Gabor functions (the Gaussian beams). Analytic solutions can be used as a reference and often lead to efficient schemes.

A second drawback is related to the fact that the propagators are not rendered very sparsely in the wavelet domain. The oscillatory nature of the monochromatic Green's functions forbids real sparseness, as discussed in section 4.3, and hence a really efficient migration scheme. In my view it is not easy to overcome this drawback. A good way to, at least partially, deal with this problem is to compute only those coefficients that are above a certain threshold, as pointed out by several authors (Goswami et al., 1995; Wagner and Chew, 1995). To my knowledge such schemes have not been developed yet. Here, I followed another road to obtain efficiency. I neglected the off-diagonal parts of the recursive propagator matrices in the wavelet domain, which is a rather crude step. This step can be refined by looking for wavelets which minimize the scale interaction.

A third drawback of the presented method is related to the fact that the method is not gather-oriented. A whole data set divided into monochromatic components is used as input. Such an input requires quite some data reordering. Note here that a shot gather oriented approach has been developed as well (Dessing, 1995). This shot gather oriented approach allows for an imaging at multiple lateral scales, but it does not give a possibility for a recursive build up, as discussed in the present chapter.

A fourth, easily surmountable, drawback is related to the fact that periodic boundary conditions have been used in the wavelet transform. Since in general the two opposite ends of a subsurface geometry can be quite different, periodic boundary conditions are not correct. Alternatively, one can use orthonormal wavelet bases defined on an interval, see for example Cohen et al. (1993) and Goswami et al. (1995).

Alternatives Especially since people are aware of the great opportunities the wavelet transform provides for seismic data compression (see section 2.6), the interest in wavelet transform based migration algorithms has increased. I am aware of

two definite alternatives, which are both based on a wavelet transform applied to the depth or time coordinate, instead of an application to the lateral spatial coordinates presented here. Very interesting, Wang and Pann (1996) develop a Kirchhoff migration scheme for seismic data that are efficiently represented with the help of a matching pursuit decomposition scheme. The matching pursuit decomposition scheme represents a seismic trace as a relatively sparse sequence of reflection events. Especially, for velocity analysis the authors argue that their approach might be advantageous. Another approach is followed by Song et al. (1996). They discuss a multiscale version of the Kirchhoff based reflector imaging method (Bleistein, 1987).

4.7 Summary

In this chapter, I have presented a wavelet transform approach to seismic migration in the frequency domain. The point of departure is the representation of primary reflection data in arbitrary domains. The space domain representation leads to an interpretation of the seismic data in terms of point source experiments, the spatial Fourier domain representation leads to an interpretation in terms of plane wave experiments, the Gabor domain representation leads to an interpretation of the seismic data in terms of Gaussian beam experiments, and the wavelet transform leads to an interpretation of the seismic data in terms of multiscale experiments.

For migration, the application of the non-standard form of the 2-D wavelet transform provides a scheme that is in principle efficient. The migration scheme in the wavelet transform domain consists of similar steps as the migration scheme in the space-frequency domain, i.e. redatuming and imaging.

The asset of the wavelet transform for migration lies in the fact that the reflectivity can be extracted step by step if one neglects the scale interaction in the propagator matrices. The process starts with a coarse approximation of the reflectivity. By adding detail parts of the reflectivity operator the resolution can be improved. The step by step approach gives the user a handle to choose between resolution and efficiency. The applicability of the proposed method has been illustrated with a synthetic data set and a real marine data set.

In the discussion a number of drawbacks of the presented scheme has been pointed out. Some of them are easily surmountable, others less. Especially, the absence of analytic solutions or direct numerical methods in relation to the fact that the propagator is not rendered sparsely in the wavelet domain is seen as an important drawback.

Chapter 5

Boundary description by singularity characterization

The present thesis is completely written in the I-form. Especially, in this chapter the we-form would have been more appropriate. The current chapter can be seen as the cumulative effect of a long and intensive collaboration with former Ph.D. students and M.Sc. students, most notably, Felix Herrmann, Edo Hoekstra and Joes Staal. It deals with the effect of singularities on waves reflecting at those singularities. Besides an amplitude effect it is revealed that singularities –especially a-symmetric singularities cause a deterministic change of the signature of the seismic wavelet. The signature change can be extracted from the seismic data via a complex trace analysis. The extraction of the strength and signature of the singularity allows for the formulation of a singularity driven inversion scheme (SDI). *The results have been submitted for publication in two parts. In Hoekstra and Dessing (1998), the first half of this chapter is roughly covered. Dessing and Hoekstra (1998) deal with the second half of this chapter.*

5.1 Introduction

Reflection and propagation are the two processes underlying the seismic reflection method. Propagation is required to bring energy into the subsurface, reflection is required to revert the propagation direction of the energy transfer. The better a geophysicist understands the physical mechanisms underlying these processes, the more successful the seismic reflection method can be applied to infer the structure and the material properties of the subsurface. Whereas in the previous chapter the attention has been focused on the role of the wavelet transform in delineating the structure of the subsurface, in the present chapter multiscale analysis tools will be utilized to better understand and characterize the complex process of reflection.

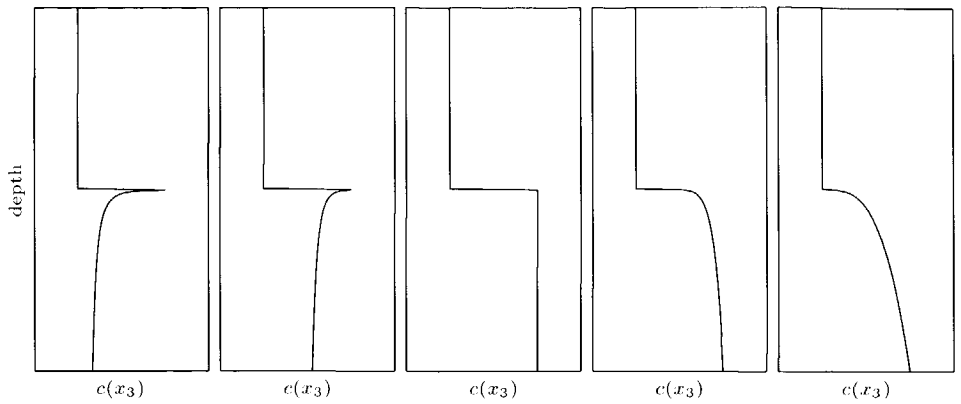


Fig. 5.1 *Five representations for a transition in the earth subsurface. The middle transition is the commonly considered step function. On the left and on the right 'sharper' and 'smoother' transitions are plotted. These transitions are a priori as likely to occur as the step function. The reflectivity of the latter four profiles turns out to be clearly distinctive from the reflectivity related to the step function.*

The point of departure is the work of Felix Herrmann (1997), who showed via a wavelet transform modulus maxima analysis of well-log data that the earth's subsurface behaves as a multifractal, which means that it consists of a hierarchy of singularities[1]. This observation has two important consequences. First, the description of a seismic reflector as a discontinuity, a step function, is just a special case of the more general concept singularity. Secondly, a singular medium necessarily implies that the measurement process depends on the scale of the measuring device (Nottale et al., 1997). The measuring or observing object is, for the geophysicist, an acoustic or elastic wave, which has interacted with a singular point in the subsurface.

Consequently, the thesis of Herrmann (1997) necessitates the introduction of more general transitions than step functions alone, as a model for the boundaries in the earth. Figure 5.1 shows four examples of the kind of generalization I am referring at. In the middle, the commonly used step function is shown, on the left and on the right, four different velocity or density profiles are displayed: the left-hand side shows two 'sharper' transitions; the right-hand side shows two 'smoother' transitions[2]. It can not be expected that the latter four profiles reflect an incoming wave in the same way as the step function does. What is the nature and size of the difference in the reflectivity? The geophysical relevance of this question stems

[1] A function $f(t)$ is singular at t_0 if its derivative is not bounded at t_0. The function $f(t)$ is said to have a singularity at t_0. In section 5.2 the concept singularity will be dealt with in more detail.

[2] The notions 'sharp' and 'smooth' will get a meaning in the sequel via the definition of the Hölder exponent.

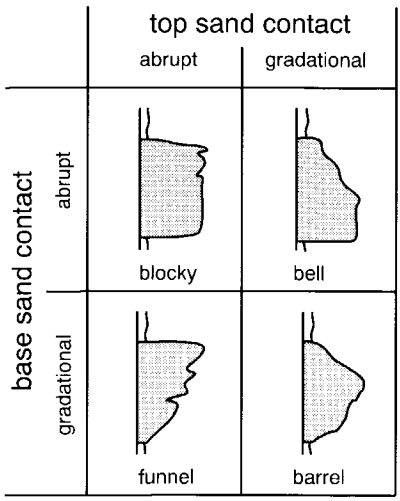


Fig. 5.2 *Four clearly distinctive geological depositional settings. Kindly provided by Frederic Verhelst.*

from the fact that geological boundaries do in general not behave as simple step functions. In Figure 5.2, schematic models of four depositional settings are given (provided by Frederic Verhelst). Note that different depositional environments yield clearly distinctive profiles.

The refinement of the model of the earth causes a further complication of the process of reflection. Let me briefly summarize the main complexities:

1. Since Snell in 1621, while he was professor in Leiden, empirically discovered his law of refraction, it is known that the reflection of a wave at a discontinuity or step function in the velocity is in general *angle dependent*. The actual angle-dependent behavior depends on the size and the direction of the transition.

2. A single singular transition either in the density or in the compressional wave speed is *scale dependent*. A discontinuity, however, does not result in a scale-dependent reflection behavior[3].

3. Another important process ruling the reflection process is *interference*. Two or more singular transitions closely to each other, result in a different reflection pattern than all of them apart. In the highly complex earth, which is probed by a wave with a pulse train the central wavelength of which lies in the seismic scale range, the interference process plays a significant role.

[3]In this chapter, another situation will be discussed which results in a scale-independent reflectivity as well.

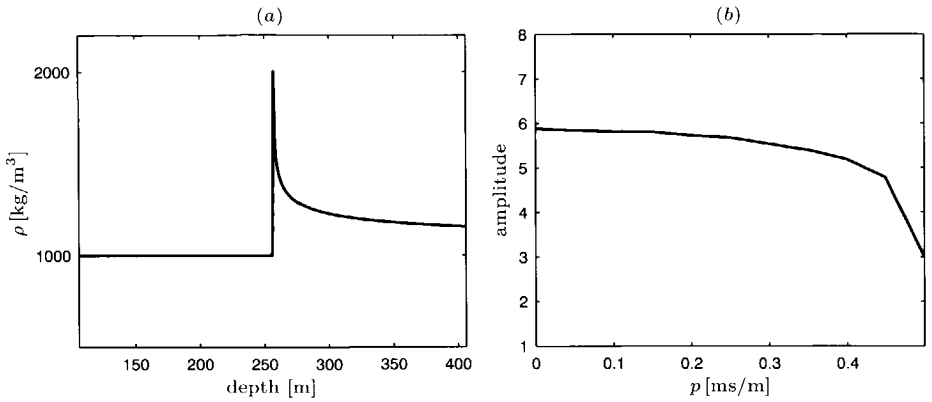


Fig. 5.3 *The amplitude of the reflectivity as a function of ray parameter, shows a clear angle-dependent behavior (b) for the density profile shown in (a). The velocity is kept constant. The ray parameter is related to the angle and the local velocity by* $p = \frac{\sin\beta}{c}$, *where* β *is the angle of incidence and* c *is the local velocity.*

In order to be able to derive the properties of geological boundaries in a deterministic way, understanding and disentangling the different effects are of great importance. Otherwise confusing situations can occur. An example of a confusing situation is shown in Figure 5.3. It is the response of a constant velocity medium with a varying density. For a constant velocity medium having only step functions as transitions in the density, the reflection does not depend on the angle of incidence. The transition of Figure 5.3 shows, however, an angle-dependent reflection behavior. This effect has to be attributed to the fact that a singularity is intrinsically scale dependent, and to the fact that the scale of a seismic wave increases with increasing angle of incidence. Herrmann (1995) made this observation, that is closely related to the effect of interference, and Wapenaar et al. (1997) proposed an angle-dependent imaging technique to separate angle dependency resulting from interference and intrinsic angle-dependent behavior. Another confusing situation can occur in the case that two quite different structures result in the same angle-dependent behavior. A situation like that is shown in Figure 5.4. One of the aforementioned effects might be the reason that the commonly used Zoeppritz model for angle-dependent reflection does not always work in the real world (Herrmann, 1997; van Wijngaarden, 1997), especially if a single inversion is carried out per single reflection event[4].

In this chapter, the attention is mainly focused on the reflection at singular structures in a laterally-invariant medium. I will not deal with the effects of interference. Whereas the importance of singularities has been recognized for a long time (see for example the quote from Maxwell on page 15), techniques to extract both

[4]Alternatively, for high quality data one can invert per sample, in which case it should be possible to recover a band-limited version of the medium.

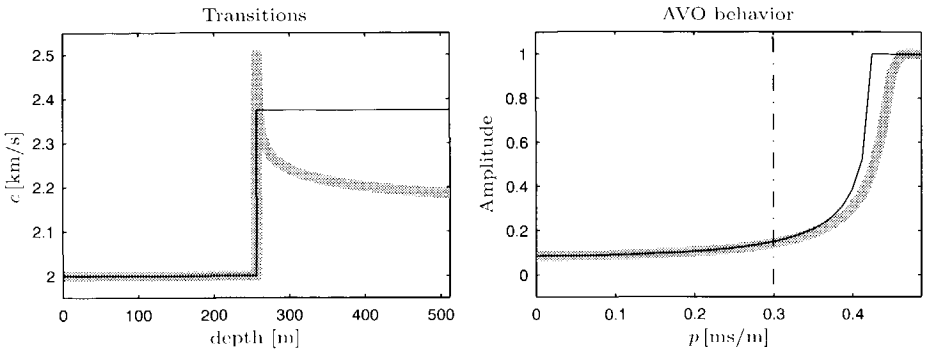


Fig. 5.4 *The angle-dependent reflectivity for a step function and for a singular structure (shown in grey) approximate each other very closely in the region where the linearized Zoeppritz model for angle-dependent reflectivity is assumed to be valid. This example will be discussed in more detail in section 5.6.*

the strength and the signature of a transition with the help of seismic reflection data are not well-developed yet. From several points of departure, it is tried to make life more deterministic. Most notably, Wapenaar (1997a) has taken the wave equation as a point of departure to derive the angle-dependent reflectivity for a specific type of singularity. In this chapter, another approach is taken. It is based upon the observation that homogeneous distributions, which are algebraic singularities of the form[5]

$$f(t) = t^{\alpha}, \qquad \alpha \subset \mathbb{R}, \quad t > 0, \tag{5.1}$$

carry out a fractional integration or differentiation. The fractional integration or differentiation manifests itself, amongst others, as a phase change in the reflected wave field[6]. The phase, which is associated with the travel time, is a stable element in seismic reflection data. It can be easily extracted with the help of a complex trace analysis (Gabor, 1946). The use of phase information has been recognized for a long time as a promising seismic attribute (Neidell and Taner, 1971; Taner et al., 1979; Robertson and Nogami, 1984; Berkhout, 1984). In this chapter, a theory is presented relating the induced phase change of the probing seismic wave field to the underlying transition in the subsurface in a clarifying way. It is expected that the results will help to better understand the pressing question how singularities in depth map to singularities in time.

The contents of the chapter is as follows. In section 5.2 the concepts regularity and singularity will be introduced. The role of the wavelet transform in character-

[5] A more refined definition follows in section 5.2.4.

[6] The phase of the transmitted wave field, however, is not directly effected.

izing the singularity strength of isolated and non-isolated singularities will be explained. Here, the results of Mallat and Hwang (1992), Staal (1995) and Herrmann (1997) will be utilized. At the end of section 5.2 it is argued that the singularity strength is far from sufficient to characterize algebraic singularities of the form (5.1). In section 5.3 the wavelet transform of those singularities will be treated in full detail. It will be clear that the phase and amplitude together completely characterize a one-sided singularity. A complex trace analysis is introduced to extract the phase information. In section 5.4 the influence of a seismic wavelet will be analyzed. Whereas in sections 5.2 and 5.3 the attention is focused on characterizing singularities themselves, in section 5.5 the attention is tuned to the reflectivity related to singular medium profiles for normal-incident waves. Both analytic and numerical derivations will be presented. The reflectivity at singularities for oblique-incident waves is discussed in section 5.6. In section 5.7 a singularity driven inversion (SDI) scheme is proposed. In section 5.8 I will briefly discuss what has been accomplished.

5.2 Local regularity and the wavelet transform

Local points of rapid variation can be intuitively associated with reflection. A wave will be reflected if the change in the medium parameters is fast with respect to the scale of the propagating wave[7]. In the current models used in seismic processing, the points of rapid variation are represented by step functions. Herrmann (1997), however, showed that the earth subsurface does not consist of step functions alone. It merely consists of a hierarchy of singularities, of which the step function is just a special case. This section is devoted to a brief introduction of the concept singularity and to the way the wavelet transform is utilized to measure the strength of a singularity. The mathematical theory will be briefly dealt with; for a more extensive treatment of this subject the reader is referred to Mallat and Hwang (1992) or Herrmann (1997). For a more general treatment of the wavelet transform the reader is referred to chapter 2.

5.2.1 Local and global regularity

Let me start with a definition to precisely introduce the concepts singularity and regularity for positive Hölder exponents.

Definition 5.1: Local and global regularity with positive Hölder exponents (Mallat and Hwang,1992)

- *Let n be a positive integer and $n \leq \alpha \leq n+1$. A function $f(t)$ is said to have a* ***local*** *Hölder exponent α, at t_0, if and only if there exist two constants A*

[7]Of course, as soon as the medium parameters are changing in the x_3-direction some energy is reflected. However a rapid change with respect to the size of the probing wave field is required to reflect a significant amount of energy in a coherent way. See also the remarks on page 68.

and $h_0 > 0$, *and a polynomial of order* n, $P_n(t)$, *such that for* $h < h_0$

$$|f(t_0 + h) - P_n(h)| \leq A|h|^{\alpha}. \tag{5.2}$$

- *A function* $f(t)$ *has a* **global** *Hölder exponent* α *over the interval* (a_1, a_2), *if and only if equation* (5.2) *is satisfied for* $t_0 + h \in (a_1, a_2)$.
- *The* n*th derivative of a function* $f(t)$ *is singular at* t_0 *if* $f(t)$ *has a Hölder exponent* $n < \alpha < n+1$. *The* n*th derivative is said to have a singularity at* t_0.
- *The Hölder regularity of* $f(t)$ *at* t_0 *is the superior bound of all values* α *such that* $f(t)$ *has a Hölder exponent* α *at* t_0.

Note that in this definition a clear distinction is made between local and global regularity. Local regularity is defined in a point, whereas global regularity is defined in an interval. The relevance of this distinction lies in the fact that for propagation global regularity is important, and for reflection local regularity, because propagation can be seen as a global averaging process, and reflection as a local averaging process (Herrmann, 1997). As I want to focus the attention on boundary characterization, the local regularity is here of greater interest.

The above definition is only valid for positive Hölder exponents α. For negative values of α a tempered distribution[8] is on the right-hand side of equation (5.2). For tempered distributions, a separate definition of the Hölder exponent is required.

Definition 5.2: Negative global Hölder exponents (Mallat and Hwang,1992)
Let $f(t)$ *be a tempered distribution of finite order. Let* α *be a non-integer real number and* $[a_1, a_2]$ *an interval of* $\mathbb{R}$. *The distribution* $f(t)$ *is said to have a global Hölder exponent* α *on* (a_1, a_2), *if and only if its primitive has a global Hölder exponent* $\alpha + 1$ *on* (a_1, a_2).

The proper incorporation of negative Hölder exponents is important, because they can be naturally associated with sudden outbursts in the measured phenomena, and they are frequently encountered in practice, for example in well-log measurements. Definition 5.2 applies to global negative Hölder exponents. It does not allow to draw conclusions on the local regularity. Local negative Hölder exponents are dealt with in the microlocalization theory described by Bony (1983) and Jaffard (1991). In the case of an isolated negative Hölder exponent it is possible to assign locally negative Hölder exponents in the sense of Definition 5.1[9].

[8] For a concise and clear treatment of tempered distributions, introduced by Schwartz (1951, 1952), see Zemanian (1965). Functions form a subclass of the set of distributions.

[9] A distribution $f(t)$ is said to have an isolated Hölder exponent α at t_0, if and only if $f(t)$ has a global Hölder exponent α over an interval (a_1, a_2), with $t_0 \in (a_1, a_2)$, and if $f(t)$ has a global Hölder exponent $\beta \geq 1$ over any subinterval of (a_1, a_2) that does not include t_0 (Mallat and Hwang, 1992).

The following three examples illustrate the concepts described in Definitions 5.1 and 5.2.

Example 5.1 *Consider the function $f_1(t)$ given by*

$$f_1(t) = \begin{cases} 0 & \text{for} \quad t \leq 0 \\ t & \text{for} \quad t > 0. \end{cases} \tag{5.3}$$

The derivative of $f_1(t)$ is the Heaviside or step function. Since the step function is bounded, $f_1(t)$ is not singular according to Definition 5.1. It has a Hölder exponent $\alpha = 1$ at $t = 0$. The step function is singular and has a Hölder exponent $\alpha = 0$ at $t = 0$.

Example 5.2 *Consider the function $f_2(t)$ given by*

$$f_2(t) = \begin{cases} 0 & \text{for} \quad t \leq 0 \\ t^\beta & \text{for} \quad t > 0, \end{cases} \tag{5.4}$$

with $1 < \beta < 2$. This function has a local Hölder exponent of $\alpha = \beta$ at $t = 0$, but is not singular. Its derivative has a local Hölder exponent $\alpha = \beta - 1$ at $t = 0$; it is singular at $t = 0$. The second derivative of $f_2(t)$ is not bounded and can only be defined in the sense of distributions. Considered as a tempered distribution

$$f_3(t) = \partial_t^2 f_2(t) = \begin{cases} 0 & \text{for} \quad t \leq 0 \\ \beta(\beta-1)t^{\beta-2} & \text{for} \quad t > 0 \end{cases} \tag{5.5}$$

can be shown to have an isolated, hence local, negative Hölder exponent of $\alpha = \beta - 2$. Hence $f_3(t)$ is singular at $t = 0$.

Example 5.3 *Consider the δ-distribution, $\delta(t)$. The second primitive of the δ-distribution is a function which is piece-wise linear, i.e the function described in the first example above. Because of Definition 5.2 it can be concluded that the δ-distribution has a global Hölder exponent of $\alpha < -1$. To conclude that the local Hölder exponent $\alpha = -1$ at $t = 0$, the fourth item of the first definition is required, and the elements in the discussion after Definition 5.2 should be used.*

5.2.2 Measuring Hölder exponents

A well-known tool to measure the Hölder regularity of a tempered distribution $f(t)$ is the Fourier transform. From the discussion on page 43 in chapter 2, it is clear that regularity in one domain corresponds to decay properties in the associated Fourier domain. It can be shown that the Hölder exponent is the superior bound of all values α for which the Fourier transform $\hat{f}(\omega)$ of $f(t)$ satisfies

$$\int \hat{f}(\omega)(1 + |\omega|)^\alpha \, \mathrm{d}\omega < +\infty. \tag{5.6}$$

Unfortunately, the estimates provided via the Fourier transform are global. It is not allowed to draw conclusions on the local regularity properties, because one cannot say whether an analyzed function is locally more regular at a particular point. For a local regularity analysis the wavelet transform has turned out to be well-equipped. Let me recall the continuous wavelet transform of $f(t)$ with respect to a wavelet $\psi(t)$, see chapter 2 for more details,

$$\mathfrak{W}\{f,\psi\}(\sigma,b) = \frac{1}{\sigma^{1/\mathrm{p}}}\int f(t)\psi^*\left(\frac{t-b}{\sigma}\right)\mathrm{d}t, \tag{5.7}$$

where p refers to the normalization of the wavelet. In this chapter I will work with $\mathrm{p} = 1$, which means that an L^1-normalization will be utilized[10]. Knowing that the scale parameter σ is inversely proportional to the frequency parameter ω (see section 2.3.3), it can be understood that the wavelet transform can serve as a tool for measuring local regularity properties.

The relationship between the local regularity at a point $t = t_0$ and the decay properties of the wavelet transform of the function or distribution is expressed in the following theorem, which is taken from Holschneider and Tchamitchian (1990), and which was proved by Jaffard (1989)[11].

Theorem 5.1: Local regularity and the wavelet transform
Suppose that $\psi(t)$ is a wavelet with n vanishing moments and that $\psi(t)$ is n times continuously differentiable. The wavelet $\psi(t)$ is, moreover, assumed to be of compact support.

(i) Let $f(t) \in L^2(\mathbb{R})$ and let $\alpha \leq n$. If $f(t)$ has a local Hölder exponent α at t_0, then there exists a positive constant A such that for $t_0 + c$ in the neighborhood of t_0, and for any σ

$$|\mathfrak{W}\{f,\psi\}(\sigma,t_0+c)| \leq A(\sigma^\alpha + |c|^\alpha). \tag{5.8}$$

(ii) Let $\alpha < n$ be a non-integer value and let A and B be two positive constants. The function $f(t)$ is said to have a local Hölder exponent α at t_0, if for some $\epsilon > 0$ the following two conditions hold for $t_0 + c$ in the neighborhood of t_0 and any scale σ

$$|\mathfrak{W}\{f,\psi\}(\sigma,t_0+c)| \leq A\sigma^\epsilon \tag{5.9}$$

and

$$|\mathfrak{W}\{f,\psi\}(\sigma,t_0+c)| \leq B(\sigma^\alpha + \frac{|c|^\alpha}{|\log|c||}). \tag{5.10}$$

[10] An L^2-normalization can be used equally well (Daubechies, 1992; Kaiser, 1994).

[11] Simpler theorems for the relation between global regularity in an interval and the wavelet transform can be found in the cited references.

The proof of the first part of Theorem 5.1 can be easily obtained, by writing out the wavelet transform around t_0, and by using equation (5.2). The second part of the proof is slightly more involved and I will not discuss it here. What does Theorem 5.1 mean? Equation (5.9) denotes that the distribution $f(t)$ has a global Hölder exponent ϵ in a neighborhood of t_0. For the interpretation of (5.8) and (5.9), the definition of the cone of points (σ, b) in the scale-time space according to

$$|b - t_0| \leq \sigma \quad \Longrightarrow \quad |c| \leq \sigma \tag{5.11}$$

is required. For points within the cone, the right-hand sides of equations (5.8) and (5.10) is dominated by the behavior of σ^{α}. For $|c| > \sigma$, the decay behavior is dominated by c. Equations (5.8) and (5.10) behave differently for these regions. However, for isolated singularities it suffices to inspect the decay behavior of the wavelet transform within the cone[12].

The requirement of enough vanishing moments and of smoothness for $\psi(t)$ has been discussed on page 43 in chapter 2. There, I have described, that it is possible, via partial integration, to transfer derivatives or integrations from the distribution to be analyzed, to the analyzing wavelet such that the resulting Hölder exponent is in the range $(-1, 0)$. The smoothness and the number of vanishing moments guarantee that the action of partial integration or differentiation is well-behaving.

Theorem 5.1 shows that estimates of the local regularity of a distribution can be obtained with the wavelet transform. It does, however, not provide an effective algorithm, because a measurement in the two dimensional (σ, b)-plane in the neighborhood of $t = t_0$ is required to estimate the local Hölder exponent at t_0. Mallat and Hwang (1992) proposed an effective partitioning of the (σ, b)-plane via the local maxima of the wavelet transform. The amplitudes along a connected set of local maxima, forming a wavelet transform modulus maxima line, characterize the strength of the local singularity.

5.2.3 Wavelet Transform Modulus Maxima Lines

From chapter 2, Figure 2.8, it can be concluded that the wavelet transform reaches a local maximum at or near a point of rapid variation, i.e. a locally irregular point. In the previous subsection, it has been shown that the decay behavior of the wavelet transform in the (σ, b)-plane is related to the regularity of the function. These two elements made Mallat and Hwang (1992) decide to consider the modulus maxima of the wavelet transform for regularity estimates.

[12]For rapidly oscillating functions, for example $\sin(1/t)$ near $t = 0$, the decay behavior within the cone cannot be utilized. Mallat and Hwang (1992) and Arneodo et al. (1997) discuss the detection of oscillating singularities. I will not deal with oscillating singularities.

Definition 5.3: Wavelet Transform Modulus Maxima (WTMM)

- *A local extremum of* $\mathfrak{W}\{f,\psi\}(\sigma_0,b)$ *is that point* (σ_0,t_0) *such that its partial derivative* $\partial\mathfrak{W}\{f,\psi\}(\sigma_0,b)/\partial b$ *has a zero-crossing at* $b=t_0$*, when* b *varies.*
- *A modulus maximum is any point* (σ_0,t_0) *such that the following inequality* $|\mathfrak{W}\{f,\psi\}(\sigma_0,b)| < |\mathfrak{W}\{f,\psi\}(\sigma_0,t_0)|$ *holds, when* b *belongs to either the right or the left neighborhood of* t_0*, and* $|\mathfrak{W}\{f,\psi\}(\sigma_0,b)| \leq |\mathfrak{W}\{f,\psi\}(\sigma_0,t_0)|$ *when* b *belongs to the other neighborhood of* t_0*. A modulus maximum is referred to as Wavelet Transform Modulus Maximum or* WTMM.
- *A Wavelet Transform Modulus Maxima Line or* WTMML *is any connected curve in the* (σ,b)*-plane along which all points are* WTMMs.

Figure 5.5 illustrates the concept of the WTMMLs. In Figure 5.5a a function is shown with three local singularities. The wavelet transform of the functions is given in Figure 5.5b. Figure 5.5c shows on top of the wavelet transform the connected WTMMs forming four WTMMLs in total. Let me further elucidate the relation between a WTMM and a singularity: a function is not singular in a neighborhood where its wavelet transform does not have modulus maxima at fine scales (Mallat and Hwang, 1992, Theorem 3). The reverse, however, is not always true: a modulus maximum in the wavelet transform of a function does not necessarily imply that the function or one of its derivatives is singular. For example, the wavelet transform of $\sin t$ certainly has WTMMs, although it is infinitely many times continuously differentiable. The following theorem relates the decay behavior along a WTMML to the local Hölder exponent in the case that the function does not oscillate in the neighborhood of the singularity. This theorem has been proven by Mallat and Hwang (1992).

Theorem 5.2: Isolated singularities and WTMMLs (Mallat and Hwang, 1992)
Suppose that the wavelet $\psi(t)$ *is* n *times continuously differentiable, and that it is the* n*th derivative of a smoothing function (hence, it has* n *vanishing moments), and it is of compact support. Let* $f(t)$ *be a distribution whose wavelet transform is well defined over* (a_1,a_2) *and let* $t_0 \in (a_1,a_2)$*. It is assumed that there exists a scale* $\sigma_0 > 0$ *and a constant* K*, such that for* $b \in (a_1,a_2)$ *and* $\sigma < \sigma_0$*, all the modulus maxima of* $\mathfrak{W}\{f,\psi\}(\sigma,b)$ *belong to a cone defined by*

$$|b-t_0| \leq K\sigma. \tag{5.12}$$

Then at all points $t_1 \in (a_1,a_2)$*,* $t_1 \neq t_0$*,* $f(t)$ *has a global Hölder exponent* n *in a neighborhood of* t_1*. Let* $\alpha < n$ *be a non-integer. The function* $f(t)$ *has a local Hölder exponent* α *at* t_0*, if and only if there exists a constant* C *such that at each modulus maximum* (σ,b) *in the cone defined by equation* (5.12)

$$|\mathfrak{W}\{f,\psi\}(\sigma,b)| \leq C\sigma^{\alpha}. \tag{5.13}$$

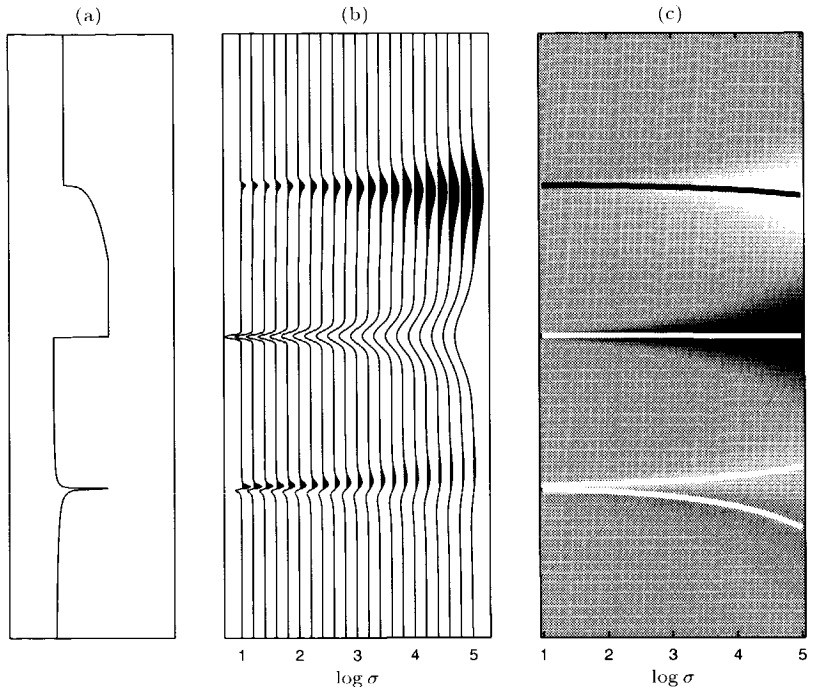


Fig. 5.5 *The wavelet transform of a signal and the extracted* WTMMLs. *(a) The function contains three singular points. (b) The wavelet transform of the function shown in (a) has only significant amplitudes at points where the function is changing rapidly. (c) The extracted* WTMMLs *on top of the same wavelet transform, now depicted in grey scale colors.*

The latter inequality can also be written as

$$\log |\mathfrak{W}\{f, \psi\}(\sigma, b)| \leq \log C + \alpha \log \sigma, \tag{5.14}$$

which is a linear relationship between the natural logarithm of the amplitude of the wavelet transform and the natural logarithm of the scale parameter σ. Theorem 5.2 proofs that the local Hölder exponent can be estimated by taking the maximum slope of straight lines that remain above $\log |\mathfrak{W}\{f, \psi\}(\sigma, b)|$. I will refer to the set of all modulus maxima lines of $f(t)$ fulfilling the cone distribution as

$$\text{WTMMR} = \{X_i(\sigma)\}_{i \in J}, \tag{5.15}$$

with $J \subseteq \mathbb{N}$, as the Wavelet Transform Modulus Maxima Representation WTMMR. Each element $X_i(\sigma)$ of the set WTMMR is a WTMML. Following Staal (1995) and Herrmann (1997), a wavelet transform modulus maxima analysis consists of the following steps:

Procedure 5.1: Wavelet Transform Modulus Maxima Analysis

1. *Compute the continuous wavelet transform of f with an appropriately chosen wavelet, i.e. having enough vanishing moments and being enough times continuously differentiable for the selected goal.*

2. *Find the* WTMMs *for each scale.*

3. *Interconnect the* WTMMs *such that the cone distribution is fulfilled, thus yielding the set* WTMMR *of* WTMMLs.

4. *Study the behavior of $\mathfrak{W}\{f,\psi\}(\sigma, X_i(\sigma))$ as a function of σ. In particular, find, via a linear regression, the maximum slope of the straight lines that remain above the logarithm of the amplitude of the modulus maxima line, on a logarithmic scale.*

Remark 5.1: Terminology
The local Hölder exponent is also referred to as the local scaling exponent or the strength α of the singularity. The strength of a singularity should not be confused with the offset log C *in equation (5.14).*

There are a few practical and theoretical aspects relevant for the further application of the described algorithm:

1. The number of vanishing moments of the analyzing wavelet determines the maximum Hölder exponent that can be detected (see also the discussion following Theorem 5.1). Hence, if the detected slope yields a Hölder exponent equal to the number of vanishing moments m than the result tells us that the analyzed function is at least m times differentiable (Bacry et al., 1993).

2. The number of vanishing moments should be kept as small as possible. More vanishing moments yield more WTMMLs for the same analyzed function. The algorithm will consequently be less efficient.

3. The cone distribution of equation (5.12) implies that we are dealing with isolated singularities. For non-oscillating functions or distributions the theory can be extended relatively easy to non-isolated singularities (Mallat and Hwang, 1992). In this situation it may be better to talk about an effective (and possibly scale-dependent) singularity strength (Herrmann, 1997).

4. Although counter examples can be formulated, the set WTMMR of all WTMMLs is generally complete (Mallat and Hwang, 1992). The function f can be reconstructed up to the finest scale where WTMMs are available, under the condition that a coarse approximation of the function is kept in memory, in addition to the set WTMMR.

5.2.4 Homogeneous distributions

The introduction of positive Hölder exponents in Definition 5.1 and its extension to negative Hölder exponents in Definition 5.2 suggests that a special role is played by functions or distributions of the form t^α as parameterizing objects. The wavelet transform of these objects should exactly fulfill the decay behavior derived in sections 5.2.2 and 5.2.3. In practice, a signal will never exactly behave as such an object, but a good understanding of the behavior of the idealized objects under the wavelet transform might improve the interpretation of the wavelet transform of more general distributions.

Consider the homogeneous distributions

$$\chi_\pm^\alpha(t) = \frac{|t|_\pm^\alpha}{\Gamma(\alpha+1)}, \tag{5.16}$$

where[13]

$$|t|_+^\alpha = \begin{cases} 0 & t \leq 0 \\ t^\alpha & t > 0 \end{cases} \qquad |t|_-^\alpha = \begin{cases} (-t)^\alpha & t < 0 \\ 0 & t \geq 0. \end{cases} \tag{5.17}$$

An overview of some important properties of homogeneous distributions is given in appendix B. Here the most relevant aspects are discussed. The distributions $\chi_\pm^\alpha(t)$ are well defined if they are set to work on highly localized and regular functions. More specifically, they are set to work on the set of wavelets $\psi(t)$, that is rapidly decreasing in one domain and arbitrarily polynomial localized in the associated Fourier domain. If the Fourier transform of $\psi(t)$, $\hat{\psi}(\omega)$, is only supported for $\omega > 0$, then this set is denoted by $S_+(\mathbb{R})$. On the other hand, if the Fourier transform $\hat{\psi}(\omega)$ of $\psi(t)$ is only supported for $\omega < 0$, then this set is denoted by $S_-(\mathbb{R})$. Wavelets that are only supported for $\omega > 0$ are called progressive wavelets. Wavelets whose Fourier transform is only supported for $\omega < 0$ are referred to as regressive wavelets (Holschneider, 1995). Progressive and regressive wavelets are complex. The direct sum of $S_+(\mathbb{R})$ and $S_-(\mathbb{R})$ is denoted by $S_0(\mathbb{R})$. The Morlet wavelet (Morlet et al., 1982) given by

$$\psi_p(t) = e^{-t^2/2} e^{j\omega_0 t}, \tag{5.18}$$

with $\omega_0 > 5$, is, for example, a wavelet in $S_+(\mathbb{R})$ [14]. The subscript p in $\psi_p(t)$ refers to the fact that the Morlet wavelet is a progressive wavelet. First or higher order derivatives of the Gaussian are wavelets in $S_0(\mathbb{R})$.

[13] The introduction of the Euler-Gamma function $\Gamma(\alpha+1)$ is nicely explained by Hörmander (1983, Section 3.2). It takes care of the simple poles in $|t|_\pm^\alpha$ for $\alpha = -1, -2, \ldots$.

[14] I have chosen the Morlet wavelet as an example, because it is important from a historical perspective, and because it is very easy to use. Strictly speaking, it is not a progressive wavelet due to the fact that it does not completely vanish for $\omega < 0$. In practice, however, it does vanish.

Although the distributions $\chi_{\pm}^{\alpha}(t)$ are well defined for $\alpha \in \mathbb{C}$, I will only consider $\alpha \in \mathbb{R}$[15]. The interpretation of the homogeneous distributions can be most easily carried out by first discussing a specific subset. Consider $\chi_{+}^{\alpha}(t)$ with $\alpha \in \mathbb{R}$ and $\alpha \notin \mathbb{Z}$[16]. The distributions $\chi_{+}^{\alpha}(t)$ are invariant under dilations, except for an amplitude factor, i.e.

$$\chi_{+}^{\alpha}(\lambda t) = \lambda^{\alpha} \chi_{+}^{\alpha}(t), \qquad \lambda > 0. \tag{5.19}$$

The exponent α is the (local) scaling exponent. The wavelet transform of a homogeneous distribution can be easily derived to be (using equation (5.7), with $\mathrm{p} = 1$)

$$\begin{aligned} \mathfrak{W}\{\chi_{+}^{\alpha}, \psi\}(\sigma, b) &= \frac{1}{\sigma} \int_{0}^{\infty} \chi_{+}^{\alpha}(t)\, \psi^{*}\left(\frac{t-b}{\sigma}\right) \mathrm{d}t \\ &= \sigma^{\alpha} \int_{0}^{\infty} \frac{t^{\alpha}}{\Gamma(\alpha+1)} \psi^{*}(t - b/\sigma)\, \mathrm{d}t \\ &= \sigma^{\alpha} U_{+}(b/\sigma), \end{aligned} \tag{5.20}$$

where

$$U_{+}(u) = \int_{0}^{\infty} \frac{t^{\alpha}}{\Gamma(\alpha+1)} \psi^{*}(t-u)\, \mathrm{d}t. \tag{5.21}$$

It is interesting to see that $U_{+}(u)$ can also be interpreted as the Mellin transform of the wavelet ψ. From equation (5.20) it is clear that along a line of b/σ is constant the amplitudes of the wavelet transform is only varying as a function of the scale parameter σ, according to the factor σ^{α}. It can be seen that WTMMLs of homogeneous distributions fulfill the prescription b/σ constant. Hence, for (σ, b) chosen such that they form a WTMML of a homogeneous distribution it is found that

$$\log |\mathfrak{W}\{\chi_{+}^{\alpha}, \psi\}(\sigma, b)| = \log |U(b/\sigma)| + \alpha \log \sigma, \tag{5.22}$$

which is relationship (5.14) derived in the previous subsection.

Another interesting subclass is obtained by choosing $\alpha = n \in \mathbb{Z}$. For these values of α, the distributions $\chi_{+}^{n}(t)$ can be shown (Gel'fand and Shilow, 1960; Hörmander, 1983) to reduce to the distributions $\delta^{(-n-1)}$. They can be seen as differentiated or integrated δ-distributions. For $n = -1$, hence $m = -n - 1 = 0$, $\delta^{(0)}$ is the well-known δ-distribution, which fulfills in distributional sense the following scaling relation:

$$\delta(\lambda t) = \lambda^{-1} \delta(t) \qquad \lambda > 0. \tag{5.23}$$

[15] For $\alpha \in \mathbb{C} \backslash \mathbb{R}$, oscillating homogeneous distributions are obtained.

[16] Note that for $\alpha > -1$, $\chi_{\pm}^{\alpha}(t)$ is a regular distribution. A regular distribution is locally integrable (Zemanian, 1965).

In case $n = -2, -3, \ldots$, hence $m = -n - 1 = 1, 2, \ldots$, $\delta^{(m)}(t)$ should be interpreted as an mth order derivative in the weak sense, i.e.

$$\int \delta^{(m)}(t)\, f(t)\, \mathrm{d}t = (-1)^m \left[\partial_t^m f\right]_{t=0} \tag{5.24}$$

$$\triangleq (-1)^m\, \partial_t^m f(0). \tag{5.25}$$

In case $n = 0, 1, \ldots$, hence $m = -n - 1 = -1, -2, \ldots$, $\delta^{(m)}(t)$ should be interpreted as a $-m$th order integration operator, i.e.

$$\int \delta^{(m)}(t)\, f(t)\, \mathrm{d}t = \frac{(-1)^m}{(-m)!} \int_0^\infty t^{-m} f(t)\, \mathrm{d}t \triangleq (-1)^m\, \partial_t^m f(0). \tag{5.26}$$

The scaling relation of the distributions $\delta^{(n)}(t)$ reads

$$\delta^{(n)}(\lambda t) = \lambda^{-n-1}\, \delta^{(n)}(t) \qquad \lambda > 0. \tag{5.27}$$

The wavelet transform of $\chi_+^{-n-1}(t)$, and hence of $\delta^{(n)}(t)$, thus reads

$$\mathfrak{W}\{\chi_+^{-n-1}(t), \psi\}(\sigma, b) = \sigma^{-n-1} \int \delta^{(n)}(t) \psi^*(t - b/\sigma)\, \mathrm{d}t \tag{5.28}$$

$$= \sigma^{-n-1}\, (-1)^n\, \partial_t^n \psi^*(-b/\sigma). \tag{5.29}$$

Note again that along lines of b/σ is constant, the amplitude of the wavelet transform of $\chi_+^{-n-1}(t)$ is ruled by the factor σ^{-n-1}. The WTMMLs are among the lines for which b/σ is constant. Equations (5.20) and (5.29) describe together the decay behavior of the wavelet transform of homogeneous distributions χ_+^α for $\alpha \in \mathbb{R}$. In section 5.3.1 the wavelet transform of homogeneous distributions will be worked out in more detail.

The most general homogeneous distribution of degree α is given by

$$\chi_{\text{cusp}}^\alpha(t) = c_+ \chi_+^\alpha(t) + c_- \chi_-^\alpha(t), \tag{5.30}$$

which fulfills indeed the following scaling relation

$$\chi_{\text{cusp}}^\alpha(\lambda t) = \lambda^\alpha \chi_{\text{cusp}}^\alpha(t), \qquad \lambda > 0. \tag{5.31}$$

For $c_+ = 0$ or $c_- = 0$, $\chi_{\text{cusp}}^\alpha(t)$ reduces to $c_- \chi_-^\alpha(t)$ and $c_+ \chi_+^\alpha(t)$ respectively[17]. The wavelet transform of the cusp has a decay behavior equal to the one-sided homogeneous distribution. It is interesting to observe that independent of the shape, the cusp and the one-sided homogeneous distributions will yield equal Hölder exponents. In order to be able to pronounce upon the signature of a singularity with a specific Hölder exponent, more information is required. In the examples this issue will be addressed again.

[17]Note that in case that $c_+ \neq 0$ as well as $c_- \neq 0$, a slightly different denominator is required, in order to get rid of the simple poles (Gel'fand and Shilow, 1960, page 64). To circumvent this problem it is assumed that $\alpha \notin \mathbb{Z}$ for the cusp. See also appendix B.

5.2.5 Examples

The examples in this subsection serve a threefold goal. First, I want to pay attention to the numerical implementation of homogeneous distributions. Secondly, I want to illustrate that the presented theory for determining the Hölder exponent via a WTMM analysis works. As this has been done by quite some authors before me as well (Grossmann et al., 1987; Mallat and Hwang, 1992; Staal, 1995; Herrmann, 1997), this should not be a big surprise. Thirdly, the limitations of the multiscale amplitude analysis will be pointed out, especially in determining the exact signature of the singularity.

Numerical aspects —Numerically, it seems not to be useful to talk about singularities, or discontinuities. A set of samples with a uniform sampling distance can always be interpolated to form a smooth function on a finer sampling grid. This kind of reasoning is not wrong, but it is too limited. Mallat and Hwang (1992) and Herrmann (1997) argue that one should consider a set of samples at the resolution it is observed, or at coarser resolutions. One can pronounce upon the decay behavior for resolutions coarser than the observed resolution, but one can not enunciate the decay behavior in the scale range smaller than the resolution. Questions about the latter scale range can be regarded as metaphysical, unless a priori information of the type of the signal is available.

Another numerical aspect is specifically related to the interpretation of singularities with negative Hölder exponents. Negative Hölder exponents are associated with distributions, which do not have an explicit functional form. They can, however, be numerically implemented in a consistent way. Consider, for example, the δ-distribution. The numerical representation of the δ-distribution consists of a set of zeros and a single $1/\Delta t$ at the location of the δ-distribution, where Δt is the sampling distance. If implemented like that, the expected scaling behavior given by a Hölder exponent of -1 will be found.

For other homogeneous distributions, with $\alpha > -1$, I have chosen a numerical implementation[18], which yields Hölder exponents very close to the theoretical results (see examples below). Note that the straightforward numerical implementation of $|t|_{\pm}^{\alpha}$ yields deviating Hölder exponents, especially for negative scaling exponents.

Examples —Figure 5.6a shows the numerical implementation of five homogeneous distributions according to the numerical implementation described above. The wavelet transform (in grey scale) with the extracted WTMMLS on top are shown in Figure 5.6b. The wavelet transform is carried out with the first derivative of a

[18]The numerical implementation takes in easily checkable MATLAB-code the following form:

$\chi_{+}^{\alpha}(t) \longrightarrow$ `chi = diff([0:N].^(alpha+1));chi=chi/max(chi);`

where `N` is the number of samples. The corresponding sample values should be chosen according to $[1/2, 3/2, \ldots, N - 1/2]$.

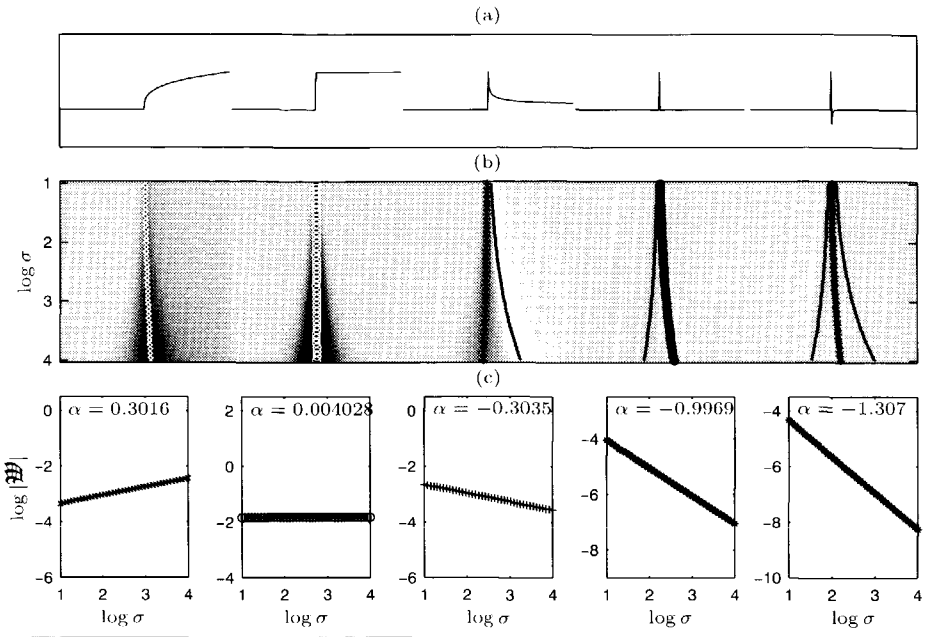


Fig. 5.6 *(a) Five homogeneous distributions with, from left to right, local scaling exponents* $\alpha = 0.3, 0, -0.3, -1, -1.3$. *(b) The wavelet transform of the five distributions with the extracted* WTMMLs *on top. (c) The natural logarithm of the amplitudes along five* WTMMLs *plotted as a function of the natural logarithm of the scale parameter. The estimated local Hölder exponents are very close to the actual exponents.*

Gaussian function. The amplitudes along five WTMMLs are shown in the bottom row of Figure 5.6. The estimated Hölder exponents are very close to the expected values. Looking at the number of WTMMLs emerging at the singularities with Hölder exponents $-1 \leq \alpha < 0$, it can be concluded that for this type of singularity the analysis could have been done, more efficiently, with the Gaussian function itself as well. Hereinafter, when seismic reflection data is analyzed this property will be used (see also section 5.4).

Limitation —Life seems to be beautiful, with the above described theory in one hand and the correct numerical implementation in the other hand, but the method has one clear limitation: it is not able to distinguish between singularities of a different signature but with the same Hölder exponent. Figure 5.7 shows the modulus maxima analysis for four homogeneous distributions:

$$\chi_+^{-0.3}(t) \qquad \chi_-^{-0.3}(t) \qquad -\chi_+^{-0.3}(t) \qquad -\chi_-^{-0.3}(t).$$

The distributions are shown in Figure 5.7a; their wavelet transforms and the ex-

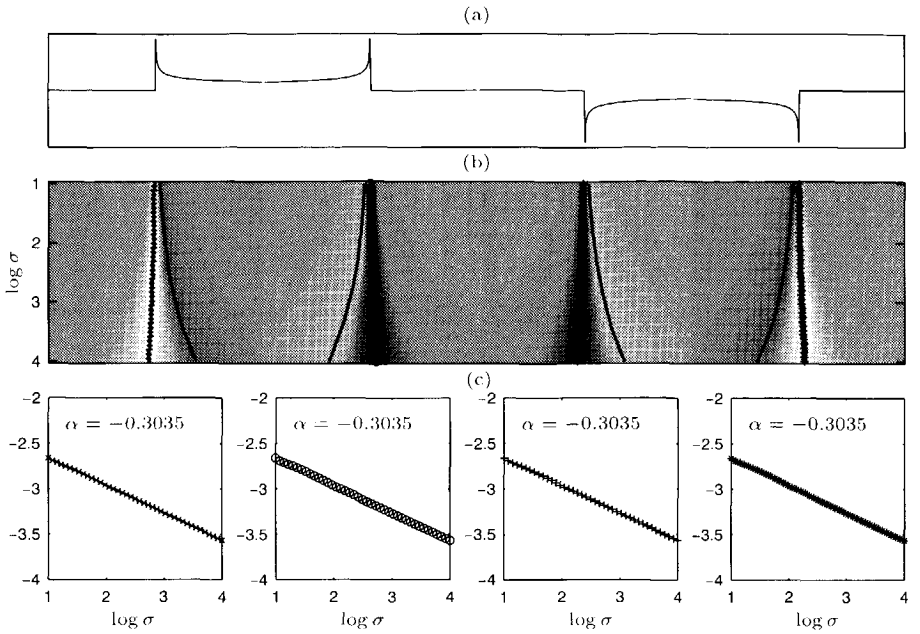


Fig. 5.7 *(a) Four homogeneous distributions all with the same Hölder exponent of* $\alpha = -0.3$*, but with clearly distinctive signature. (b) The wavelet transform of the four distributions with the extracted* WTMMLs *on top. (c) The natural logarithm of the amplitudes along four* WTMMLs *plotted as a function of the natural logarithm of the scale parameter. The four singularities cannot be discriminated on the basis of their Hölder exponents.*

tracted WTMMLs are shown in Figure 5.7b. On the basis of the estimated Hölder exponents the four distributions can not be discriminated (bottom row of Figure 5.7). If the distributions in Figure 5.7a are interpreted as the acoustic propagation velocity or the acoustic impedance as a function of depth, a geologist would associate them with notably different geological settings (Verhelst, 1997).

Is there more information in the wavelet transform of a homogeneous distribution or, more generally, in a signal that has interacted with a homogeneous distribution (such as a wave), which can be used advantageously to analyze singularities more deterministically? In the wavelet transform of the homogeneous distribution in equation (5.20), I have used up to now only the factor σ^{α} and I have not looked at the actual structure of $U(u)$. In the next section, the structure of $U(u)$ will be analyzed, and it will be found that the clue is the *instantaneous phase*.

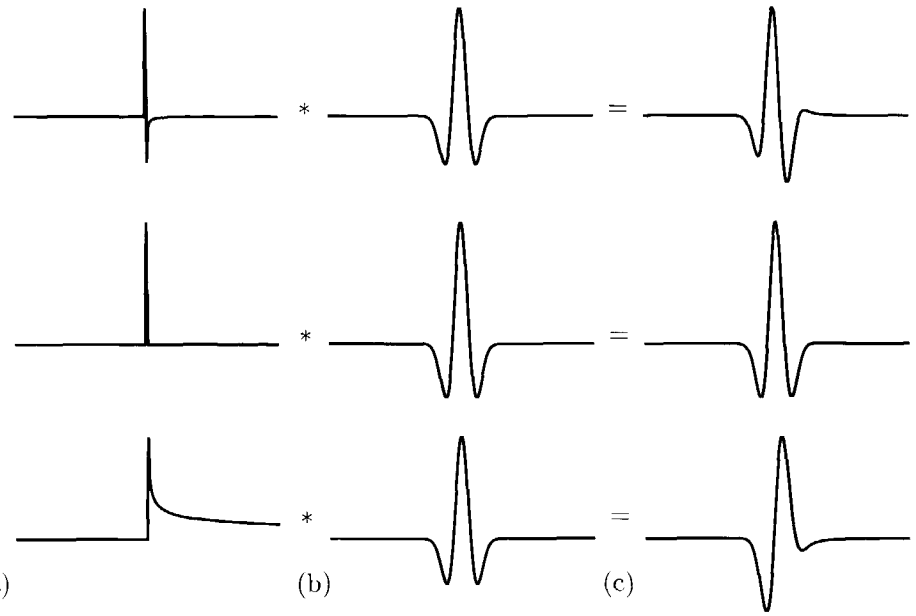


Fig. 5.8 *The basic phase idea. (a) A column of three numerical representations of three homogeneous distributions of the type $\chi_+^\alpha(t)$. The top one has $\alpha = -1.33$; the middle is a δ-distribution, hence here is $\alpha = -1$; the bottom one has a scaling exponent $\alpha = -0.33$. (b) The three homogeneous distributions in the first column are convolved with a Ricker wavelet. (c) The result of the convolution. It shows that a δ-distribution does not affect the shape of the Ricker wavelet, whereas the other two do.*

5.3 Phase and homogeneous distributions

In 1992, Mallat and Hwang wrote:

> *"It is thus necessary to combine the modulus and the phase information to characterize the different singularities, but no effective method has been derived yet."*

Not aware of this statement and a little bit tired of looking at the rather popular amplitudes alone, Hoekstra (Hoekstra and Dessing, 1998) decided to look how singularities induce phase changes. Rather simple relations came out. It was later discovered that similar results were obtained by Holschneider (1995) as well, although Holschneider did not utilize the results.

The idea of the singularity-induced phase changes is illustrated in Figure 5.8. It shows three homogeneous distributions: one with a scaling exponent of -1.33, a δ-distribution, and one with a scaling exponent of -0.33. The convolution of the distributions with a Ricker wavelet (Figure 5.8b), i.e. the second derivative of a Gaussian, yields the signals in Figure 5.8c. As expected, the δ-distribution does

not affect the shape of the Ricker wavelet. The other two distributions do change the shape of the Ricker wavelet. Two effects can be distinguished: (1) a *fractional* differentiation (for $\alpha = -1.33$), or a *fractional* integration (for $\alpha = -0.33$); (2) a phase change. It will be shown in the sequel that these effects are closely linked. The simple example of Figure 5.8 suggests that singularities of the type $\pm\chi_{\pm}^{\alpha}(t)$ are uniquely characterized by looking at:

- the amplitude behavior across different scales for determining the *singularity strength*, in terms of the local Hölder exponent α;
- the induced phase change to determine the *signature of the singularity*.

In order to make the latter statement more precise, I will present in section 5.3.1 the analytic wavelet transform of a homogeneous distribution with respect to a progressive wavelet[19], in full detail. In section 5.3.2, a method to detect the phase will be presented. It is based on the analytic signal (Gabor, 1946). In that section, I will also show that singularities of different strength and signature can be uniquely placed in a so-called $(\varphi - \alpha)$-diamond.

5.3.1 Wavelet transform of homogeneous distributions

In section 5.2.4 the wavelet transform of homogeneous distributions has been discussed for the first time in this chapter. I ended the analysis of the wavelet transform of homogeneous distributions for an arbitrary wavelet ψ in $S_{\pm}(\mathbb{R})$ or $S_0(\mathbb{R})$ with (cf. equation 5.20)

$$\mathfrak{W}\{\chi_{\pm}^{\alpha}, \psi\}(\sigma, b) = \sigma^{\alpha} U_{\pm}(b/\sigma). \tag{5.32}$$

The factor $U_{\pm}(u)$ can be most easily resolved, if a progressive wavelet, $\psi_p(t)$, is used in the wavelet transform. I consider $U_{+}(u)$ given by equation (5.21) first. Substitution of the inverse Fourier transform of $\psi_p(\omega)$ for $\psi_p(t)$, and using the fact that $\psi_p(\omega) = 0$ for $\omega \leq 0$, yield

$$U_{+}(u) = \frac{1}{\Gamma(\alpha+1)} \frac{1}{2\pi} \int_0^{\infty} t^{\alpha} \left[\int_0^{\infty} \psi_p(\omega)\, \mathrm{e}^{j\omega(t-u)}\, \mathrm{d}\omega \right]^{*} \mathrm{d}t \tag{5.33}$$

$$= \frac{1}{\Gamma(\alpha+1)} \frac{1}{2\pi} \int_0^{\infty} \psi_p^{*}(\omega)\, \mathrm{e}^{j\omega u} \left[\int_0^{\infty} t^{\alpha}\, \mathrm{e}^{-j\omega t}\, \mathrm{d}t \right] \mathrm{d}\omega, \tag{5.34}$$

where the order of integration has been changed on account of Fubini's theorem (Daubechies, 1992). The integral $\int_0^{\infty} t^{\alpha}\, \mathrm{e}^{-j\omega t}\, \mathrm{d}t$ should be read as (Holschneider, 1995)

$$\lim_{\lambda \downarrow 0} \int_0^{\infty} t^{\alpha}\, \mathrm{e}^{-\lambda t - j\omega t}\, \mathrm{d}t, \tag{5.35}$$

[19] Progressive and regressive wavelets are discussed on page 156.

for convergence reasons. Using the fact that (Abramowitz and Stegun, 1970)

$$\begin{aligned}\Gamma(\alpha+1) &= \int_0^\infty t^\alpha \, \mathrm{e}^{-t} \, \mathrm{d}t \\ &= \gamma^{\alpha+1} \int_0^\infty t^\alpha \, \mathrm{e}^{-\gamma t} \, \mathrm{d}t,\end{aligned}$$

one can find, with $\gamma = \mathrm{e}^{j\frac{\pi}{2}}\omega$, that

$$U_+(u) = \frac{1}{2\pi} \int_0^\infty \frac{\mathrm{e}^{-j\frac{\pi}{2}(\alpha+1)}}{\omega^{\alpha+1}} \, \psi_p^*(\omega) \, \mathrm{e}^{j\omega u} \, \mathrm{d}\omega. \tag{5.36}$$

Along similar lines one can find for $U_-(u)$

$$U_-(u) = \frac{1}{2\pi} \int_0^\infty \frac{\mathrm{e}^{+j\frac{\pi}{2}(\alpha+1)}}{\omega^{\alpha+1}} \, \psi_p^*(\omega) \, \mathrm{e}^{j\omega u} \, \mathrm{d}\omega. \tag{5.37}$$

Putting equations (5.32), (5.36), and (5.37) together, the wavelet transform of $\chi_\pm^\alpha(t)$ with respect to a progressive wavelet[20] can be written as

$$\begin{cases} \mathfrak{W}\{\chi_\pm^\alpha, \psi_p\}(\sigma, b) = \sigma^\alpha U_\pm(b/\sigma) \\ \\ U_\pm(u) = \dfrac{1}{2\pi} \displaystyle\int_0^\infty \frac{\mathrm{e}^{\mp j\frac{\pi}{2}(\alpha+1)}}{\omega^{\alpha+1}} \, \psi_p^*(\omega) \, \mathrm{e}^{j\omega u} \, \mathrm{d}\omega. \end{cases} \tag{5.38}$$

The important elements in the wavelet transform, depending on the signature and strength of the singularity, are:

- *Amplitude*: σ^α
  This factor has been discussed in the previous section; it is directly related to the singularity strength and it expresses the decay behavior as a function of the scale parameter σ.

- *Fractional integration or differentiation*: $A = \frac{\mathrm{e}^{+j\frac{\pi}{2}(\alpha+1)}}{\omega^{\alpha+1}}$

  1. The complex factor A in the Fourier integral results in a fractional integration for $\alpha > -1$, and in a fractional differentiation for $\alpha < -1$. For $\alpha = -1$, $\chi_\pm^\alpha(t)$ reduces to the δ-distribution, and, consequently, the shape of the analyzing wavelet is not affected.
  2. The homogeneous distributions $\chi_\pm^\alpha(t)$ induce, in addition, a phase change. Observe that this phase change cannot be seen apart from the fractional integration or differentiation. Along a WTMML the phase change is constant. This can be easily seen since b/σ is constant along a WTMML.

[20] Note that equivalent derivations can be carried out for non-progressive, but real wavelets. In that case it is used that $\psi(t) = \frac{1}{\pi}\Re\left[\int_0^\infty \psi(\omega)\exp(j\omega t)\,\mathrm{d}\omega\right]$, and that $\chi_\pm^\alpha$ is real.

	$+\chi_+^\alpha$	$+\chi_-^\alpha$	$-\chi_+^\alpha$	$-\chi_-^\alpha$
φ	$\frac{\pi}{2}(-\alpha-1)$	$\frac{\pi}{2}(\alpha+1)$	$\frac{\pi}{2}(-\alpha+1)$	$\frac{\pi}{2}(\alpha-1)$
φ	$\frac{\pi}{2}(-\alpha+q+1)+\theta$	$\frac{\pi}{2}(\alpha+q-1)+\theta$	$\frac{\pi}{2}(-\alpha+q-1)+\theta$	$\frac{\pi}{2}(\alpha+q+1)+\theta$

Table 5.1 *Relation between the signature of the singularity and the induced phase change. The top row gives the type of singularities. The middle row gives the induced phase changes for the different singularity types as a function of the scaling exponent α. The bottom row is discussed in section 5.4. It gives the induced phase changes of the singularity convolved with a seismic wavelet with q vanishing moments (defined according to equation 5.53) and an additional phase change θ.*

Therefore the induced phase change can be derived from the numerator of A to be

$$\varphi = \mp\frac{\pi}{2}(\alpha+1). \tag{5.39}$$

In Table 5.1, the induced phase changes for singularities with the same α, but with different signatures are concisely put together.

Figure 5.9 summarizes the described relation between the phase change and the Hölder exponent on the one hand, and the signature and strength of the singularity on the other hand, in a $(\varphi-\alpha)$-diamond. In analyzing the reflectivity of waves at singularities the $(\varphi-\alpha)$-diamond will be utilized to uniquely determine the strength and signature of a singularity at which a wave is reflected.

In section 5.2.4, the most general homogeneous distribution of degree α has been introduced (cf. equation 5.30)

$$\chi_{\text{cusp}}^\alpha(t) = c_+\chi_+^\alpha(t) + c_-\chi_-^\alpha(t).$$

Using the scaling relation for the cusp of equation (5.31), and the results of equation (5.38), one can easily find that the wavelet transform of $\chi_{\text{cusp}}^\alpha(t)$ with respect to a progressive wavelet is

$$\begin{cases} \mathcal{W}\{\chi_{\text{cusp}}^\alpha, \psi_p\}(\sigma, b) = \sigma^\alpha U_{\text{cusp}}(b/\sigma) \\ \\ U_{\text{cusp}}(u) = \left(c_+ + c_- \, \mathrm{e}^{j\pi(\alpha+1)}\right) U_+(u). \end{cases} \tag{5.40}$$

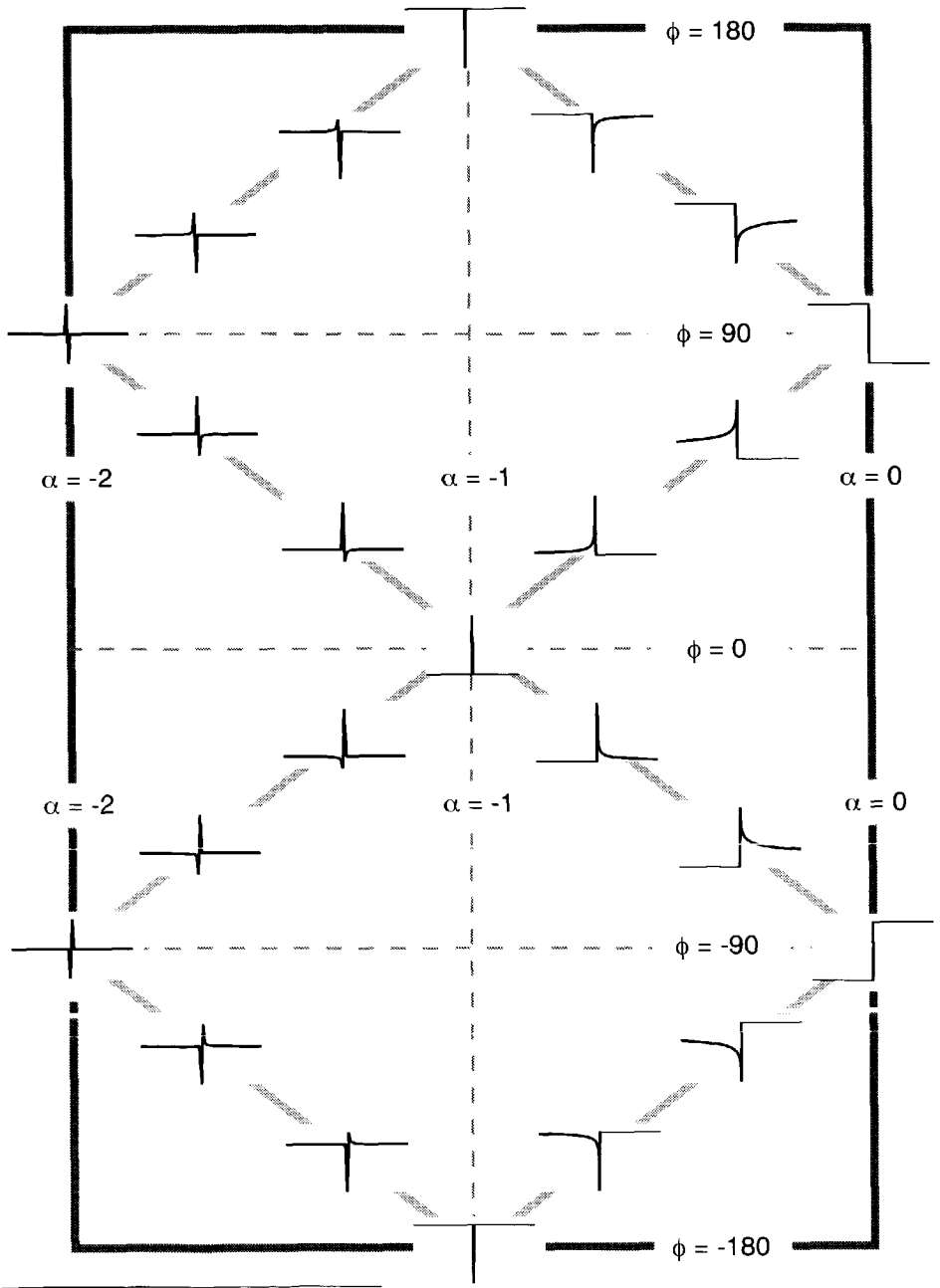


Fig. 5.9 *Relation between the induced phase change on the one hand, and the singularity strength and signature for a one-sided homogeneous distributions on the other hand. The vertical axis denotes the induced phase and the horizontal axis the scaling exponent of the homogeneous distribution, i.e. the singularity strength. For each scaling exponent four signatures are considered, in line with the four profiles for one exponent depicted in Figure 5.7. Note that the phase is in degrees.*

Following the interpretation of the factors of the wavelet transform of the one-sided homogeneous distributions, one can easily identify:

- *Amplitude*: σ^α
  See the discussion for the one-sided homogeneous distributions on page 164.
- *Fractional integration or differentiation*: $A = \left(\frac{e^{-j\frac{\pi}{2}(\alpha+1)}}{\omega^{\alpha+1}}\right)\left(c_+ + c_- e^{j\pi(\alpha+1)}\right)$

  1. The first factor of A between brackets carries out the fractional integration or differentiation according to the discussion for the one-sided homogeneous distributions on page 164.
  2. On top of the phase change induced by the numerator of the first factor of A, the second factor influences the signature of the wavelet that has interacted with the cusp. For the actual induced phase change φ, the following two cases are discriminated only:
     (a) $c_- = 0$ or $c_+ = 0$: $\chi^\alpha_{\text{cusp}}(t)$ reduces to $c_+\chi^\alpha_+(t)$ or $c_-\chi^\alpha_-(t)$ respectively, and the apparatus developed on page 164 can be utilized.
     (b) $c_+ = c_- \neq 0$: for this case, A reduces to

$$A = \frac{2c_+ \cos(\pi(\alpha+1)/2)}{\omega^{\alpha+1}}, \tag{5.41}$$

     which means that the induced phase change equals $\varphi = 0$ or $\varphi = \pi$.
     (c) $c_+ = -c_- \neq 0$: for this case A reduces to

$$A = \frac{-2jc_+ \sin(\pi(\alpha+1)/2)}{\omega^{\alpha+1}}, \tag{5.42}$$

     which means that the induced phase change equals $\varphi = \pm\pi/2$.

5.3.2 Complex trace analysis

Several methods exist to extract phase information. A possible method is to carry out a wavelet transformation with a progressive wavelet exactly according to the theory described in section 5.3.1. The Morlet wavelet (Morlet et al., 1982; Daubechies, 1992) is a good and easy to use candidate for such an analysis; Holschneider (1995) discusses a number of alternative progressive wavelets. The phase $\varphi(\sigma, b)$ in the whole (σ, b)-plane of the wavelet transform of a function f is given by

$$\varphi(\sigma, b) = \arctan\left(\frac{\Im[\mathfrak{W}\{f, \psi_p\}(\sigma, b)]}{\Re[\mathfrak{W}\{f, \psi_p\}(\sigma, b)]}\right). \tag{5.43}$$

The phase change induced by a singularity should be extracted along the WTMMLs. For a single homogeneous distribution the induced phase change along a WTMML is independent of the scale (Grossmann et al., 1987; Holschneider, 1995). Although

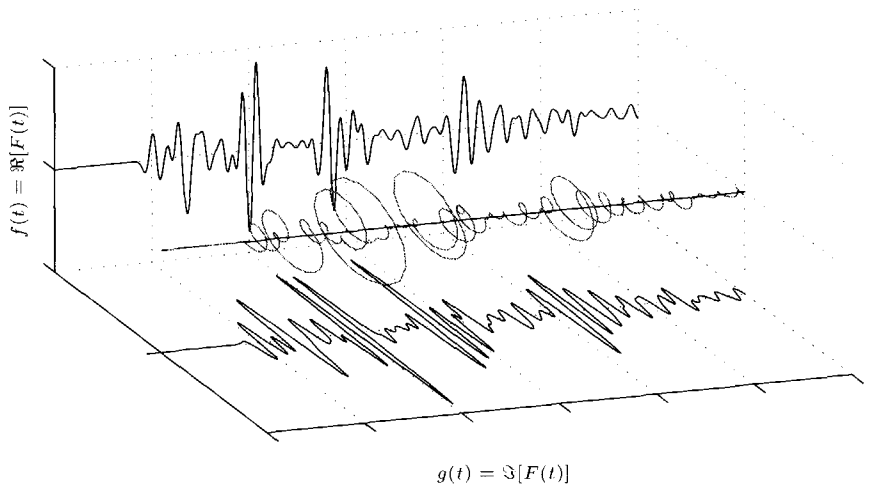


Fig. 5.10 *Original function $f(t)$ and the function in quadrature $g(t)$. Together they form a complex trace $F(t)$, which can be characterized by an instantaneous amplitude and an instantaneous phase.*

the results obtained according to this method are fully consistent with the theory, I prefer to use a different method based upon the analytic signal, primarily because the signal to which the method will finally be applied has already interacted with the singularity. The method to be proposed hereinafter turns out to be slightly simpler to use in that situation. It requires less user dependent options.

Gabor (1946), him again, introduced the concept analytic signal for the analysis of time-dependent signals. The relevance of the analytic signal as the basis for complex trace analysis and attribute extraction has long been recognized in the seismic community (see for example, Neidell and Taner, 1971; Taner et al., 1979; Robertson and Nogami, 1984; Berkhout, 1984).

Following Engelhard (1996) and Gabor (1946), consider a non-periodic, real time signal $f(t)$. It can be expressed in the form

$$f(t) = a(t) \cos \varphi(t), \tag{5.44}$$

where $a(t)$ is the instantaneous amplitude or envelope, and where $\varphi(t)$ is the instantaneous phase. A representation like that is far from being unique. One can associate an infinity of pairs $(a(t), \varphi(t))$ to a given real signal $f(t)$, such that (5.44) holds (Picinbono and Martin, 1983; Delprat et al., 1992). Among the possible pairs, the couple $(a(t), \varphi(t))$ for which

$$F(t) = a(t)\, e^{j\varphi(t)} = f(t) + jg(t), \tag{5.45}$$

with $g(t)$ the Hilbert transform of $f(t)$ is a canonical choice. $F(t)$ is called the analytic signal. The real part of $F(t)$ is the original time signal, and the imaginary

part is called the signal in quadrature, i.e.

$$g(t) = a(t) \sin \varphi(t). \tag{5.46}$$

The function $g(t)$ is 90 degrees phase shifted with respect to $f(t)$; it is orthogonal to $f(t)$. Figure 5.10 illustrates the idea of a complex trace. The complex trace $F(t)$ forms a rotating vector. With the help of the rotating complex vector, one can uniquely assign an amplitude and a phase to $f(t)$ for all values of t. The instantaneous angular frequency of $f(t)$ is defined by the derivative of the instantaneous phase, i.e.

$$\Omega(t) \triangleq \frac{\mathrm{d}\varphi(t)}{\mathrm{d}t}. \tag{5.47}$$

Though the signal in quadrature can be computed with the help of the Hilbert transform, a more practical road uses the Fourier transform. The analytic signal is related to the Fourier spectrum of the signal as follows:

$$F(t) = \frac{1}{\pi} \int_0^\infty \hat{f}(\omega)\, \mathrm{e}^{j\omega t}\, \mathrm{d}\omega. \tag{5.48}$$

Given the analytic signal, it is easy to derive the envelope $a(t)$ and the instantaneous phase $\varphi(t)$, according to

$$a^2(t) = f^2(t) + g^2(t) = F(t)F^*(t) \tag{5.49}$$

$$\tan \varphi(t) = \frac{g(t)}{f(t)} = \frac{1}{j} \frac{[F(t) - F^*(t)]}{[F(t) + F^*(t)]}. \tag{5.50}$$

The phase of an event —The described algorithm provides the instantaneous phase, $\varphi(t)$, and the envelope, $a(t)$. Actually, one wants to assign to a single event, for example a reflection from a single boundary, a unique phase. The unique phase is extracted, via an interpolating algorithm, at the point where the envelope reaches its maximum. The phase at this point exactly corresponds with the phase one would intuitively assign to an event. It corresponds to the phase that will be determined at a WTMM point according to the theory of section 5.3.1. Figure 5.11 illustrates how a unique phase is assigned to an event.

5.4 Influence of seismic wavelet

Before I turn to the analysis of the reflectivity of waves at isolated singularities one issue has to be addressed. A seismic measurement consists of the impulse response of the earth convolved with a band-limited seismic wavelet. How does a seismic wavelet

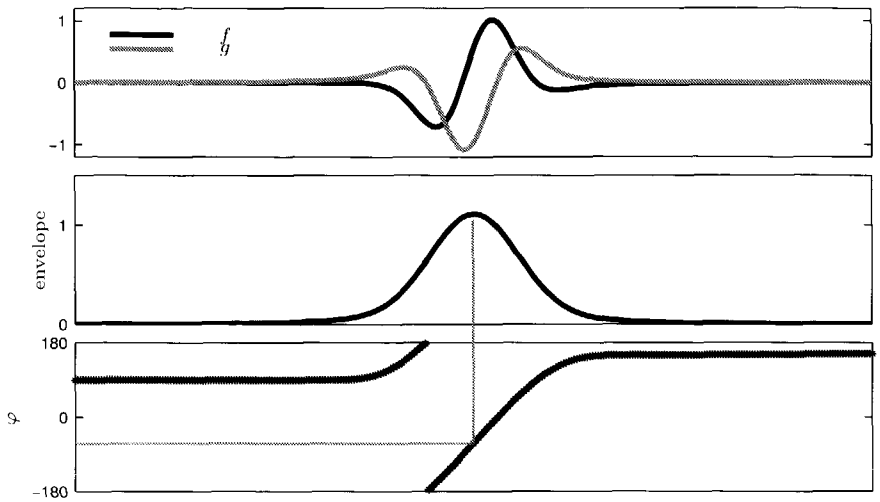


Fig. 5.11 *To a single seismic event, one can uniquely assign a phase by extracting the phase at the point where the envelope reaches its maximum. The top picture shows the original function $f(t)$ and the function in quadrature (its Hilbert transform), $g(t)$. The middle picture shows the instantaneous amplitude or envelope, $a(t)$. The bottom picture shows the instantaneous phase, $\varphi(t)$; the phase at the maximum of the envelope is assigned as the phase of the event.*

influence the detection of a singularity? Seemingly a strange question, because the seismic wavelet is actually the measuring object. However, since different seismic wavelets yield different reflection responses for the same subsurface structure, it is important to exactly characterize the influence of the seismic wavelet. For that reason, the scale and phase behavior of an isolated singularity (as an example of the impulse response of the earth) is compared with the scale and phase behavior of the isolated singularity convolved with a seismic wavelet. More mathematically, I want to compare $\mathfrak{W}\{\chi_+^\alpha, \psi_p\}(\sigma, b)$, for which an explicit expression has been derived in equation (5.38), and $\mathfrak{W}\{\Xi_+^\alpha, \psi_p\}(\sigma, b)$, where

$$\Xi_+^\alpha(t) = (\chi_+^\alpha * \psi_\theta)(t) \tag{5.51}$$

is the convolution of the homogeneous distribution $\chi_+^\alpha(t)$ with a seismic wavelet $\psi_\theta(t)$. Equation (5.51) can also be written as the real part of the convolution of the homogeneous distribution with the analytic signal, $z_\theta(t)$, of the seismic wavelet $\psi_\theta(t)$, i.e.

$$\Xi_+^\alpha(t) = \Re\left[(\chi_+^\alpha * z_\theta)(t)\right]. \tag{5.52}$$

In the sequel of this section a realistic seismic wavelet is constructed first. Then $\mathfrak{W}\{\Xi_+^\alpha, \psi_p\}(\sigma, b)$ is computed and compared with $\mathfrak{W}\{\chi_+^\alpha, \psi_p\}(\sigma, b)$. I finalize this section with an example.

For the seismic wavelet $\psi_\theta(t)$ a broad class of functions is allowed. They are constructed in the following way:

Procedure 5.2: Construction of realistic seismic wavelet $\psi_\theta(t)$

1. *Define a basis seismic wavelet $\psi_s(t)$ as the qth derivative of a smoothing function of compact support, according to*

$$\psi_s(t) = -\partial_t^q \phi(t), \tag{5.53}$$

where the minus-sign is for later convenience. Hence, the basis seismic wavelet $\psi_s(t)$ has q vanishing moments[21].

2. *Derive the analytic signal $z_s(t)$ of $\psi_s(t)$ according to equation (5.48), yielding*

$$z_s(t) = a(t)\, \mathrm{e}^{j\varphi(t)}.$$

3. *An additional phase term θ is allowed to be present in the actual seismic wavelet. The additional phase change is easily accounted for by multiplying $z_s(t)$ by $\mathrm{e}^{j\theta}$, i.e.*

$$z_\theta(t) = z_s(t)\, \mathrm{e}^{j\theta}.$$

The actual seismic wavelet, $\psi_\theta(t)$, is now given by the real part of the analytic signal, $z_\theta(t)$.

4. *Optionally, one can allow for an additional dilation to allow for propagation losses or to bring the wavelet in the appropriate frequency range. This operation can be easily carried out with the dilation operator $\mathcal{E}_\sigma$ (see section 2.2).*

The Fourier transform of $z_\theta(t)$ is given by

$$Z_\theta(\omega) = \begin{cases} -2(j\omega)^q \phi(\omega)\, \mathrm{e}^{j\theta} & \omega > 0 \\ 0 & \omega \leq 0. \end{cases} \tag{5.54}$$

Now, using equations (5.38), with $\sigma = 1$, one can see that

$$\begin{aligned} \Xi_+^\alpha(t) &= \Re\left[(\chi_+^\alpha * z_\theta)(t)\right] \\ &= \Re\left[\frac{1}{2\pi}\int_0^\infty \frac{\mathrm{e}^{-j\frac{\pi}{2}(\alpha+1)}}{\omega^{\alpha+1}} Z_\theta(\omega)\, \mathrm{e}^{j\omega t}\, \mathrm{d}\omega\right]. \end{aligned} \tag{5.55}$$

[21] The parameter q in the definition of the basic seismic wavelet rules its decay behavior towards $\omega = 0$. The decay properties for $\omega \to \infty$ are not specified and should be adjusted to the source signature and the presence of non-elastic losses.

Substitution of (5.54), and the fact that $j = \mathrm{e}^{j\pi/2}$ yield

$$\Xi_+^\alpha(t) = \Re\left[\frac{1}{2\pi}\int_0^\infty \frac{\mathrm{e}^{-j\frac{\pi}{2}(\alpha+1)}}{\omega^{\alpha+1}}(-2)\,\mathrm{e}^{j\pi q/2}\omega^q\phi(\omega)\,\mathrm{e}^{j\theta}\,\mathrm{e}^{j\omega t}\,\mathrm{d}\omega\right] \tag{5.56}$$

$$= \Re\left[\frac{1}{2\pi}\int_0^\infty \frac{\mathrm{e}^{-j\frac{\pi}{2}(\alpha-q+1)}}{\omega^{\alpha-q+1}}\,\mathrm{e}^{j\theta}(-2\phi(\omega))\,\mathrm{e}^{j\omega t}\right] \tag{5.57}$$

$$= \Re\left[(\,\mathrm{e}^{j\theta}\chi_+^{\alpha-q} * z_\phi)(t)\right], \tag{5.58}$$

where I have introduced $z_\phi(t)$ as the analytic signal of $-2\phi(t)$. The last expression shows that $\Xi_+^\alpha(t)$ can also be interpreted as the convolution of a homogeneous distribution $\chi_+^{\alpha-q}(t)$ and $-\phi(t)$, with an additional phase change θ. For $-\chi_+^\alpha(t)$, and for $\pm\chi_-^\alpha(t)$ similar expressions can be found. Hence, using equations (5.38) again, it is found that

$$\begin{cases} \mathfrak{W}\{\Xi_\pm^\alpha, \psi_p\}(\sigma, b) = \sigma^{\alpha-q}U_\pm(b/\sigma) \\ \\ U_\pm(u) = \mathrm{e}^{j\theta}\displaystyle\int_0^\infty \frac{\mathrm{e}^{\mp j\frac{\pi}{2}(\alpha\mp q-1)}}{\omega^{\alpha-q-1}}(2\phi(\omega))\psi_p(\omega)\,\mathrm{e}^{j\omega u}\,\mathrm{d}\omega. \end{cases} \tag{5.59}$$

The influence of the seismic wavelet $\psi_\theta(t)$ can be seen by comparing this result with the results for the homogeneous distributions without a prior convolution with a seismic wavelet, as derived in equation (5.38) on page 164. The different elements can be interpreted exactly according to the interpretation given after equation (5.38). It is clear that the seismic wavelet has severe impact on the multiscale amplitude behavior, which changes from σ^α to $\sigma^{\alpha-q}$. The bottom row of Table 5.1 summarizes the phase changes induced by the combination of a homogeneous distribution χ_+^α, and the seismic wavelet, characterized by q, θ and $\phi(t)$, derived from equation (5.59).

In Figure 5.12 the effect of a seismic wavelet is illustrated for the homogeneous distribution $\chi_+^{-0.7}(t)$[22]. Two different seismic wavelets $\psi_\theta(t)$ have been used. The first is given by $\psi_1(t) = -\partial_t\phi(t)$, and the second by $\psi_2(t) = -\partial_t^2\phi(t)$. The smoothing function $\phi(t)$ is a Gaussian function. Hence, $\psi_2(t)$ is a Ricker wavelet, which is from a physical point of view a well-chosen representation for a seismic wavelet[23]. Note that the additional phase change is zero, i.e. $\theta = 0$. Figures 5.12a-c show $\chi_+^{-0.7}(t)$, the convolution of $\chi_+^{-0.7}(t)$ with the first derivative of the Gaussian, and the convolution of $\chi_+^{-0.7}(t)$ with the Ricker wavelet, respectively. In Figure 5.12d the modulus maxima analysis has been carried out with a Gaussian on the functions of 5.12a, 5.12b and 5.12c , and in Figure 5.12e with the first derivative of a Gaussian

[22]Hence, the homogeneous distribution $\chi_+^{-0.7}(t)$ is assumed to be a (band-limited) impulse response of the earth.

[23]The second derivative has a decay behavior of 12dB/octave towards $\omega = 0$, which corresponds to two vanishing moments.

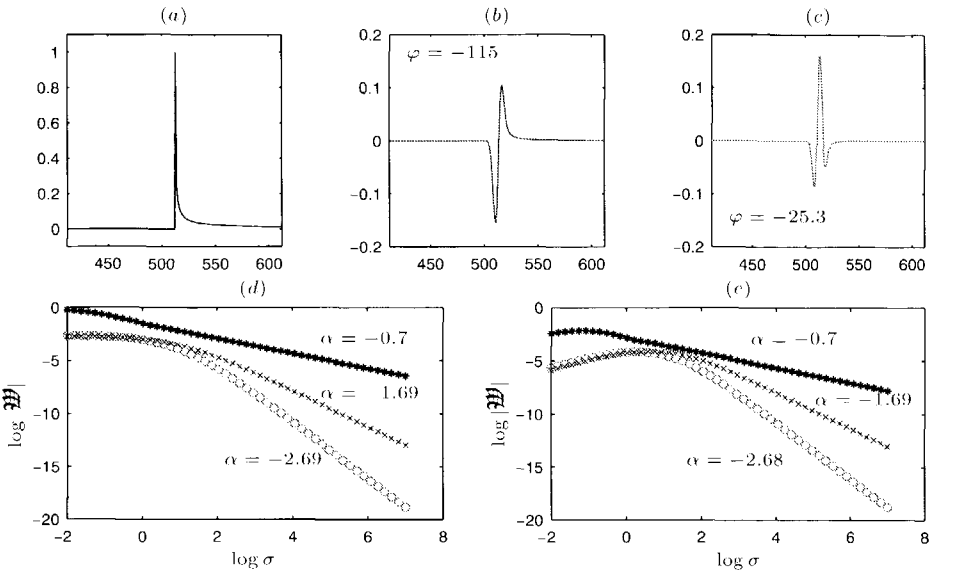


Fig. 5.12 *Influence of a seismic wavelet on the estimated strength of a singularity. (a) The homogeneous distribution* $\chi_+^{-0.7}(t)$. *(b) The homogeneous distribution of (a) convolved with* $\psi_1(t)$. *(c) The homogeneous distribution of (a) convolved with* $\psi_2(t)$. *(d) Amplitude along* WTMML *of (a)-(c). The 'wavelet transform' has been carried out with a Gaussian, which is allowed because we are dealing with negative singularities only. (e) Amplitude along* WTMML *of (a)-(c). The wavelet transform has been carried out with the first derivative of a Gaussian. The slopes of the lines is clearly in agreement with the theory, as is the phase. The extracted phases are shown in the lower left corners of (b) and (c). They are in agreement with the theory.*

on the same three functions[24]. Both analyses yield the results predicted by the theory, both with respect to the decay behavior as a function of σ and with respect to the induced phase change. Due to the fact that a Gaussian gives rise to less WTMMLs it is preferred to use the Gaussian instead of any of its derivatives in a WTMM analysis of seismic data. Finally note that the $(\varphi-\alpha)$-diamond for $(\pm\chi_\pm^\alpha * \psi_\theta)$, with ψ_θ a Ricker wavelet, has exactly the same structure as the $(\varphi-\alpha)$-diamond for $\pm\chi_+^\alpha$ as shown in Figure 5.9.

Remark 5.2 *It is conjectured that the wavelet* $\psi_\theta(t)$ *that results from carrying out procedure 5.2 is the ideal decomposing block for analyzing a seismic reflection response.*

[24]The double analysis is carried out to show that the analyzing wavelet does not influence the analysis results.

5.5 Normal-incidence reflectivity at isolated singularities

Suppose that locally a vertical acoustic velocity profile can be written as an infinite asymptotic series of homogeneous distributions on top of a smooth (polynomial) background

$$c(x_3) = c_0 + \sum_{n=1}^{\infty} \left[c_{+,n}\, \chi_+^{\alpha_n}(x_3 - x_{3,0}) + c_{-,n}\, \chi_-^{\alpha_n}(x_3 - x_{3,0}) \right] + \sum_{n=1}^{\infty} \gamma_n (x_3 - x_{3,0})^n, \quad (5.60)$$

with $\alpha_n \neq 0, 1, 2, \ldots$, and $\alpha_i > \alpha_{i-1}$, and γ_n arbitrary constants[25]. The simplest local expansion takes the following form

$$c(x_3) = c_0 + c_+\, \chi_+^{\alpha}(x_3 - x_{3,0}) + c_-\, \chi_-^{\alpha}(x_3 - x_{3,0}), \quad (5.61)$$

in which case the principal part of (5.60) is considered only, hence $c_+ = c_{+,1}$, $c_- = c_{-,1}$, and $\alpha = \alpha_1$. Equation (5.61) describes a vertical velocity profile with a constant background velocity c_0; a distortion in the form of a homogeneous distribution is superposed on the background velocity, yielding a singularity at $x_3 = x_{3,0}$. The strength of the singularity α is not disturbed by adding a background velocity, although a strict scaling behavior of the form $c(\lambda x_3) = \lambda^{\alpha} c(x_3)$ is not present anymore; the scaling behavior is present for $(c(x_3) - c_0)$. The signature of the singularity is set by the values c_+ and c_-. Here, I only consider $-1 < \alpha < 1$, and the following cases for c_+ and c_-:

- $c_+ = 0$ or $c_- = 0$, giving rise to a one-sided singularity,
- $c_+ = c_- > 0$, giving rise to a cusp, and
- $c_+ = c_- < 0$, giving rise to an 'inverse' cusp.

I am interested in the signature and the amplitude of a wave reflected at singularities of the form given by equation (5.61), both for normal and oblique incidence. Especially, the relation with and the deviation from the reflectivity of waves impinging on traditionally considered step functions (for which $\alpha = 0, c_+ \neq c_-$) will be studied.

[25] One can object to a model consisting of a smooth background and a distortion for two major reasons. A first objection is the fact that such a model does not accord with the fact that the velocity as a function of depth in the earth's subsurface is multifractal (Herrmann, 1997), which implies that it is changing at all scales. Another objection is the fact that for negative $c_{\pm}$ the velocity profile can become negative, which is regarded to be unphysical. One has to take special care to prevent negative velocities.

The road I have followed to derive the forthcoming results can be called heuristic. It is not based on a solid derivation from first principles. It is merely based on advanced back-of-the-envelope calculations supported with appropriately chosen numerical experiments. The results are consistent with the wave equation based derivation of Wapenaar (1997a), which is a derivation for the limit $c_0 \to 0$.

The reader has to remember that the velocity profile of (5.61) is a mathematical construct, which gets a meaning only in the sense of distributions. It should be tested against a properly chosen set of test functions (Schwartz, 1951, 1952; Zemanian, 1965). The natural test function is the wave which interacts at a certain scale depending on the local velocity, with the singular structure. But, since the local velocity depends again on the spatial scale of the wave, the wave is seemingly trapped in a vicious circle of scale and velocity. Since I wish to carry out some analytic computations, which require a tested distribution, I enforce a certain spatial scale σ_1, according to (cf. equation 2.84)

$$c(\sigma_1, x_3) = \frac{1}{\sigma_1} \int c(x_3')\, \phi\left(\frac{x_3 - x_3'}{\sigma_1}\right) \mathrm{d}x_3', \tag{5.62}$$

or, using equation (5.61),

$$c(\sigma_1, x_3) = c_0 + c_+ \chi_+^{\alpha}(\sigma_1, x_3 - x_{3,0}) + c_- \chi_-^{\alpha}(\sigma_1, x_3 - x_{3,0}). \tag{5.63}$$

For the expansions in the next subsection it turns out to be convenient to define a (scale-dependent) peak distortion velocity c_{pd} by[26]

$$c_{pd} = \begin{cases} \max\left[c(\sigma_1, x_3) - c_0\right] & \text{if} \quad c_\pm > 0 \\ \min\left[c(\sigma_1, x_3) - c_0\right] & \text{if} \quad c_\pm < 0. \end{cases} \tag{5.64}$$

Figure 5.13 shows how the peak distortion velocity can be interpreted. For homogeneous distributions with negative exponents, the peak distortion velocity is reached close to the singular point $x_3 = x_{3,0}$ (Figure 5.13a and 5.13b). For homogeneous distributions with positive exponents, the peak distortion velocity is reached at the boundaries of the depth range (Figure 5.13c and 5.13d).

A few remarks apply to the definition of a peak distortion:

1. The actual peak distortion is determined by both the value of $c_\pm$ and the value of σ_1. Hence, a small peak distortion does not necessarily mean a coarse scale parameter σ_1, it can also mean that $c_\pm$ is relatively small.

2. For scales larger than the scale σ_1 at which the velocity profile is considered, the scaling behavior is not affected. For smaller scales the regularity of the scaling function will be observed.

[26] The definition of a peak distortion is more difficult in the case c_+ and c_- have opposite signs. Such a situation is not considered here (see the items on the facing page).

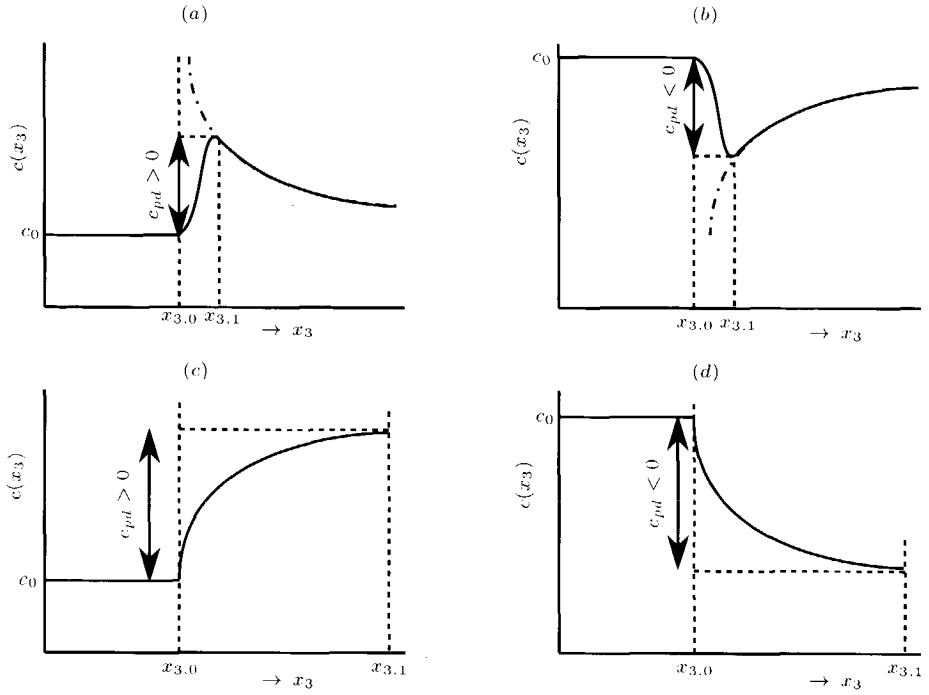


Fig. 5.13 *Interpretation of the peak distortion velocity of homogeneous distributions with negative and positive exponents. (a-b) For negative exponents the peak distortion is reached close to the singular point. The exact location depends on the scaling function ϕ and the scale σ. (c-d) For positive exponents the peak distortion is reached at the boundaries of the depth range.*

3. The introduction of a peak distortion at a specific scale σ_1 does not mean that the wave is interacting with the homogeneous distribution at that specific scale. The reasoning should be turned around: if the wave is living at that specific scale than it will see a velocity profile corresponding to that scale. The wave is in the driving seat!

4. The approach which Verhelst and van Wijngaarden (1997) follow, might be useful to refine the peak distortion value. They iteratively go through the mutually related pair of scale and velocity to find a representation of well-log measurements at a seismic scale range.

5.5.1 Analytic considerations

The reflectivity for a normal-incident acoustic wave is given by (Brekhovskikh and Godin, 1990)

$$R(p=0,x_3) = \frac{\partial \log \rho c}{2\partial x_3}, \tag{5.65}$$

where p denotes the ray parameter; for normal incidence, $p = 0$. Throughout this section normal-incident waves are considered. Hence, in the sequel the dependence on the ray parameter will be omitted, for notational convenience. The density ρ is assumed to be constant and the vertical acoustic velocity profile is given by $c(\sigma_1, x_3)$ of equation (5.63). Under these conditions, the reflectivity takes the following form:

$$R(x_3) = \frac{c_+ \chi_+^{\alpha-1}(\sigma_1, x_3 - x_{3,0}) - c_- \chi_-^{\alpha-1}(\sigma_1, x_3 - x_{3,0})}{2\left(c_0 + c_+ \chi_+^{\alpha}(\sigma_1, x_3 - x_{3,0}) + c_- \chi_-^{\alpha}(\sigma_1, x_3 - x_{3,0})\right)}. \tag{5.66}$$

For the further analysis it is assumed that $c_+ = 0$ or $c_- = 0$. At the end of the analysis the effects for the two-sided homogeneous distributions will be pointed out. Equation (5.66) reduces to

$$R(x_3) = \frac{\pm c_\pm \chi_\pm^{\alpha-1}}{2\left(c_0 + c_\pm \chi_\pm^{\alpha}\right)}, \tag{5.67}$$

where the dependency of $\chi_\pm^{\alpha}$ on $(\sigma_1, x_3 - x_{3,0})$ has been omitted, for notational convenience. The introduction of the peak distortion, c_{pd}, naturally allows for a distinction between $c_{pd} < c_0$ and $c_{pd} > c_0$. Furthermore, a distinction has to be made between positive ($c_\pm > 0$) and negative ($c_+ < 0$) distortions. Let me first consider two cases with positive distortions.

CASE I : *Positive distortion; peak distortion smaller than background velocity*
Mathematically the following situation is considered:

$$c_\pm > 0, \quad \text{and} \quad c_{pd} < c_0. \tag{5.68}$$

Because c_0 is now always larger than the distortion $c_\pm \chi_\pm^{\alpha}(\sigma_1, x_3 - x_{3,0})$, the reflectivity $R(x_3)$ can be written in a series expansion according to

$$\begin{aligned} R(x_3) &= \frac{\pm c_\pm \chi_\pm^{\alpha-1}}{2c_0}\left(1 + \frac{c_\pm \chi_\pm^{\alpha}}{c_0}\right)^{-1} \\ &= \frac{\pm c_\pm \chi_\pm^{\alpha-1}}{2c_0}\left(1 - \frac{c_\pm \chi_\pm^{\alpha}}{c_0} + \frac{c_\pm^2 \chi_\pm^{2\alpha}}{c_0^2} - \cdots\right). \end{aligned} \tag{5.69}$$

In the limit that $c_0 \gg c_{pd}$ the local scaling behavior of the reflectivity at $x_3 = x_{3,0}$ is dominated by the first term, i.e.

$$R(x_3) \simeq \pm \frac{c_\pm \chi_\pm^{\alpha-1}}{2c_0}. \tag{5.70}$$

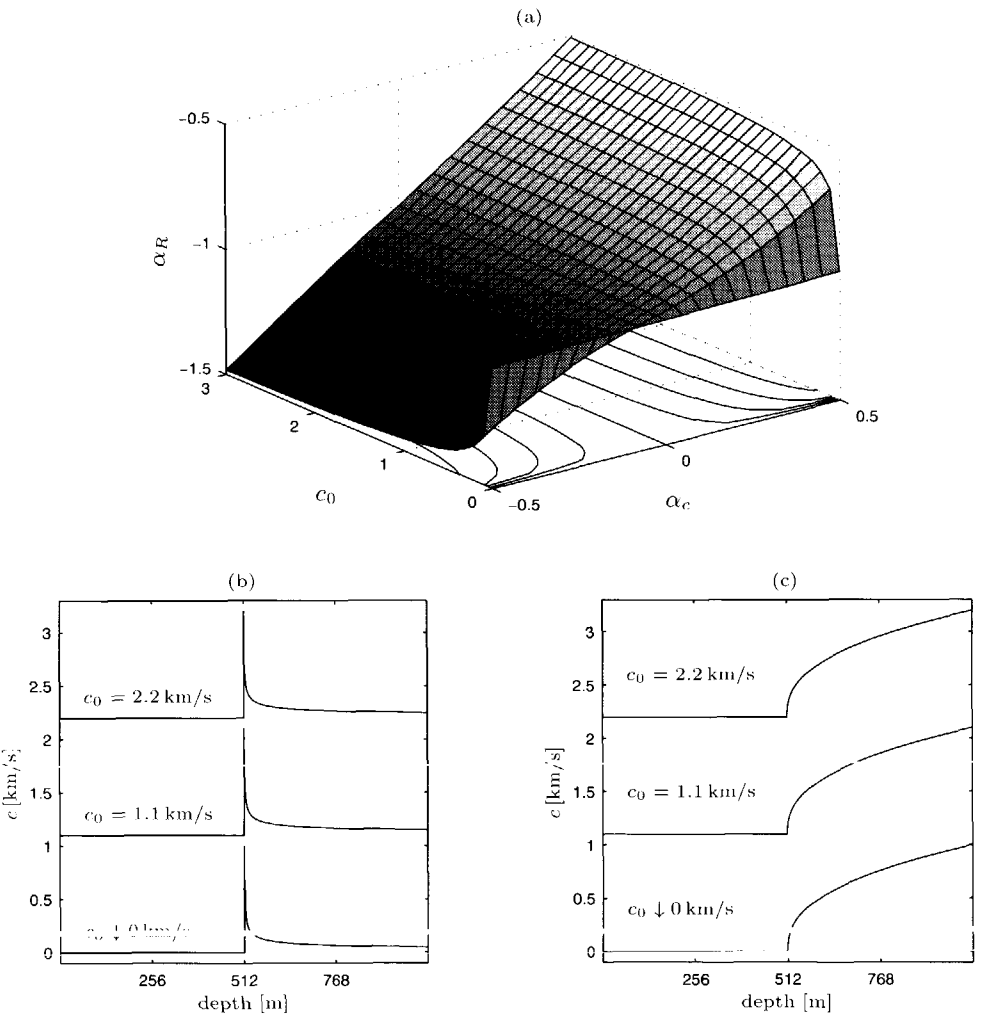


Fig. 5.14 *Influence of the background velocity c_0 on the singularity strength in the reflectivity for normal-incident waves for the situation that $c_\pm > 0$. (a) The singularity strength of $R(x_3)$ at $x_3 = x_{3,0}$ (in the figure denoted by α_R) as a function of the scaling exponent α in the velocity profile (in the figure denoted by α_c) and the background velocity c_0 . Note that in the limiting case that the background velocity goes to zero the scaling in the reflectivity is independent of the singularity strength of the velocity profile. (b-c) Representative examples of the considered distributions for different background velocities, for $\alpha = -0.4$ (b) and for $\alpha = 0.4$ (c).*

Hence, the reflectivity has a local scaling behavior $\alpha - 1$ at $x_3 = x_{3,0}$. Following the derivations in section 5.3.1, it can be concluded that the singularity will also induce a phase change. The phase change is given by Table 5.1.

CASE II : *Positive distortion; peak distortion larger than background velocity*
Mathematically, the following situation is considered

$$c_\pm > 0, \quad \text{and,} \quad c_{pd} > c_0.$$

Because the background velocity c_0 is now smaller than the distortion $c_\pm \chi_\pm^\alpha(\sigma_1, x_3 - x_{3,0})$ around $x_3 = x_{3,0}$, the reflectivity can be written in a series expansion according to

$$\begin{aligned} R(x_3) &= \frac{\pm c_+ \chi_\pm^{\alpha-1}}{2c_\pm \chi_\pm^\alpha} \left(1 + \frac{c_0}{c_\pm \chi_\pm^\alpha}\right)^{-1} \\ &= \frac{\pm c_\pm \chi_\pm^{\alpha-1}}{2c_\pm \chi_\pm^\alpha} \left(1 - \frac{c_0}{c_\pm \chi_\pm^\alpha} + \frac{c_0^2}{c_\pm^2 \chi_\pm^{2\alpha}} - \cdots\right). \end{aligned} \tag{5.71}$$

which is valid for x_3 in the neighborhood of $x_{3,0}$, and for $x_3 > x_{3,0}$ if a right-sided distribution (plus-sign) is considered, and for $x_3 < x_{3,0}$ if a left-sided distribution (minus-sign) is considered. In the limit $c_{pd} \gg c_0$, for example in the case $c_0 \downarrow 0$, the reflectivity takes the following form

$$R(x_3) \simeq \frac{\pm c_\pm \chi_\pm^{\alpha-1}(\sigma_1, x_3 - x_{3,0})}{2c_\pm \chi_\pm^\alpha(\sigma_1, x_3 - x_{3,0})}, \tag{5.72}$$

where the dependency of $\chi_\pm^\alpha$ on $(\sigma_1, x_3 - x_{3,0})$ has been included again, for completeness. It is expected that the scaling behavior of the reflectivity will be -1 at $x_3 = x_{3,0}$, independent of the local scaling behavior of the velocity distribution. This remarkable behavior for $c_{pd} \gg c_0$ has earlier been observed by Wapenaar (1997a).

In Figure 5.14, the derivations are numerically validated with the help of the WTMM analysis of section 5.2.2. The scaling exponent of the velocity profile is $-0.5 \leq \alpha \leq 0.5$, the peak distortion $c_{pd} = 1\,\text{km/s}$, and the background velocity is $0 < c_0 < 3\,\text{km/s}$. The sampling distance is taken $\Delta x_3 = 1\,\text{m}$. Figure 5.14a shows the computed local scaling exponent of $R(x_3)$ at $x_3 = x_{3,0}$, denoted by α_R, as a function of the local scaling exponent α of the velocity profile, denoted by α_c, and as a function of the background velocity. The limits of equation (5.70) of CASE I and equation (5.72) of CASE II are found for $c_0 = 3\,\text{km/s}$, and $c_0 \downarrow 0\,\text{km/s}$, respectively. Figures 5.14b and 5.14c show examples of the considered velocity profiles for $\alpha = -0.4$ and $\alpha = 0.4$, respectively.

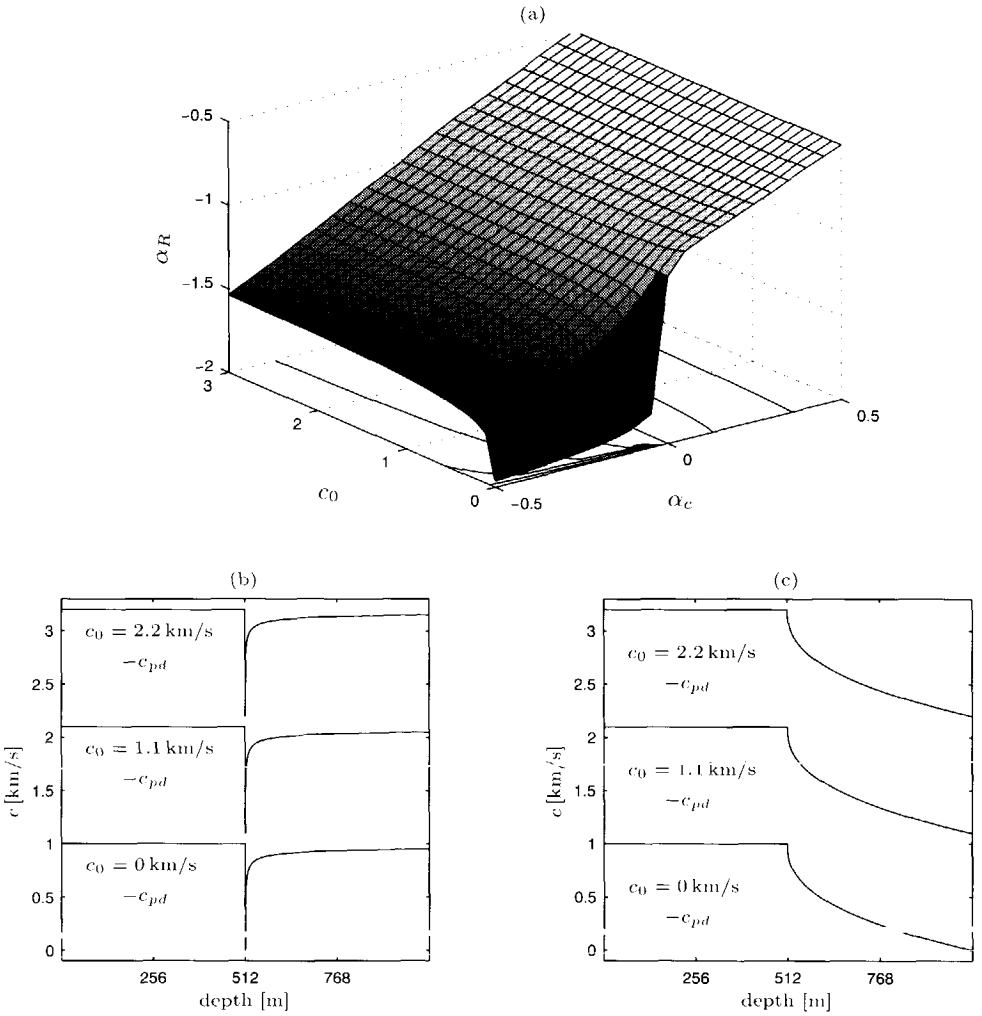


Fig. 5.15 *Influence of the background velocity c_0 on the singularity strength in the reflectivity for normal-incident waves for the situation that $c_\pm < 0$. (a) The singularity strength of $R(x_3)$ at $x_3 = x_{3,0}$ (in the figure denoted by α_R) as a function of the scaling exponent α in the velocity profile (in the figure denoted by α_c) and the background velocity c_0. (b-c) Representative examples of the considered distributions for different background velocities, for $\alpha = -0.4$ and for $\alpha = 0.4$. Note that c_0 is always larger than $-c_{pd}$.*

For negative distortions ($c_\pm < 0$), one is always bound to a background velocity larger than the absolute value of the peak distortion. Negative velocities are generally not considered physical. For negative distortions the following situation applies:

CASE III : *Negative distortions*
Consider

$$c_\pm < 0, \quad \text{and} \quad c_0 > -c_{pd}.$$

Now, the series expansion described in CASE I, equation (5.69), applies, i.e.

$$R(x_3) = \frac{\pm c_\pm \chi_\pm^{\alpha-1}}{2c_0}\left(1 - \frac{c_\pm \chi_\pm^{\alpha}}{c_0} + \frac{c_\pm^2 \chi_\pm^{2\alpha}}{c_0^2} - \cdots\right). \tag{5.73}$$

Hence, one expects in the limit of $c_0 \gg -c_{pd}$, a local scaling behavior $\alpha - 1$ at $x_3 = x_{3,0}$.

Figure 5.15 shows the results of a numerical validation of the derivation for negative distortions. Here, the peak distortion $c_{pd} = -1\,\mathrm{km/s}$; the scaling exponent in the velocity distribution varies according to $-0.5 \leq \alpha \leq 0.5$, and the background velocity varies according to $-c_{pd} < c_0 \leq -c_{pd} + 3\,\mathrm{km/s}$. The derived behavior is confirmed, i.e. a scaling behavior of $\alpha - 1$ is derived. Hence, the scaling exponent of the medium is inherited by the reflected normal-incident wave. The behavior for $\alpha < 0$ and $c_0 \to -c_{pd}$ is not completely understood yet. It is however a situation where the required conditions do not apply. Moreover, here the velocity goes to zero. I will not bother about this part. The limits of CASE I, CASE II, and CASE III are put together in Table 5.2.

Remark 5.3: Reflection at a cusp I
For the cusp with $c_+ = c_-$, the situation is slightly different. In the case that the conditions of the limit of CASE I *or* CASE III *apply, it can be derived from equation* (5.66) *that*

$$R_{cusp}(x_3) = \frac{c_+\left(\chi_+^{\alpha-1} - \chi_-^{\alpha-1}\right)}{2c_0}, \tag{5.74}$$

where I have introduced the reflectivity $R_{cusp}(x_3)$ as the reflectivity of a cusp. Now the results of equation (5.40) *can be utilized. According to that result the reflectivity of a cusp can be expected to depend on the scale parameter as $\sigma^{\alpha-1}$, and to induce a phase change of either $\pi/2$ or $-\pi/2$.*

In the case that the limiting conditions of CASE II *apply, I obtain*

$$R_{cusp}(x_3) = \frac{\chi_+^{\alpha-1} - \chi_-^{\alpha-1}}{2\left(\chi_+^{\alpha} + \chi_-^{\alpha}\right)}, \tag{5.75}$$

	Conditions	$R(x_3)$
CASE I	$c_\pm > 0 \quad c_0 \gg c_{pd}$	$\pm\frac{c_\pm \chi_\pm^{\alpha-1}}{2c_0}$
CASE II	$c_\pm > 0 \quad c_0 \ll c_{pd}$	$\pm\frac{c_\pm \chi_\pm^{\alpha-1}}{2c_\perp \chi_\pm^{\alpha}}$
CASE III	$c_\pm < 0 \quad c_0 \gg -c_{pd}$	$\pm\frac{c_\pm \chi_\pm^{\alpha-1}}{2c_0}$

Table 5.2 *The limiting situations of* CASE I, CASE II, *and* CASE III *as discussed in section 5.5.1.*

which can be expected to have a decay behavior of σ^{-1} *as a function of the scale parameter* σ, *and to carry out an additional phase change of* $\pm\pi/2$. *Numerical experiments carried out for the cusps, with the same parameters as those yielding Figure 5.14 and 5.15 for the one-sided singularities, confirm the above derivations.*

5.5.2 Numerical considerations: $\varphi - \alpha$ diamonds of the reflectivity

The conclusions of the previous section require a further numerical validation. The validation is necessary in order to see whether the regions considered in CASE I, CASE II, and CASE III make sense. These regions can be discriminated because of the enforcement of a spatial scale via equation (5.63). The enforced scale is not necessarily the scale at which the wave is considering the singular structures. For the numerical validation the $(\varphi - \alpha)$-diamond introduced in section 5.3 will be utilized, and a layercode modeling scheme.

The $(\varphi - \alpha)$-diamond has revealed a unique relationship between the signature and strength of a one-sided singularity on the one hand, and the phase change the singularity induces on an analyzing wavelet on the other hand. The diamond, which is derived from a convolution of a wavelet with singularities, can not be transferred directly to the interaction of a seismic wavelet with the same singularities. For that reason the derivations in the previous subsection have been carried out. The derivations show that if CASE I applies, the reflectivity behaves as if it is a homogeneous distributions of degree $\alpha - 1$ with an additional amplitude factor in

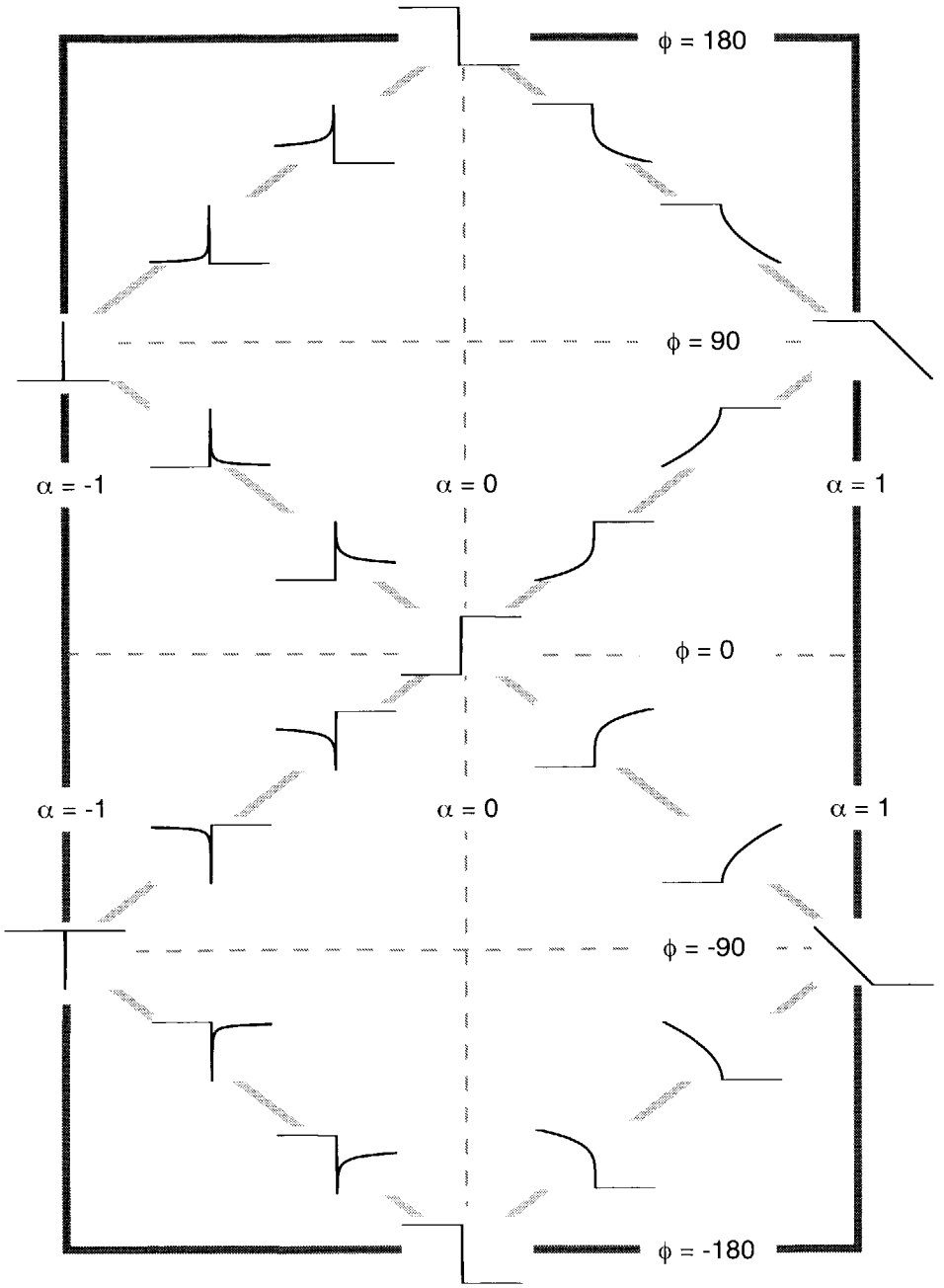


Fig. 5.16 *Expected relation between the velocity $c(x_3)$, and the induced phase change by the corresponding reflectivity. The velocity is given by $c(x_3) = c_0 + c_\pm \chi_\pm^\alpha$ with the scaling exponent α in the range $-1 < \alpha < 1$, $c_0 \gg c_{pd}$ for $c_\pm > 0$, and $c_0 \gg -c_{pd}$ for $c_\pm < 0$.*

front. For the velocity profiles, homogeneous distributions of degree $-1 < \alpha < 1$ have been examined, hence the reflectivity acts as a homogeneous distribution with $-2 < \alpha - 1 < 0$. Consequently, the phase change induced in this situation should be such that the lines of the $(\varphi - \alpha)$-diamond of Figure 5.9 are closely followed. In order to make the relation with the underlying velocity profile manifest, an alternative $(\varphi - \alpha)$-diamond is shown in Figure 5.16. It shows the expected relation between the singularity strength and signature in the velocity profile on the one hand, and the induced phase changes in the reflected data on the other hand[27].

In a layercode modeling scheme the earth is represented by a set of discrete layers, each with a constant density, a constant compressional-wave velocity, and, if applicable, a constant shear-wave velocity. A representation of the homogeneous distributions by a set of discrete layers is intuitively difficult. However, the consistency of the results obtained with the discrete representations in the convolution approach in section 5.2 and 5.3, and the fact that the point to point decay behavior is the actual ruling factor, gives sufficient trust in the numerical approach (see also the discussion in section 5.2.5).

The following procedure is carried out in the numerical experiments.

Procedure 5.3: Validation of $(\varphi - \alpha)$-diamonds

1. *A velocity profile*

$$c(x_3) = c_0 + c_\pm \chi_\pm^\alpha (x_3 - x_{3,0})$$

is constructed according to the implementation discussed on page 159. The subsequent layers are 1 m thick[28]*. The background velocity is chosen such that*

$$\begin{aligned} c_0 &= 2\,km/s && \text{for} \quad c_\pm \geq 0 \\ c_0 &= 2\,km/s - c_{pd} && \text{for} \quad c_\pm < 0. \end{aligned}$$

The degree of the homogeneous distribution varies according to $-1 < \alpha < 1$, which means that the most singular velocity profile is the representation of a δ-distribution, and that the least singular velocity profile is piecewise linear. The factor $c_\pm$ varies such that the peak distortion is in the following range:

$$\begin{aligned} 0.1\,km/s \leq c_{pd} &\leq 10\,km/s && \text{for} \quad c_\pm \geq 0 \\ -0.1\,km/s \geq c_{pd} &> -10\,km/s && \text{for} \quad c_\pm < 0. \end{aligned}$$

[27] Note that Figure 5.16 does not introduce anything new: it is just another way of presenting the diamond-concept.

[28] Note that a specific sampling distance necessarily implies a certain minimum scale.

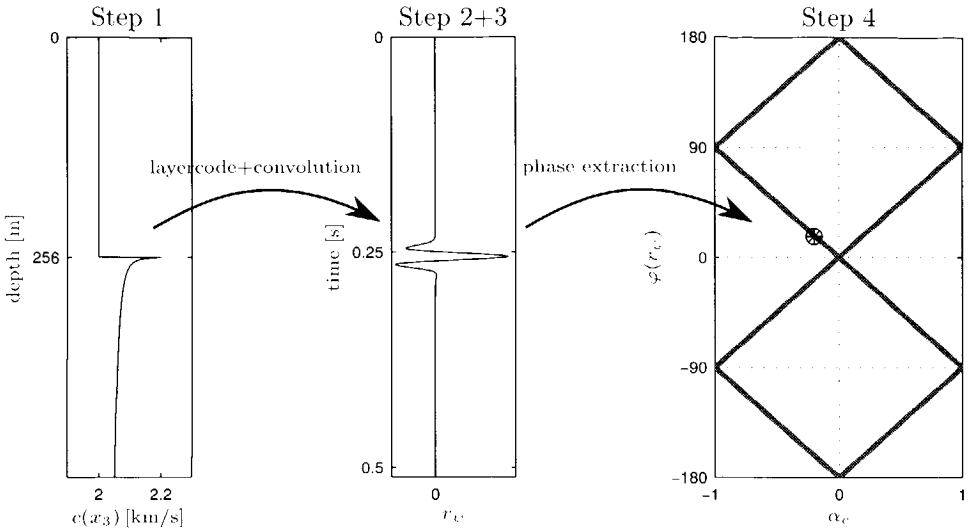


Fig. 5.17 *Example of procedure 5.3 for a single one-sided homogeneous distribution on top of a background velocity $c_0 = 2\,km/s$. Step 1: The homogeneous distribution is of the form $\chi_+^{-0.2}(x_3 - x_{3,0})$ with $x_{3,0} = 256\,m$. The peak distortion is given by $c_{pd} = 0.2\,km/s$. Step 2,3: The computed response for normal incidence is convolved with a Ricker wavelet with a central frequency of 40 Hz. Step 4: Phase estimation; for a relatively small peak distortion the (φ, α)-point falls on top of one of the diamond lines.*

2. *A layercode modeling is carried out for normal-incident waves only, yielding the reflection response $r(p = 0, t)$*[29]*. The temporal sampling rate is 1 ms. The reflection response is convolved with a Ricker wavelet ψ with a central frequency of 40 Hz. The convolved response is represented by $r_\psi(t) = (r * \psi)(t)$.*

3. *The reflection response $r_\psi(t)$ is fed into the phase detection scheme as described in section 5.3.2, Figure 5.11, in order to determine the phase φ of the reflection response.*

4. *The set of points (φ, α) is plotted in a $(\varphi - \alpha)$-diamond.*

The procedure is summarized for $\alpha = -0.2$, $c_+ = 0.2\,\mathrm{km/s}$ (sampling 1 m), $c_- = 0\,\mathrm{km/s}$ in Figure 5.17. The results of procedure 5.3 are presented in Figures 5.18 and 5.19.

The results of passing through the procedure for

$$\begin{cases} c_{pd} = 0.1\,\mathrm{km/s} & c_\pm > 0 \quad (c_0 = 2\,\mathrm{km/s}) \\ c_{pd} = -0.1\,\mathrm{km/s} & c_\pm < 0 \quad (c_0 = 2.1\,\mathrm{km/s}) \end{cases}$$

[29] The ray parameter p will again be omitted.

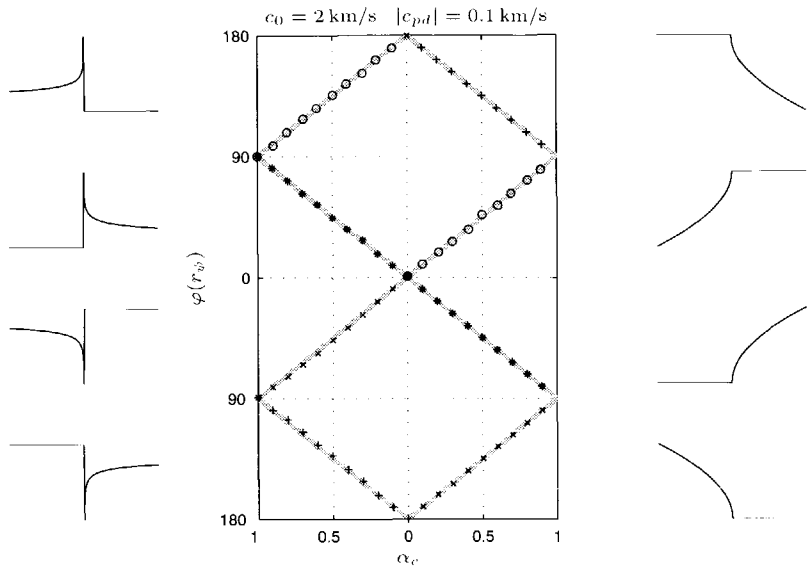


Fig. 5.18 *$(\varphi-\alpha)$-diamond for one-sided singularities. The phase is given in degrees. The background velocity $c_0 = 2\,km/s$. The peak distortion is $0.1\,km/s$. Note, that their is a very clear relationship between the structure, the scaling exponent in the velocity function and the resulting phase change.*

is shown in Figure 5.18. The derived values for φ for a given α correspond with the limits of CASE I, CASE II, and CASE III of the previous subsection, which are summarized in Table 5.2. On the left-hand and right-hand side of Figure 5.18 typical velocity profiles are given.

Figure 5.19 shows the results of passing through procedure 5.3 for

$$\begin{cases} c_{pd} = 1, 2, 5, 10\,\mathrm{km/s} & c_{\pm} > 0 \quad (c_0 = 2\,\mathrm{km/s}) \\ c_{pd} = -1, -2, -5, -10\,\mathrm{km/s} & c_{\pm} < 0 \quad (c_0 = 3, 5, 7, 12\,\mathrm{km/s}). \end{cases}$$

The following conclusions can be drawn:

- Irrespective of the values of c_0 and c_{pd} the step functions are located quite accurately at $(\varphi, \alpha) = (\pm 180, 0)$ and $(0, 0)$. Hence, a step function induces a phase change of 180 degrees for a step downwards, and does not induce a phase change at all for a step upwards, which is not a big surprise.

- The distortions in the shape of the $(\varphi - \alpha)$-diamonds is relatively small for $|c_{pd}| < 2\,\mathrm{km/s}$. The deviations are in the order of 25 percent maximum. Hence, the peak distortion for a sampling distance of $\Delta x_3 = 1\,\mathrm{m}$, and for the

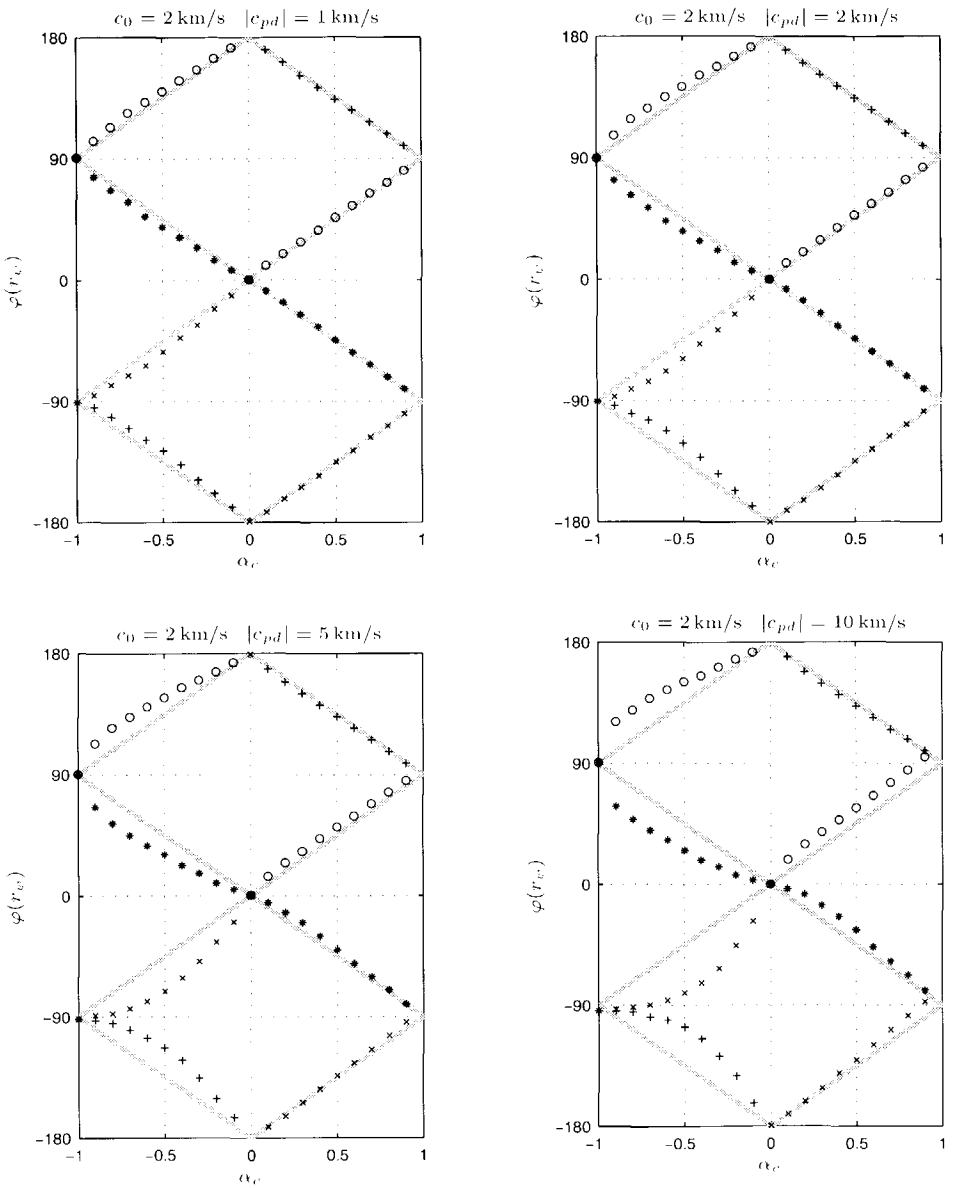


Fig. 5.19 *($\varphi-\alpha$)-diamonds for different values of the peak distortion velocity. The phase is given in degrees. The background velocity is kept constant at $c_0 = 2\,km/s$. For larger peak distortion values the diamonds deviate from the "ideal" diamond structure shown in Figure 5.18 and shown as grey lines on the background. Note, however, that for a peak distortion in the order of the background velocity (top right picture) the distortions are still small.*

chosen frequency range is a reasonable measure for the distinction between the different cases.

- For $|c_{pd}| = 5, 10\,\mathrm{km/s}$ the points (φ, α) move away from their "ideal" positions, especially for negative α. For $\alpha < 0$ and $\varphi > 0$, CASE II of section 5.5.1 applies. In this part the extracted phases move to phases related to a scaling exponent of -1. For $\alpha < 0$ and $\varphi < 0$, the extracted phases tend to $\varphi = -90$ degrees, which can be associated with a scaling exponent of -2 which seems to correspond to the limit observed in Figure 5.15.

- The behavior of the points (φ, α) for $|c_{pd}| = 5, 10\,\mathrm{km/s}$ and $\alpha > 0$ is surprisingly stable. In this part the deviations from the ideal diamond are seemingly smaller than for $\alpha < 0$.

Remark 5.4: Reflectivity at a cusp II
For the reflectivity at a cusp (see Remark 5.3 on page 181) the numerical validation has to be carried out in a slightly different way due to the fact that the phase information does not give any clue. The induced phase change is either $\pi/2$ or $-\pi/2$. I have validated the limits discussed on page 181 with a WTMM *analysis on the layercode-reflection response of the cusps, and concluded that the derived limits are confirmed.*

5.5.3 Reconsideration of the peak distortion

The experiments yielding Figures 5.18 and 5.19 have been carried out for a single central frequency. It is not expected that the results are independent of the central frequency. For a larger central frequency the scale of the probing wave field reduces, due to the relation $\lambda = 2\pi c/\omega$, where λ is the wavelength. A smaller scale is expected to "observe" a larger effective peak distortion. For a Ricker wavelet with a central frequency ranging from 2 to 256 Hz this conjecture has been tested. A velocity profile of the type $c_0 + c_+ \lambda_+^{-0.25}(x_3 - x_{3,0})$ is used with a sampling distance of $\Delta x_3 = 0.25\,\mathrm{m}$, with $c_0 = 2\,\mathrm{km/s}$, and with c_+ chosen such that the peak distortion, $c_{pd} = 0.1, 0.5, 1, 2, 5\,\mathrm{km/s}$, if a sampling distance of $\Delta x_3 = 1\,\mathrm{m}$ would have been chosen. Figure 5.20 shows the results. The conjecture is confirmed, because, for larger central frequencies, the induced phase change moves towards $\varphi = 0$. A δ-distribution does not induce a phase change. Therefore, one can conclude that the singular transition is seen more and more as if it is a δ-distribution. The effect is more pronounced for larger peak distortions.

5.5.4 Peak distortion in well-log data?

The effective singularity of the reflectivity function depends on the relative peak distortion and on the effective size of the probing seismic wave field. Although

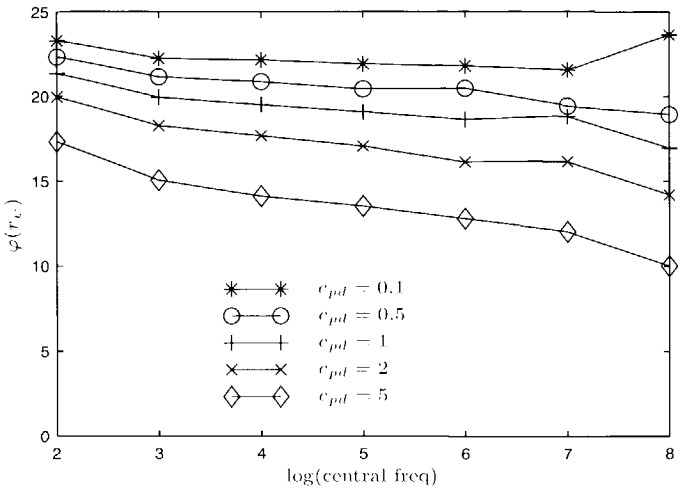


Fig. 5.20 *The phase φ as a function of the central frequency of the seismic wavelet for a velocity with a homogeneous distribution of negative degree defined according to $c_0 + c_+ \chi_+^{0.25}(x_3 - x_{3,0})$, with $c_0 = 2\,km/s$ and the peak distortion velocity varying.*

it can be argued that the data themselves tell what the wave has seen, it is in my opinion important to demarcate the possible $(\varphi - \alpha)$ combinations as much as possible to allow for a practical singularity driven inversion algorithm. Hence, it is important to know whether a normal-incident wave is inheriting the singularity in the medium or not, i.e. whether it is reflected in the situation where the conditions of CASE I or CASE III on the one hand apply, or those of CASE II on the other hand apply. For this purpose, I will consider two acoustic impedances[30] taken from real well-logs offshore Norway, one is in the Midgard field (Courtesy SAGA), and the other is in the North Viking Graben (Courtesy Mobil). More specifically, the relative distortion ϱ of the background acoustic impedance and the distortion with respect to the background will be analyzed, i.e.

$$\varrho = \mathrm{abs}\left[\frac{Z(\sigma_2, x_3) - Z(x_3)}{Z(\sigma_2, x_3)}\right], \tag{5.76}$$

where the acoustic impedance is denoted by $Z(x_3)$, and where $Z(\sigma_2, x_3)$ denotes an estimate of the background acoustic impedance. It is obtained by smoothing the acoustic impedance to a scale σ_2 with a Gaussian function. The results of the procedure for the acoustic impedances of both well-logs are compactly shown in

[30] In the analysis in the first part of this section, the density has been taken constant.

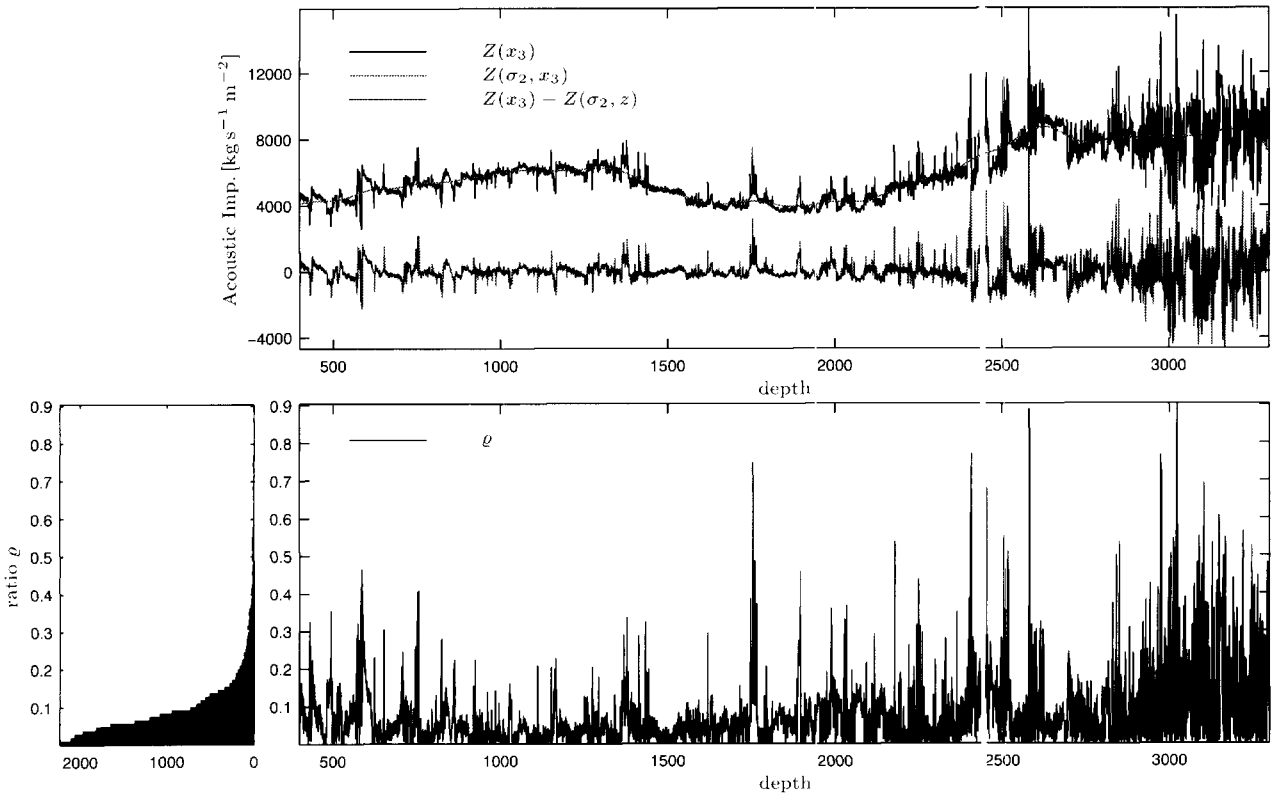


Fig. 5.21 *The relation between the acoustic impedance contrast and the "background" acoustic impedance for a well in the Midgard field, offshore Norway (Courtesy SAGA). The sampling distance is 0.1524 m. The top picture shows the acoustic impedance $Z(x_3)$, the background acoustic impedance $Z(\sigma_2, x_3)$, and their difference. The bottom picture shows the ratio ϱ given by equation* (5.76)*, together with its histogram.*

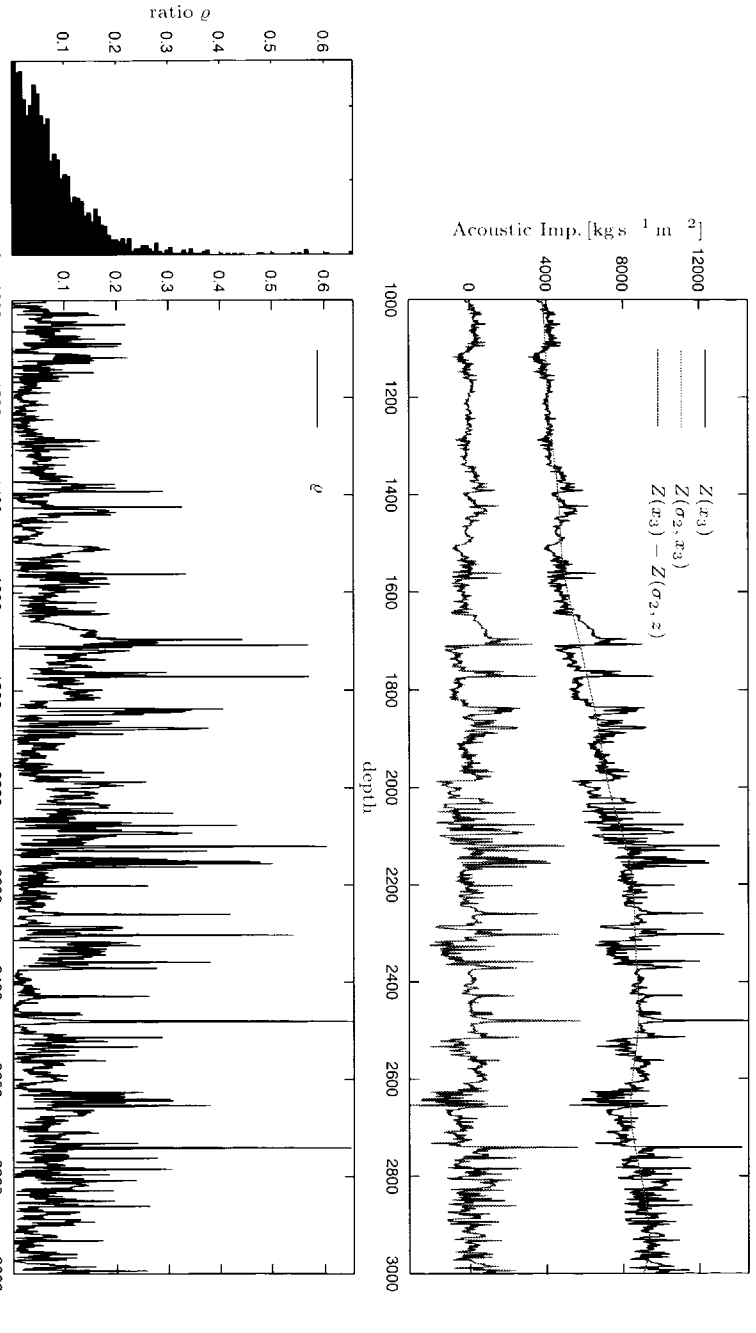


Fig. 5.22 *Relation between the impedance contrast and the "background" impedance for a well in the North Viking Graben (Courtesy Mobil). See caption of Figure 5.21.*

Figure 5.21 and 5.22 for the SAGA log and the Mobil log, respectively. The general trend in ϱ is such that $c_{pd} < c_0$. As a consequence, the analysis of the previous subsections allows to conclude that the scaling behavior of the acoustic impedance is inherited by the seismic reflection data, even for normal-incident waves. If one takes into account the numerical experiments of the previous subsection, then one can conclude that a seismic wave is generally reflected in the region described by the limits of CASE I and CASE III in Table 5.2.

5.6 Angle-dependent reflectivity at isolated singularities

The role of phase in the analysis of amplitude-versus-angle behavior of seismic reflection data is generally limited to the simple situation that the phase is either 0 or π, which means that the reflection amplitude is either positive or negative (see, for example, Verm and Hilterman, 1995). The neglect of the role of phase is also stated by Castagna et al. (1995)[31], who write

"... *the effects of phase changes are not yet readily dealt with* ...".

From the previous section, however, it can be concluded that a one-sided transition other than a step function causes a phase change to the normal-incident wave field. What happens for other than normal-incident waves with respect to the induced phase change and with respect to the amplitude? The aim of this section is to shed light on this question. Whereas for normal-incident waves analytic derivations have been carried out, the approach for oblique-incident waves is of a more phenomenological character[32]. The phenomenons subject to analysis are

- the phase change as a function of ray parameter, and
- the amplitude as a function of ray parameter.

There are two handles at my disposal to link the observed phenomenon to. The first one is the relation between the phase change and the signature of the singularity as revealed in the previous section. The second one is the Zoeppritz model for angle-dependent reflection at a step function. In the sequel, I will commence with the phenomenological analysis in section 5.6.1. I will limit the analysis to one type of structure. On the basis of the observed phenomenon, I will introduce the concept of an equivalent Zoeppritz boundary in section 5.6.2.

[31] Thanks to Aart-Jan van Wijngaarden for pointing out the AVO references to me.

[32] For the situation that the background velocity c_0 is much smaller than the peak distortion velocity c_{pd}, analytic results have been derived (Wapenaar, 1997a).

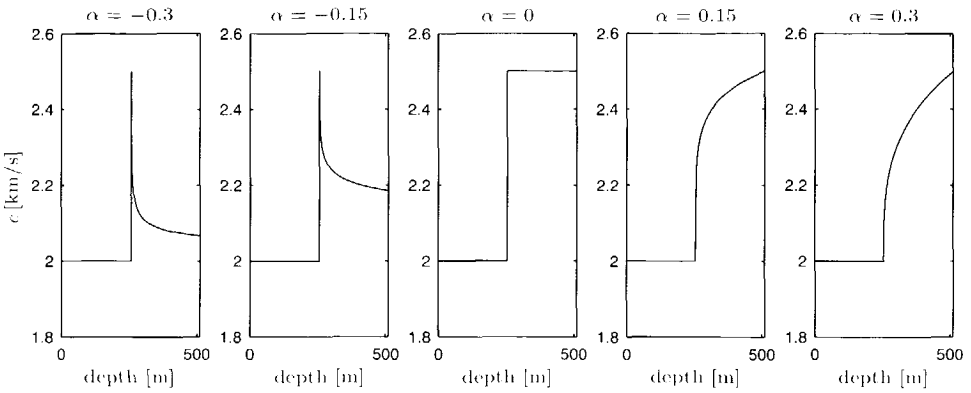


Fig. 5.23 Velocity profiles subject to an AVO analysis.

5.6.1 Phenomenological analysis

Consider a finite-scale representation of a velocity profile given by

$$c(x_3) = c_0 + c_+ \chi_+^{\alpha}(x_3 - x_{3,0}). \tag{5.77}$$

The background velocity is taken $c_0 = 2\,\mathrm{km/s}$. The singularity of the homogeneous distribution is at $x_{3,0} = 256\,\mathrm{m}$. The sampling distance is taken $\Delta x_3 = 1\,\mathrm{m}$. The factor c_+ in front of the homogeneous distribution is taken positive. It is chosen such that the peak distortion velocity takes the value $c_{pd} = 0.5\,\mathrm{km/s}$. For this choice the limit of CASE I applies (see Table 5.2). The scale parameter α in the homogeneous distribution varies from -0.5 to 0.5 with steps of 0.05. Five examples of the considered velocity profiles are shown in Figure 5.23. The middle picture shows the traditionally studied step function.

With the help of an acoustic layercode modeling scheme, the reflected pressure as a function of ray parameter and time has been computed and, subsequently, convolved with a seismic wavelet $\psi_0(t)$, yielding $r_\psi(p,t)$ for all 21 velocity functions $c(x_3)$; the density is taken constant. The general construction of a seismic wavelet has been introduced in procedure 5.2 on page 171; here a Ricker wavelet with a central frequency of 40 Hz is chosen. The response for the step function is shown in Figure 5.24b. The analysis of the amplitude-versus-ray-parameter (AVP) reflection[33] is carried out with the help of the complex trace analysis, yielding one amplitude, $a(p)$, and one phase, $\varphi(p)$, for each ray parameter value p. Normally, the AVP analysis is carried out with the local maxima for each event and for each p-value. The reassessment of the role of the phase of reflection events forced me to consider

[33]Depending on the type of the second parameter considered, AVP is also referred to as amplitude versus offset (AVO) or amplitude versus angle (AVA), see Castagna (1993) for details.

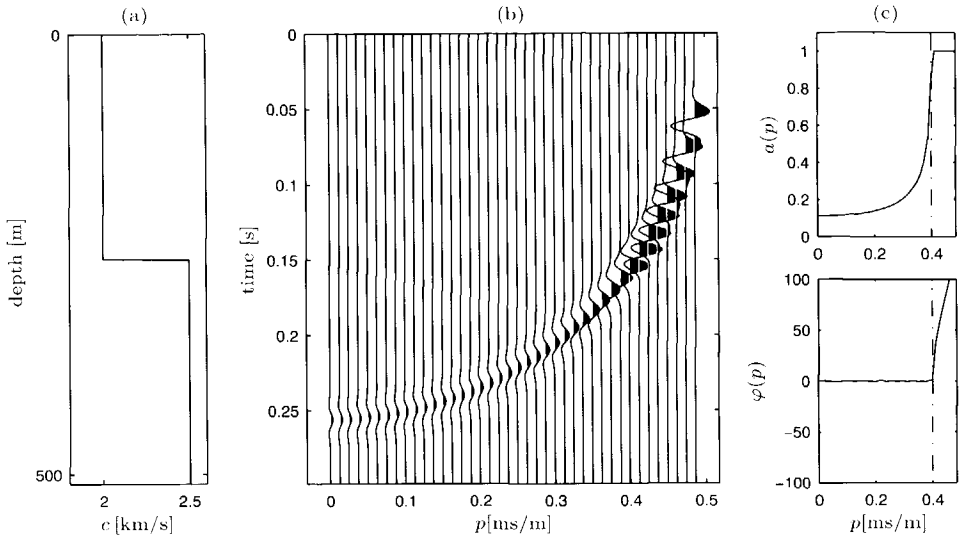


Fig. 5.24 *Reflection at a step function. (a) The velocity profile. (b) The response as a function of ray parameter and intercept time. (c) The amplitude of the maximum of the envelope as a function of ray parameter (top) and the induced phase as a function of ray parameter (bottom).*

the maximum of the envelope, and not the local maximum. The extracted phase and amplitude as a function of the ray parameter are shown in the two pictures of Figure 5.24c. The dash-dotted line in these pictures delimits pre-critical and post-critical reflection. In the pre-critical area the phase is approximately constant. In the post-critical area the amplitude is approximately constant. Note that the analysis can be carried out directly on the reflection data without doing a normal move-out correction or a redatuming step, because the reflection data consist of a single reflection event only. In generally varying media, the effects of propagation through the overburden have to be corrected for first. Note, moreover, that due to the fact that local velocities are quite difficult to estimate, the AVP analysis is normally carried out on time domain data, see the SEG-book on AVO (for example, Castagna, 1993), or the Special Issue on AVO of The Leading Edge (for example, Peddy et al., 1995).

The same analysis has been carried out to the responses of the other $c(x_3)$ functions. For a few velocity profiles the maximum of the envelope, $a(p)$ and the induced phase change, $\varphi(p)$, as a function of ray parameter are shown in Figure 5.25. In this figure the analysis results for the step function are shown as well. The following observations are made:

1. The amplitude, $a(p)$, for the reflection at a step function is larger than the

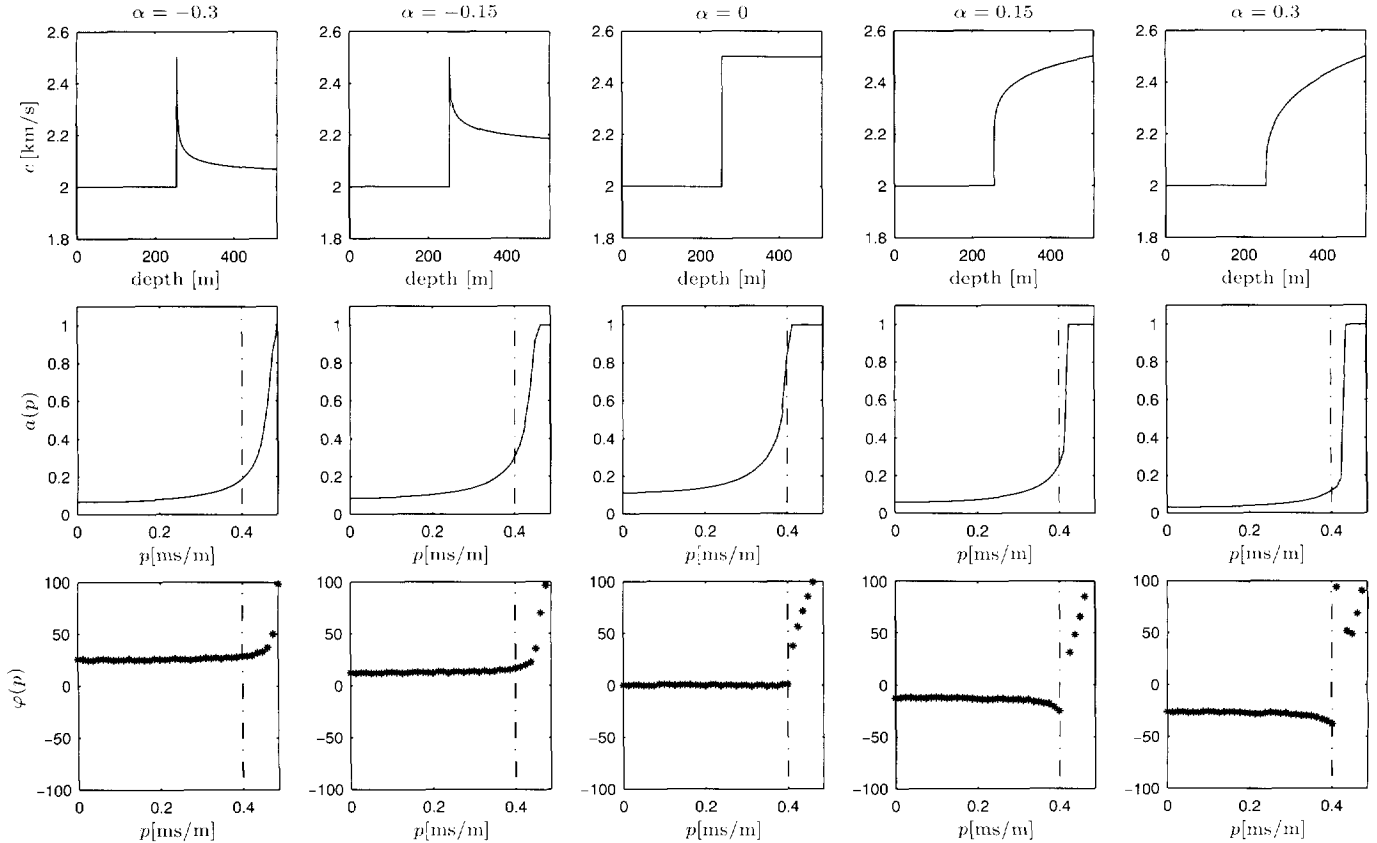


Fig. 5.25 *Analysis of the amplitude and the induced phase of the angle-dependent reflection response. The top row shows the considered velocity profiles. The middle row shows the amplitude of the envelope as a function of ray parameter for the five velocity profiles. The bottom row shows the induced phase as a function of ray parameter for the five velocity profiles.*

amplitudes of the reflection at the other profiles for a fixed peak distortion.

2. In the pre-critical reflection area, the induced phase change does hardly change as a function of ray parameter for any of the velocity profiles. This has two consequences. First, phases at multiple angles can be utilized to stably extract the induced phase change, i.e. the induced phase change can be averaged. Secondly, if the induced phase change is changing as a function of ray parameter, it either means that the wave is post-critical, or, more importantly, that interference is occuring. For larger p-values the vertical wavelength of the probing seismic wave is increasing. If the medium parameters are significantly different for larger vertical wavelengths, interference will play a significant role. Wapenaar et al. (1997) are discussing the interference effect in considerable detail. They especially pay attention to an imaging scheme correcting for the interference effect. The phase change as a function of ray parameter can be utilized to a priori assess the necessity of an adjusted imaging scheme.

3. Both for positive and for negative exponents α the actual post-critical reflection area is progressing to higher p-values.

4. For pre-critical p-values the amplitude behavior can not be distinguished a priori from that of step function, especially if the induced phase change is not considered.

The last observation suggests that in the case the phase is not considered, a Zoeppritz inversion will automatically yield an equivalent Zoeppritz boundary. An equivalent Zoeppritz boundary is considered to be a step function that behaves the same as the singular transition with respect to AVP behavior. In the next subsection, I will briefly elucidate the idea of an equivalent Zoeppritz boundary.

5.6.2 Equivalent Zoeppritz boundaries

Suppose a plane wave impinges on a plane boundary. It is partially reflected and partially transmitted at the boundary. The Zoeppritz equations are relations for the ratio between the displacement of the reflected or transmitted plane wave and the displacement of the plane wave impinging on that plane boundary. More specifically, they express the amplitude ratio between reflected or transmitted shear or compressional waves on the one hand and incident shear waves or compressional waves on the other hand (Aki and Richards, 1980). A linearized Zoeppritz model assumes relatively small angles of incidence (Bortfeld, 1961).

In Zoeppritz inversion the aim is to estimate the relative contrast parameters for the compressional-wave velocity, the shear-wave velocity and the density. Here, I only use the so-called PP-reflection for the Zoeppritz inversion, i.e. the amplitude ratio between a compressional wave reflected from a plane interface and the

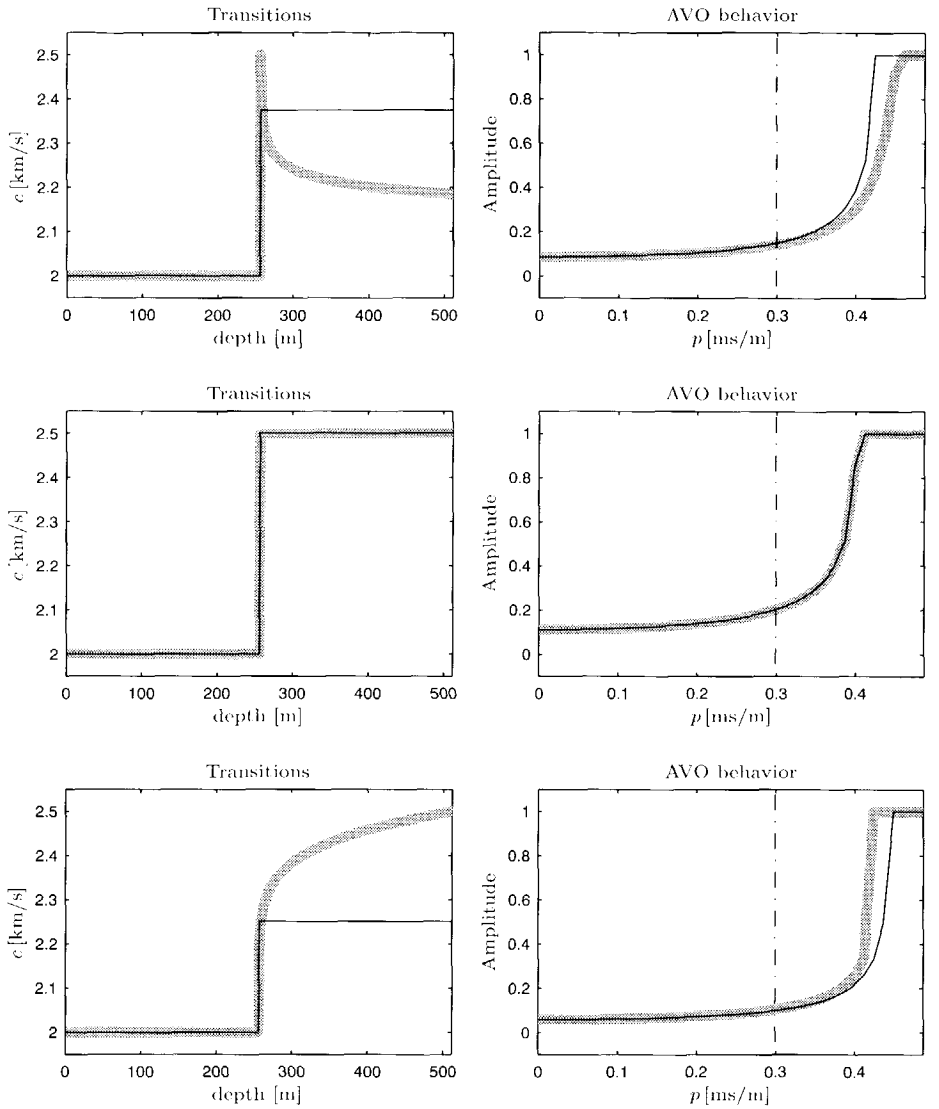


Fig. 5.26 *The left column shows in grey lines the singular structures. The right column shows in grey lines the angle-dependent amplitudes extracted from the numerically computed reflection response of the three singular structures on the left-hand side. The angle-dependent amplitudes have been used as input for a linearized (PP) Zoeppritz inversion. The black lines on the left-hand side denote the estimated step functions which have the same amplitude behavior in the linear region. The amplitude behavior as a function of ray parameter for the estimated step functions is shown in the right column in black lines. The linear region is delineated with the dotted line.*

corresponding compressional wave impinging on that plane interface. A Zoeppritz inversion with the amplitudes extracted in the previous subsection as input, yields an equivalent Zoeppritz boundary. An equivalent Zoeppritz boundary is one which has the same angle-dependent reflection behavior for relatively low p-values. Here, I will not discuss the Zoeppritz inversion scheme in detail. I refer to the cited references for more details.

For three velocity profiles, one with a negative, one with a positive and one with zero scaling exponent (hence a step function), Zoeppritz inversion has been carried out. As a result, a relative contrast is found for the density and for the compressional wave velocity. The shear wave velocity is assumed to be negligible. The estimated density contrast is very small, which is in correspondence with the constant density that has been used for the modeling. The estimated contrast of the compressional wave velocity is shown in Figure 5.26. It is clear that the equivalent Zoeppritz boundary for transitions other than step functions, is smaller than the peak distortion. It makes also manifest that a singular transition can not be distinguished from a step function if one observes it at a single scale or, more importantly, if one neglects the phase information. This qualitative observation will be subject of future quantitative research.

5.7 Singularity driven inversion

The results of sections 5.5 and 5.6 suggest the following Singularity Driven Inversion (SDI) scheme.

Procedure 5.4: Singularity Driven Inversion
A singularity driven inversion consists of the following steps:

1. *Carry out an angle-dependent imaging scheme. Keep the data in the time domain. A good candidate for such an angle-dependent imaging scheme is the bifocal imaging procedure described by Berkhout (1997b).*

2. *Extract the phases and the maxima of the envelope of all reflection events. A reflection event is a point where the envelope reaches a local maximum. The phase provides information on the signature of the singularity.*

3. *Extract the local scaling exponent to estimate the local strength of the singularity. This is done with the help of a wavelet transform modulus maxima analysis. Note that the extraction is done only for those modulus maxima lines the location of which corresponds to the location of a local maximum of the envelope.*

4. *Combine the phase and the local scaling exponent to estimate the type of transition with the help of the $(\varphi - \alpha)$-diamond. For a stable extraction of the*

phase, one can combine the extracted phase for multiple angles of incidence as long as interference does not dominate.

5. *Estimate the equivalent Zoeppritz boundary via the extracted multi-angle amplitudes.*

6. *Combine the results of (4) and (5) to obtain an estimate of the type and size of the singular transition that has caused the reflection.*

Steps (2), (3) and (4) of procedure 5.4, which are regarded especially new, are illustrated with two synthetic examples. In the examples, steps (1), (5) and (6) will not be carried out.

Consider the velocity profile in the top picture of Figure 5.27. It consists of a number of isolated one-sided algebraic singularities of the type given by equation (5.61). The normal-incidence response has been computed with a layercode modeling scheme, and subsequently convolved with a Ricker wavelet with a central frequency of 100 Hz, yielding the response in the middle picture of Figure 5.27. According to the procedure described in section 5.3.2, the induced phase change has been determined. The stars in this picture denote the locations of the computed phases. These are the locations where the envelope of the reflection trace reaches a local maximum. According to the procedure described in section 5.4, the singularity strength of the reflected wave at the location of the stars has been computed. The phase and the singularity strength have been combined to determine in the $(\varphi - \alpha)$-diamond of Figure 5.16 the most likely singular structure that has induced the phase change. The estimated structures of the interfaces have been put together in the bottom picture of Figure 5.27. Note that, without any prior knowledge, the general structure of the subsurface singularities are already very well recovered by using only the normal-incidence reflection response.

A similar procedure has been carried out, but now for a less synthetic velocity and density profile. In a part of well-log 6407/2 of the Midgard field around the target zone at $x_3 = 2500\,m$[34], the response for normal-incident waves has been modeled. The acoustic impedance in that specific part is shown in the top picture of Figure 5.28. The impulse response has been convolved with a Ricker wavelet with a central frequency of 100 Hz. The convolved impulse response is shown in the middle picture of Figure 5.28. The interface structures have been estimated from the convolved impulse response, according to the procedure described in the previous paragraph. The reconstructed interfaces have been put in a single trace as a function of time in the bottom picture of Figure 5.28. Note that the reconstructed interfaces have been plotted at a scale smaller than the scale of the seismic wavelet. If one keeps that in mind one can find quite a few striking similarities between the

[34]The well-log data have been provided by SAGA. See Ekern (1987) for an extensive description of this region.

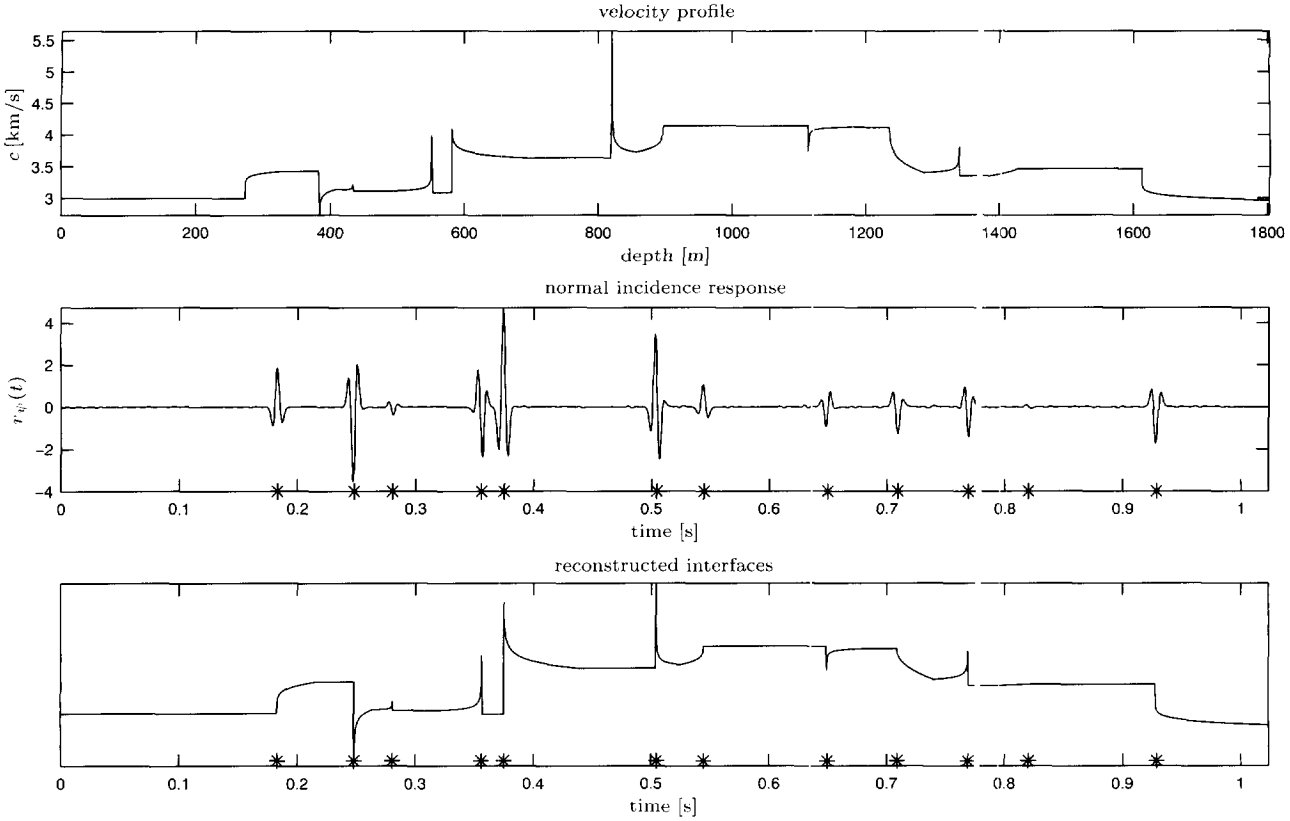


Fig. 5.27 *Top picture: synthetic velocity profile consisting of a set of one-sided singularities of different strength and signature. Middle picture: acoustic pressure response for normal-incident waves. Bottom picture: reconstructed interfaces (still in the time domain).*

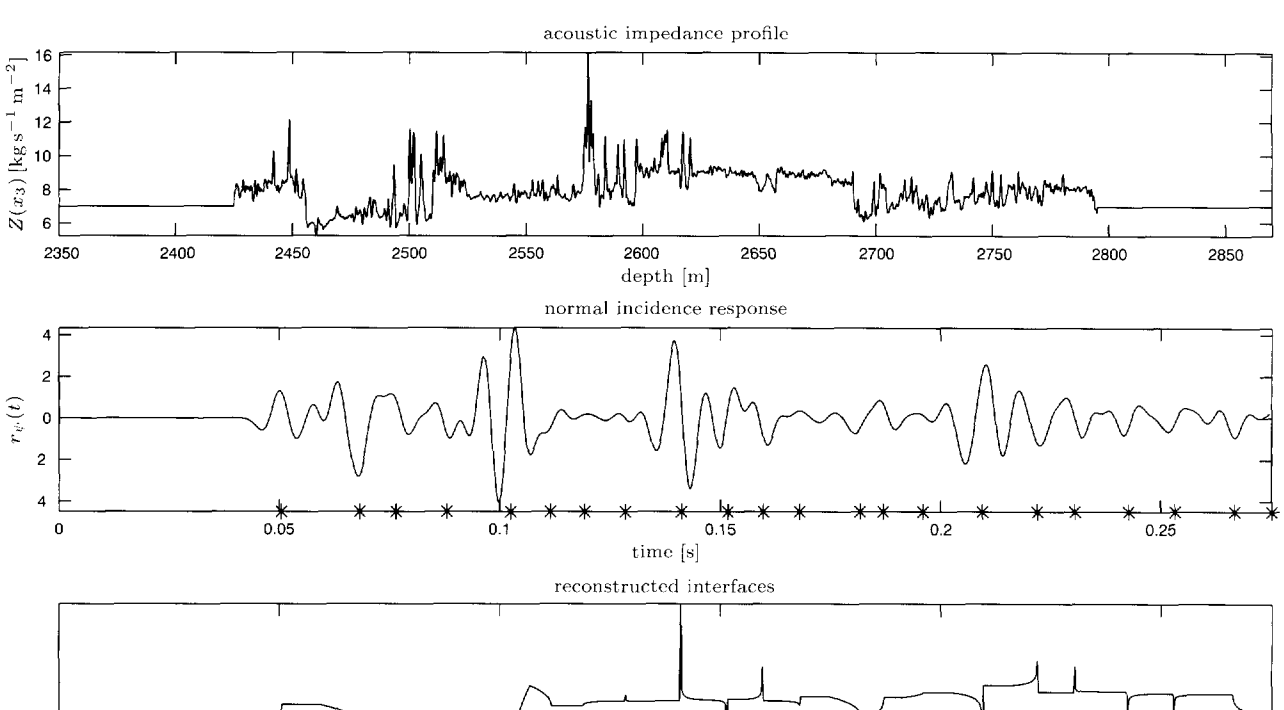


Fig. 5.28 *Top picture: The acoustic impedance $Z(x_3)$ from a part of well 6407/2 around the reservoir region in the Midgard field at $x_3 = 2500\,m$. Middle picture: the response for normal-incident waves. Bottom picture: reconstructed interfaces (still in the time domain).*

acoustic impedance profile and the reconstructed interfaces. Observe, for example, how the first reflecting layers between $x_3 = 2425\,\mathrm{m}$ and $x_3 = 2475\,\mathrm{m}$, and the large peak around $x_3 = 2575\,m$ are reasonably well estimated via the singularity driven inversion procedure.

5.8 Summary

In this chapter, I have discussed two tools that together can be utilized to characterize singularities. The first tool is the wavelet transform modulus maxima analysis to determine the strength of a singularity on the basis of the amplitude decay as a function of the scale parameter. The second tool is a phase analysis to determine the signature of the singularity. For algebraic singularities, the relation between the strength and the signature of a singularity on the one hand and the induced phase change on the other hand has been made transparent with the help of a so-called $(\varphi - \alpha)$-diamond.

A model of the subsurface is proposed consisting of a background velocity with an algebraic singularity on top. Analytic derivations are presented which show how the (algebraic) singularities in the subsurface are inherited by a normal-incident wave reflected at such a singularity. The ratio of the background velocity and the distortion velocity is the ruling factor. If the background velocity is much smaller than the distortion the reflectivity behaves as a δ-distribution, independent of the type of singularity. On the other hand, if the background velocity is larger than the distortion velocity, the reflectivity has a scaling behavior equal to the scaling behavior of the derivative of the velocity function. The analytic derivations are confirmed by numerical layercode-experiments.

The reflection at algebraic singularities for oblique-incident waves is carried out on a phenomenological level. The important observation, here, is the fact that for distortions smaller than the background velocity, the induced phase change does not depend on the angle of incidence (in the pre-critical area). A changing phase as a function of the angle of incidence either refers to post-critical angles of incidence or refers to interference. Another important observation is the fact that if the phase information is neglected, it is a priori difficult to discriminate between the reflection of a wave at a step function and the reflection at more general transitions. This observation induced me to introduce the concept of an equivalent Zoeppritz boundary. An equivalent Zoeppritz boundary is a step function having the same AVP behavior as the singular transition.

The tools presented in the first part of this chapter and the analytic and numerical derivations of the reflectivity at algebraic singularities in the second part motivated me to sketch a singularity driven inversion scheme, allowing for a reconstruction of the singularity type and the singularity strength of the subsurface boundaries. This idea is illustrated with two examples.

Chapter 6

3-D multiscale image analysis of migrated data

The present chapter is a result of a close cooperation with Edo Hoekstra. While he was writing his M.Sc. thesis, Edo Hoekstra and the author implemented wavelet transform based algorithms for 2-D image analysis of migrated data. The results of this procedure have been shown in Hoekstra (1996) and Dessing et al. (1996). For the delineation of stratigraphic features and faults the 2-D algorithm turned out not to be the appropriate tool. Triggered by other successful 3-D algorithms, Hoekstra and the author decided to develop 3-D wavelet transform based algorithms for the full 3-D analysis of migrated data. Those results are presented here and will be submitted for publication.

6.1 Introduction

In the seismic processing sequence shown in Figure 1.3 and discussed in more detail in section 2.6, I have now arrived at the last stage. In this stage the location and size of potential reservoirs have to be estimated from the migrated data, i.e. from the reflectivity $\mathcal{R}$ [1]. The very nature of the migration process implies that boundaries with a significant amount of reflection energy are clearly visible in the migrated data. The location of the reflectors is important, but from the point of view of a seismic interpreter it is even more important to locate stratigraphic features such as fluvial-deltaic systems, tidal channels, beaches, but also to locate faults. These features are essential in delineating the reservoirs. Unfortunately, they are not necessarily present as strong reflectors, think of faults. An experienced interpreter, however, is very well able to extract that "hidden" information in a migrated seismic section.

[1] Note that at this stage of the processing sequence one does not have to bother about wave propagation anymore, since the propagators $\mathcal{W}^{\pm}$ have been removed.

As soon as it comes to 3-D migrated data sets, the task an interpreter is faced with, gets much more involved. Information from multiple slices or multiple sections has to be combined. An automatic extraction procedure would significantly simplify and accelerate the interpretation task. A 3-D image analysis tool is required to make an automatic extraction feasible.

The search for useful 3-D seismic image analysis tools has boomed enormously since the presentation of the results of the coherency cube analysis (Bahorich and Farmer, 1995; Gersztenkorn and Marfurt, 1996; Gersztenkorn et al., 1996). All of a sudden the seismic community realized that an advanced image analysis can reveal significant amounts of "hidden" information. From various points of departure, it is tried to obtain similar or better results (Luo et al., 1996; Steeghs, 1997). Steeghs (1997) shows how a local Radon-Wigner decomposition yields good results. Luo et al. (1996) pursue an edge detection approach, however without using the wavelet transform. Their method has similarities to the wavelet transform approach described in the present chapter. The common denominator of the different methods is the search for local incoherences. The essence of the wavelet transform is its sensitivity to changes or incoherences at multiple scales. The wavelet transform, consequently, seems to be a natural tool to analyze migrated data for the extraction of "hidden" features.

For the analysis of a 3-D migrated volume a 1-D or 2-D wavelet transform does not suffice, although interesting information can be extracted (Hoekstra, 1996; Dessing et al., 1996). Both the 1-D and 2-D wavelet transform do not sufficiently take advantage of the 3-D coherency and 3-D incoherence present in a migrated volume. Hence, the use of the 3-D wavelet transform is essential. The extension of the wavelet transform to more dimensions is mathematically not very difficult (but certainly not unique), especially if a separable wavelet is chosen[2]. Ignorance with respect to the opportunities provided by the 3-D wavelet transform have prevented people from using the wavelet transform in three dimensions.

The utilization of the wavelet transform for multiscale image analysis is from a historical perspective a completely logical step. Mainly, because it has soon been recognized that the wavelet transform is closely related to already existing multiscale image recognition algorithms. For example, Rosenfeld and Thurston (1971), Marr (1982), Witkin (1983), and Burt and Adelson (1983) explicitly used information at multiple scales to analyze amongst others local intensity changes in an image. The relationship between the "older" multiscale image analysis techniques and the wavelet transform has been made especially manifest by the work of Mallat (1989b) and Mallat and Zhong (1992). The mathematics underlying the wavelet transform allow the formalization of a number of existing analysis techniques.

[2]A separable function $f(x_1, x_2, x_3)$ can be written as the multiplication of three functions, for each coordinate one, i.e. $f(x_1, x_2, x_3) = f^1(x_1)f^2(x_2)f^3(x_3)$, where the superscripts denote functions for the $\boldsymbol{i}_1$-, $\boldsymbol{i}_2$- and $\boldsymbol{i}_3$-direction, respectively.

The reason to take recourse to multiscale analysis techniques in image analysis is twofold. First, useful information in an image does not necessarily come to us at one specific scale. In relation to the earth's subsurface, one can think of a salt dome and a tidal channel, which are objects clearly living on different scales. The second reason is originated in the way mammals, and hence the human kind, see. The mammalian retina and visual cortex can be well modeled by a multiscale smoothing function and a set of wavelets (ter Haar Romeny, 1994). Since mammals are very well able to extract features and structures at multiple scales, mimicking the mammal's visual system is not a bad idea.

Before the wavelet transform can be used for image analysis, a number of choices has to be made. First, with respect to the type of transform, and then with respect to the type of smoothing function and wavelet. Concerning the type of wavelet transform I discriminate between the continuous wavelet transform (see sections 2.3 and 2.4, in particular), and the discrete wavelet transform (see section 2.5, in particular). The discrete wavelet transform is very efficient and easily invertible. However, the advantages of the discrete wavelet transform do not weigh up against its disadvantages. The discrete wavelet transform is not translation invariant, and the scale parameter is changing dyadically, i.e. in a rather coarse way. The first disadvantage can be removed by introducing the so-called non-downsampled discrete wavelet transform (Mallat and Zhong, 1992; Saito, 1994) at the cost of some efficiency. In the latter form, the discrete wavelet transform can compete with the continuous wavelet transform. The possibility to continuously vary the scale parameter made me decide to use the continuous wavelet transform in the first stage.

The choice of a particular wavelet and smoothing function pair depends on the type of data to be analyzed, on the number and type of features that have to be extracted, and on the amount of a priori knowledge. Although quite some a priori knowledge of migrated sections is present, the number of different features (both in type and in shape) that might be of interest, is so large, that I have decided to use the ideal pair for an uncommitted image analysis: a Gaussian smoothing function and its first or higher order derivatives as wavelets[3].

[3]The requirement of an uncommitted image analysis can be translated into requirements of causality, homogeneity and isotropy. Causality in a multiscale image analysis means that new details are not created at coarser scales (for 1-D functions). Homogeneity means that all scales are treated in a similar manner, i.e. there is not a preferred scale. Isotropy means of course that there is not a preferred direction. If a multiscale image analysis scheme has to fulfill the three requirements, the multiscale image $I(\tau, \boldsymbol{x})$, with $I(0, \boldsymbol{x})$ the original image, is the solution of the diffusion equation with the original image as initial condition, i.e.

$$\partial_\tau I(\tau, \boldsymbol{x}) = \nabla^2 I(\tau, \boldsymbol{x}) \qquad \text{with} \qquad I(0, \boldsymbol{x}) = I(\boldsymbol{x}). \tag{6.1}$$

The Gaussian function is the Green's function of the diffusion equation. Hence the Gaussian and its derivatives should be used for an uncommitted multiscale image analysis (Koenderink, 1984; ter Haar Romeny, 1994). Note that the relation between the 'time' parameter τ and the scale parameter σ is explained in equations (2.88) and (2.89) on page 43.

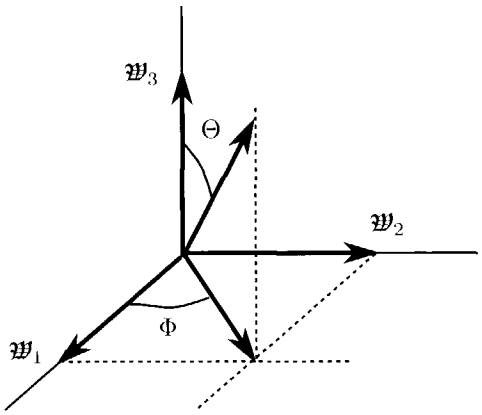


Fig. 6.1 *The interpretation of the angles Φ and Θ. The first is referred to as the azimuth and is given by equation (6.11), the second is referred to as the dip and is given by equation (6.12).*

In the sequel I will first introduce the wavelet transform for data depending on three parameters in section 6.2. The introduced material is illustrated with the help of a simple example: the analysis of a sphere in three dimensions. In section 6.3 the application of the algorithm to a 3-D migrated volume from the Gulf of Mexico is shown. The results will be qualitatively compared with the results from other methods in section 6.4.

6.2 Three-dimensional wavelet transform

Consider a 3-D image $I(\boldsymbol{x})$, with $\boldsymbol{x} = (x_1, x_2, x_3)$. An image analysis procedure is aiming at particular features in the data. The main quest for the image analyzer is to translate the search for a particular feature in a mathematical operation. In the field of scale-space computer vision, this quest is very systematically dealt with through a search for differential invariants (Florack et al., 1994). Here, I will use one of those invariants, namely the amplitude of the gradient in the image $I(\boldsymbol{x})$ at scale $\boldsymbol{\sigma}$, where the 3-D scale parameter is given by $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$. Here, it is explicitly allowed to make use of different scale parameters σ_1, σ_2 and σ_3 for the x_1- ,x_2- and x_3-axis, respectively. The amplitude of the gradient is denoted by $M\{I\}(\boldsymbol{\sigma}, \boldsymbol{x})$. Since the wavelet transform is sensitive to changes, the amplitude of the gradient can be easily derived with the help of the wavelet transform.

In a 3-D space, the wavelet transform takes the form of an inner product of the original image and three wavelets, for each direction one. Using the apparatus

developed in chapter 2 (in particular, section 2.4) I can introduce the L^1-normalized smoothing function $\phi(\boldsymbol{\sigma}, \boldsymbol{x})$ and the first-order spatial derivatives of $\phi(\boldsymbol{\sigma}, \boldsymbol{x})$ as wavelets $\psi_n(\boldsymbol{\sigma}, \boldsymbol{x})$, $n = 1, 2, 3$, according to

$$\phi(\boldsymbol{\sigma}, \boldsymbol{x}) = \frac{1}{\sigma_1 \sigma_2 \sigma_3} \phi\left(\frac{x_1}{\sigma_1}, \frac{x_2}{\sigma_2}, \frac{x_3}{\sigma_3}\right), \tag{6.2}$$

and

$$\psi_n(\boldsymbol{\sigma}, \boldsymbol{x}) = \frac{1}{\sigma_1 \sigma_2 \sigma_3} \psi_n\left(\frac{x_1}{\sigma_1}, \frac{x_2}{\sigma_2}, \frac{x_3}{\sigma_3}\right) \tag{6.3}$$

$$= \sigma_n \partial_n \phi(\boldsymbol{\sigma}, \boldsymbol{x}), \tag{6.4}$$

respectively. Remember that ∂_n refers to a derivative with respect to the x_n-variable, and note that the summation convention for Roman subscripts does not apply. By introducing the wavelet vector $\boldsymbol{\psi}(\boldsymbol{\sigma}, \boldsymbol{x})$ as

$$\boldsymbol{\psi}(\boldsymbol{\sigma}, \boldsymbol{x}) = (\psi_1, \psi_2, \psi_3)^T, \tag{6.5}$$

the wavelet transform $\vec{\mathcal{W}}\{I\}$ of the image $I(\boldsymbol{x})$ with respect to the wavelet $\boldsymbol{\psi}(\boldsymbol{\sigma}, \boldsymbol{x})$ takes the form

$$\vec{\mathcal{W}}\{I\} = \begin{pmatrix} \mathcal{W}_1\{I\} \\ \mathcal{W}_2\{I\} \\ \mathcal{W}_3\{I\} \end{pmatrix} = (I * \boldsymbol{\psi})(\boldsymbol{\sigma}, \boldsymbol{x}), \tag{6.6}$$

where $\vec{\mathcal{W}}\{I\}$ is a shorthand notation for $\mathcal{W}\{I, \boldsymbol{\psi}\}(\boldsymbol{\sigma}, \boldsymbol{x})$, which has been introduced in equation (2.75) for a 1-D wavelet transform. Each of the elements of the vector $\vec{\mathcal{W}}\{I\}$ can be written in an explicit form by (cf. equation 2.85)

$$\mathcal{W}_n\{I\} = \int I(x_1', x_2', x_3') \frac{1}{\sigma_1 \sigma_2 \sigma_3} \psi_n\left(\frac{x_1' - x_1}{\sigma_1}, \frac{x_2' - x_2}{\sigma_2}, \frac{x_3' - x_3}{\sigma_3}\right) \mathrm{d}x_1' \, \mathrm{d}x_2' \, \mathrm{d}x_3'. \tag{6.7}$$

Due to relation (6.4), it is also possible to write the wavelet transform as the gradient of a smoothed image, according to

$$\mathcal{W}_n\{I\} = -\sigma_n \partial_n (I * \phi) = -\sigma_n \partial_n I(\boldsymbol{\sigma}, \boldsymbol{x}). \tag{6.8}$$

The smoothed image $I(\boldsymbol{\sigma}, \boldsymbol{x})$ is given by the inner product of $I(\boldsymbol{x})$ with the smoothing function $\phi(\boldsymbol{\sigma}, \boldsymbol{x})$, according to (cf. equation 2.84)

$$I(\boldsymbol{\sigma}, \boldsymbol{x}) = \int I(x_1', x_2', x_3') \frac{1}{\sigma_1 \sigma_2 \sigma_3} \phi\left(\frac{x_1' - x_1}{\sigma_1}, \frac{x_2' - x_2}{\sigma_2}, \frac{x_3' - x_3}{\sigma_3}\right) \mathrm{d}x_1' \, \mathrm{d}x_2' \, \mathrm{d}x_3'. \tag{6.9}$$

The quantity $\mathcal{W}_n\{I\}$ expresses the amount of change in the $\boldsymbol{i}_n$-direction at the scale $\boldsymbol{\sigma}$ of the image I. The wavelet transform vector $\vec{\mathcal{W}}\{I\}$ is also referred to as the gradient vector.

The amplitude of the gradient vector, $M\{I\}(\boldsymbol{\sigma}, \boldsymbol{x})$, is now given by the norm of the vector $\vec{\mathfrak{W}}\{I\}$, i.e.

$$M\{I\}(\boldsymbol{\sigma}, \boldsymbol{x}) = |\vec{\mathfrak{W}}\{I\}| = \sqrt{\mathfrak{W}_1\{I\}^2 + \mathfrak{W}_2\{I\}^2 + \mathfrak{W}_3\{I\}^2}. \tag{6.10}$$

Associated to the amplitude of the gradient is the direction of the gradient. In three dimensions the direction is given by two angles. The angle $\Phi\{I\}$ is referred to as the azimuth, and is given by

$$\Phi\{I\} = \operatorname{atan}\left(\frac{\mathfrak{W}_2\{I\}}{\mathfrak{W}_1\{I\}}\right), \tag{6.11}$$

and the angle $\Theta\{I\}$ is referred to as the dip, and is given by

$$\Theta\{I\} = \operatorname{atan}\left(\frac{\mathfrak{W}_3\{I\}}{\sqrt{\mathfrak{W}_1\{I\}^2 + \mathfrak{W}_2\{I\}^2}}\right). \tag{6.12}$$

The meaning of the angles $\Phi\{I\}$ and $\Theta\{I\}$ is illustrated in Figure 6.1. In addition to its importance in its own sake, the angle information can be used to extract the modulus maximum points or edges. The edge detection can be seen as a more-dimensional generalization of the wavelet transform modulus maximum extraction dealt with in the previous chapter. In two dimensions, edge detection has been discussed by quite some authors (for example, Canny, 1986; Mallat and Zhong, 1992). In three or more dimensions the edge detection algorithms get more involved, but are conceptually not different.

The idea of a 3-D wavelet transform is illustrated with the help of a synthetic example, namely a sphere as shown in Figure 6.2. The application of the wavelet transform in the three directions according to equation (6.6) yields the amount of change in the three directions. It is clear that the result is in accordance with the expectations (Figure 6.2, bottom row). By applying either of the operations (6.10), (6.11), or (6.12), the amplitude of the gradient in the sphere, the azimuth and the dip of the changes can be extracted, respectively.

6.2.1 Implementation aspects

Although the computer memory size of a migrated cube is generally much smaller than the corresponding surface seismic data set, it can still easily occupy a GigaByte or more of computer memory. An efficient algorithm which allows for a fast, almost real time, feature extraction in a cube of that size can significantly contribute to a reduced interpretation time.

A first requirement for an efficient algorithm is a separable smoothing function and wavelet. A separable function allows for a treatment of all coordinates separately. A separable smoothing function takes the following form

$$o(\boldsymbol{\sigma}, \boldsymbol{x}) = o^1(\sigma_1, x_1)o^2(\sigma_2, x_2)o^3(\sigma_3, x_3), \tag{6.13}$$

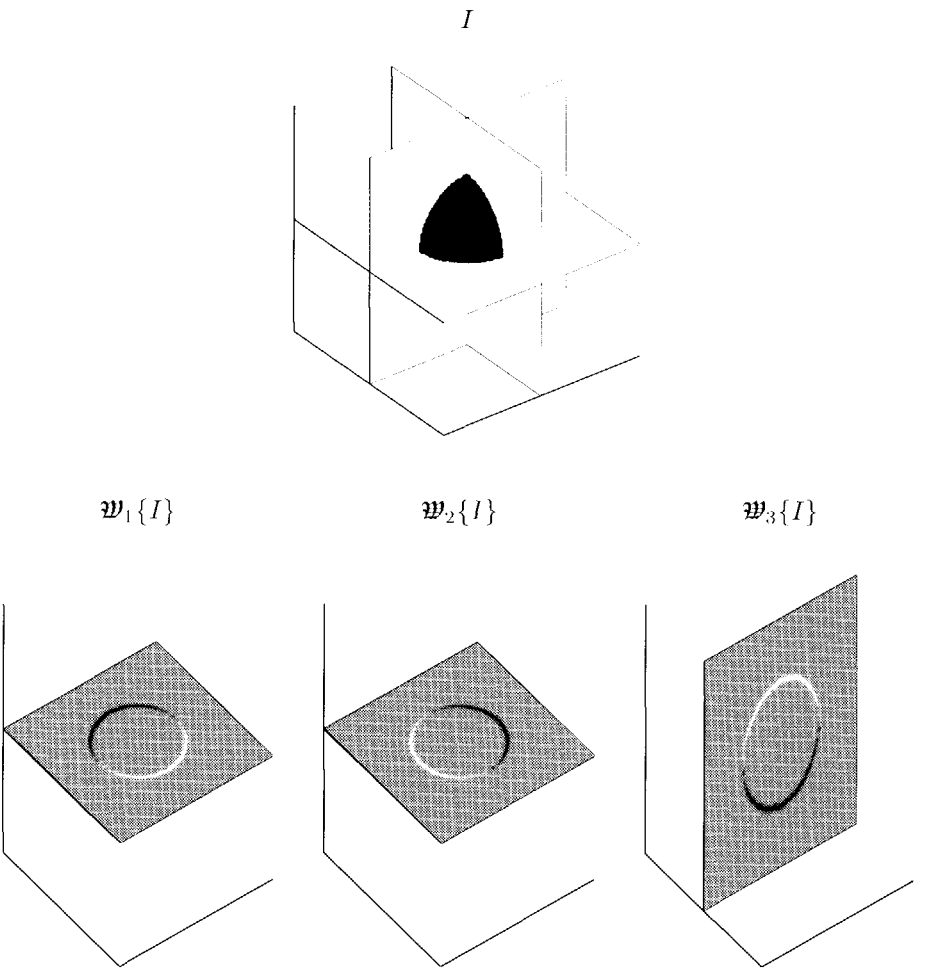


Fig. 6.2 *A sphere I in three dimensions on top. The 3-D wavelet transform of the sphere in the bottom row. From left to right: $\mathfrak{W}_1\{I\}$, $\mathfrak{W}_2\{I\}$, and $\mathfrak{W}_3\{I\}$, which are the wavelet transform in the $\boldsymbol{i}_1$-, $\boldsymbol{i}_2$-, and $\boldsymbol{i}_3$-direction respectively. It can be concluded that the three transforms are sensitive to changes in the three different directions.*

where the superscript n denotes the smoothing function for the axis x_n. The corresponding separable wavelets $\psi_n(\boldsymbol{x})$, $n = 1, 2, 3$, take the following form

$$\psi_1(\boldsymbol{\sigma}, \boldsymbol{x}) = \psi^1(\sigma_1, x_1)\phi^2(\sigma_2, x_2)\phi^3(\sigma_3, x_3) \tag{6.14}$$

$$\psi_2(\boldsymbol{\sigma}, \boldsymbol{x}) = \phi^1(\sigma_1, x_1)\psi^2(\sigma_2, x_2)\phi^3(\sigma_3, x_3) \tag{6.15}$$

$$\psi_3(\boldsymbol{\sigma}, \boldsymbol{x}) = \phi^1(\sigma_1, x_1)\phi^2(\sigma_2, x_2)\psi^3(\sigma_3, x_3), \tag{6.16}$$

where $\psi^n(\sigma_n, x_n)$ denotes the wavelet for the axis x_n. The wavelet ψ_n and the smoothing function ϕ are related according to equation (6.4). Note the difference between the superscripted and subscripted wavelets. In the introductory section of the present chapter, it has been argued that an uncommitted image analysis necessitates the use of the Gaussian as smoothing function, and, consequently, one of its derivatives as wavelet. Hence, the L^1-normalized smoothing function $\phi^n(\sigma_n, x_n)$ is given by

$$\phi^n(\sigma_n, x_n) = \frac{1}{\sigma_n\sqrt{2\pi}} \mathrm{e}^{-x_n^2/(2\sigma_n^2)}, \tag{6.17}$$

and the wavelet $\psi^n(\sigma_n, x_n)$ is given by its first derivative with respect to x_n, according to

$$\psi^n(\sigma_n, x_n) = \sigma_n \partial_n \phi^n(\sigma_n, x_n). \tag{6.18}$$

A second requirement for an efficient algorithm is the limitation of the number of scale values $\boldsymbol{\sigma}$ at which $I(\boldsymbol{\sigma}, \boldsymbol{x})$ is computed. A limitation is computationally necessary, despite the fact that a thorough integration of information obtained at multiple scales is expected to yield superior results (Koenderink, 1984; ter Haar Romeny, 1994). The type of data and the type of features of interest determine, probably via a trial-and-error cycle, the scale that reveals the information of interest[4].

For a single scale and for a Gaussian smoothing function and for corresponding wavelets, the computation of the wavelet transform and its derived quantities takes roughly $2N \log N$ operations, where N is the total number of data in the 3-D volume. Here, the wavelet transform is carried out with via the Fourier domain. The efficiency of the algorithm can be improved by choosing at a particular scale spatial filters that closely approximate the action of the Gaussian and its derivatives. The biorthogonal cubic spline filters described by Mallat and Zhong (1992) seem to be excellent candidates, which can be concluded from the results obtained on 2-D seismic sections with those filters (Hoekstra, 1996). Such an implementation will lead to a $O(N)$-algorithm. Let me summarize the algorithm.

[4] A refinement of the algorithm which is still computationally reasonably easy to accomplish, can be obtained by combining information at two or three scales. For example, the global dip and azimuth information on a coarse scale, combined with stratigraphic features and faults extracted on a fine scale.

Procedure 6.1: 3-D image analysis with the wavelet transform

1. *Choose a smoothing function and three associated wavelets for each direction one.*
2. *Choose an appropriate scale* $\boldsymbol{\sigma} = (\sigma_1, \sigma_2, \sigma_3)$.
3. *Compute the wavelet transform in all directions for that particular scale, resulting in* $\mathfrak{W}_1\{I\}$, $\mathfrak{W}_2\{I\}$, *and* $\mathfrak{W}_3\{I\}$.
4. *Compute a differential invariant and the associated quantities of interest. Here, the amplitude of the gradient, together with the dip and the azimuth.*
5. *Combine the information from multiple slices to extract the faults and stratigraphic features of interest.*

6.3 Application to a real data set

The present section is devoted to the opportunities of the presented method for the analysis of a real 3-D migrated volume. The migrated volume has been provided by Geco-Prakla[5]. It is a time migrated data set from the Gulf of Mexico. It is located on the continental shelf, west and adjacent to the present-day Mississippi delta. The data set has been used for the geological evaluation of the coherency cube analysis (Bahorich and Farmer, 1995; Haskell et al., 1995; Nissen et al., 1995; Gersztenkorn and Marfurt, 1996; Gersztenkorn et al., 1996). The general structure of the volume is shown in Figure 6.3. Due to the fact that a time migrated data set is dealt with, the third coordinate x_3 is replaced here by the time coordinate t. The fat lines denote the vertical cross sections and the time slices that will be depicted in subsequent figures. Figure 6.4 shows the vertical cross sections for constant x_2, namely Y_1 and Y_2. Figure 6.5 shows two vertical cross section for constant x_1, namely X_1 and X_2. The structure is generally quite flat except for the salt dome region. The whole volume is criss-crossed by faults. They are especially clear in vertical cross section Y_1. For example, at $x_1 = 6\,\mathrm{km}$ a normal fault is cutting through the entire section. Above and below the layer of reworked material at 900 ms, i.e. a submarine canyon in-fill (Steeghs, 1997), lots of channels from the former Mississippi delta are present[6]. These channels are, however, not clearly visible in the vertical cross sections.

Procedure 6.1 has been carried out with the first derivative of a Gauss function in all three directions at a single small scale, according to equations (6.6) and (6.7). At a coarse scale the general structure in terms of the dip and the azimuth is highlighted better, but that information is not of interest to me in this chapter. The

[5]Amoco Production Company is acknowledged for making the data available.

[6]Ir. W. Peet is acknowledged for interpreting a large portion of the 3-D volume and for pointing out a number of interesting stratigraphic features.

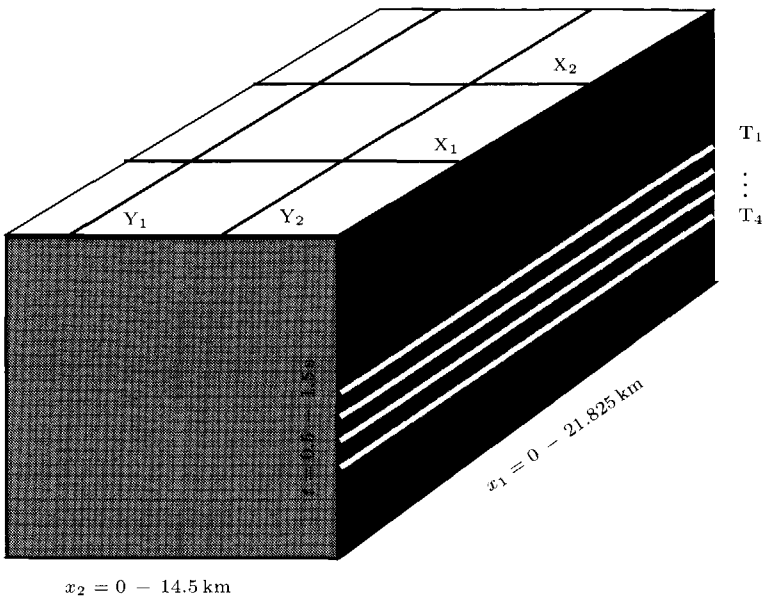


Fig. 6.3 *The general structure of the analyzed 3-D migrated volume $I(x_1, x_2, t)$. Note that a time migrated data volume is dealt with, therefore the coordinate x_3 is replaced by t. In Figure 6.4, the vertical cross sections Y_1 at $x_2 = 1\,km$ and Y_2 at $x_2 = 9\,km$ are shown. In Figure 6.5, the vertical cross sections X_1 at $x_1 = 5\,km$ and X_2 at $x_1 = 14\,km$ are shown. The results of the 3-D wavelet transform analysis is specifically illustrated with the help of the four time slices T_1, T_2, T_3 and T_4 at $t = 1284\,ms$, $1292\,ms$, $1316\,ms$ and $1348\,ms$, respectively.*

differential invariant computed here is the amplitude of the gradient for all points in the volume. The weighted addition of multiple slices is the final step.

Figure 6.7 up to and including Figure 6.14 give the results of the method for the four time slices T_1, T_2, T_3 and T_4. For comparison the original time slices are given as well. A number of features that become clearly visible by applying the 3-D wavelet transform has been highlighted with arrows. Horizontal arrows denote channels and vertical arrows denote faults. It is clear that the method significantly helps in extracting stratigraphic features and faults. Especially the channels are very well extracted. In the captions of the figures a number of interesting features are discussed in more detail. Note that the method can still be improved with respect to the separation of dip information and stratigraphic information. This can be concluded from the relatively high amplitudes in the processed time slices at the flanks of the salt dome.

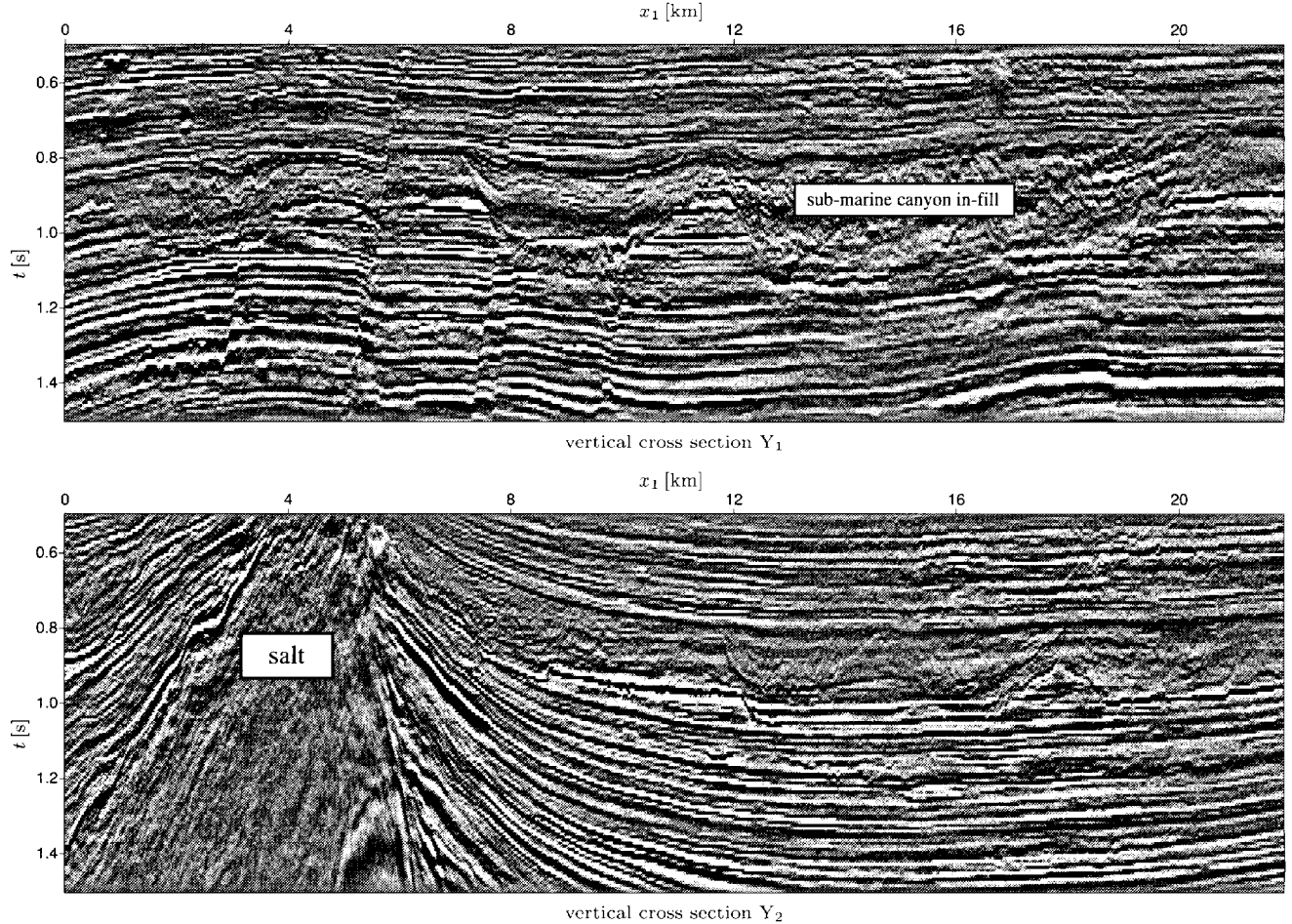


Fig. 6.4 *The two vertical cross sections* Y_1 *at* $x_2 = 1\,km$ *and* Y_2 *at* $x_2 = 9\,km$. Y_1 *clearly shows a number of faults, and the reworked material approximately at* $t = 900\,ms$. Y_2 *clearly shows the lifting influence of the salt dome. Data Courtesy Geco-Prakla.*

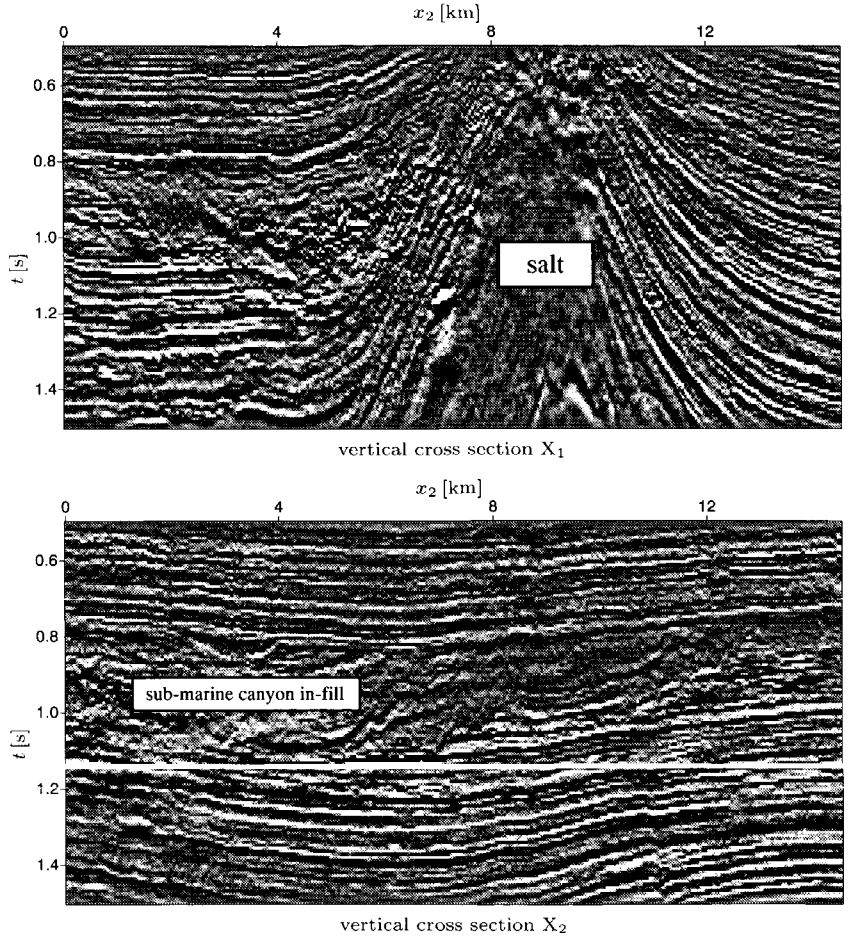


Fig. 6.5 *The two vertical cross sections* X_1 *at* $x_1 = 5\,km$ *and* X_2 *at* $x_1 = 14\,km$. X_1 *especially highlights the salt dome.* X_2 *shows a slight syncline structure and the reworked material.*

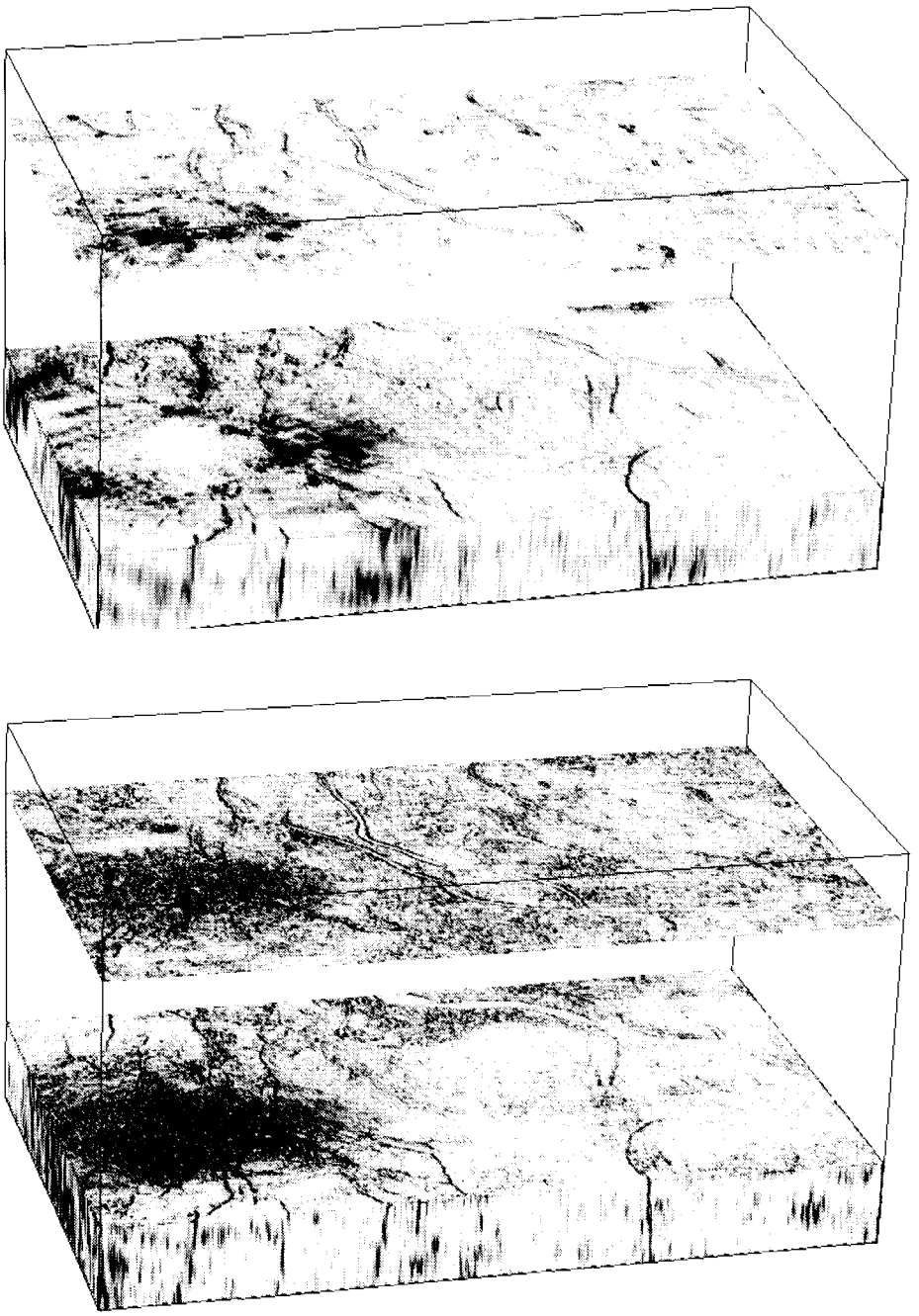


Fig. 6.6 *The application of different differential operators allows for a flexible feature-dependent processing of a migrated cube. The top picture shows in a 3-D view amplitude changes; the bottom picture shows local dips.*

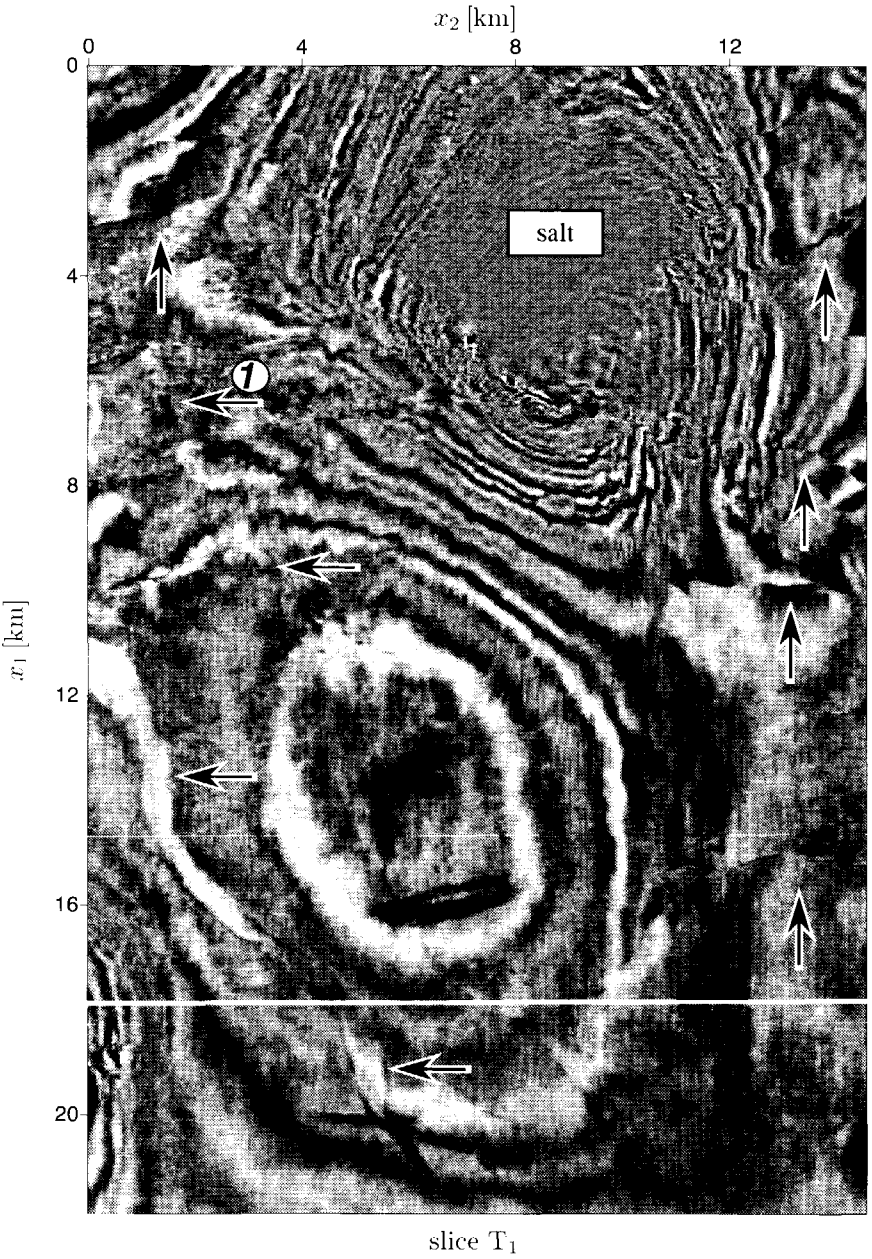


Fig. 6.7 *Time slice at $t = 1284$ ms from the unprocessed migrated volume. To be compared with the same time slice from the processed volume in Figure 6.8.*

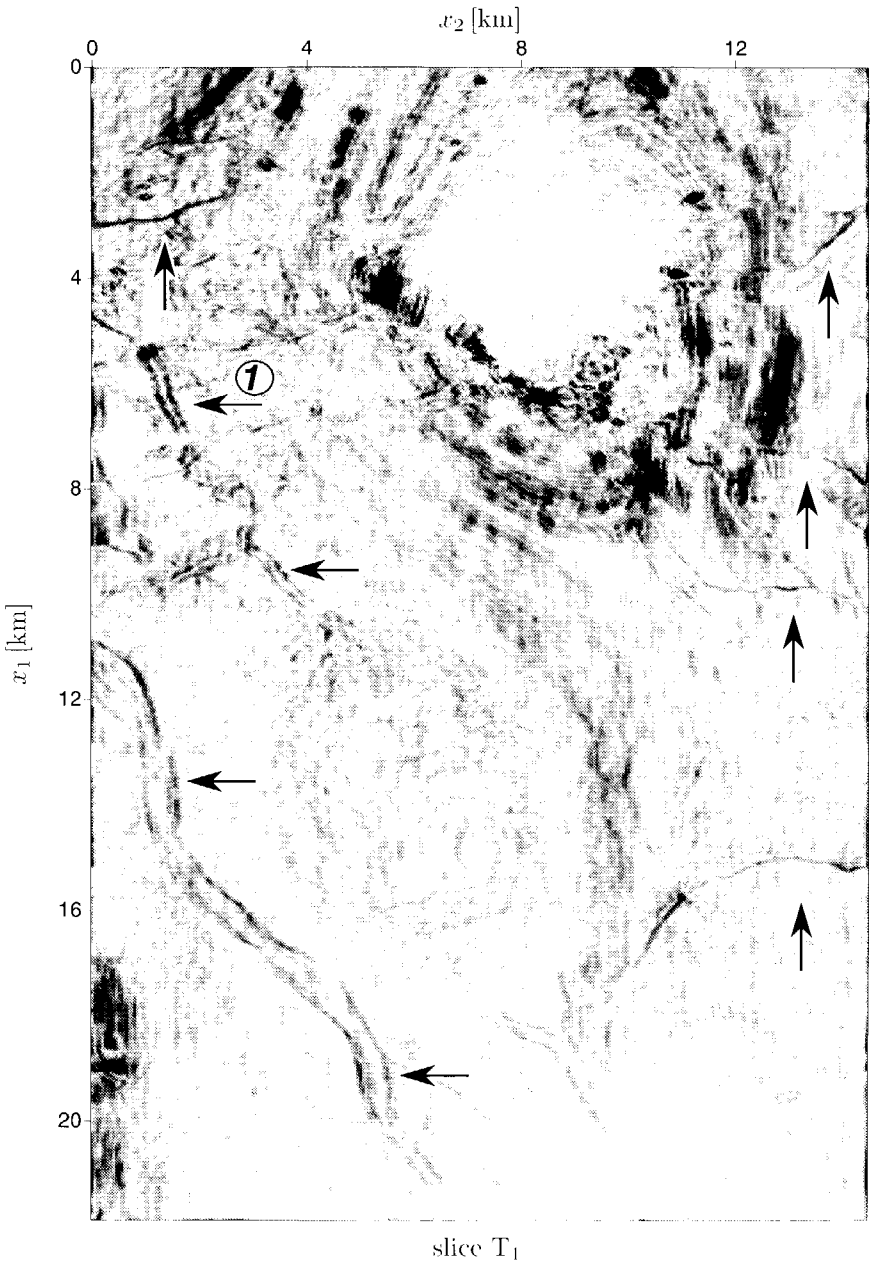


Fig. 6.8 *Time slice at $t = 1284\,ms$ from the processed migrated volume. The arrows denote interesting stratigraphic features and faults. Vertical arrows point to faults, horizontal arrows to channels. Whereas a number of features can be easily identified with the benefit of the hindsight in the time slice from the unprocessed volume, others seem to appear out of the blue. For example, the channel denoted by ① is not visible in the original time slice.*

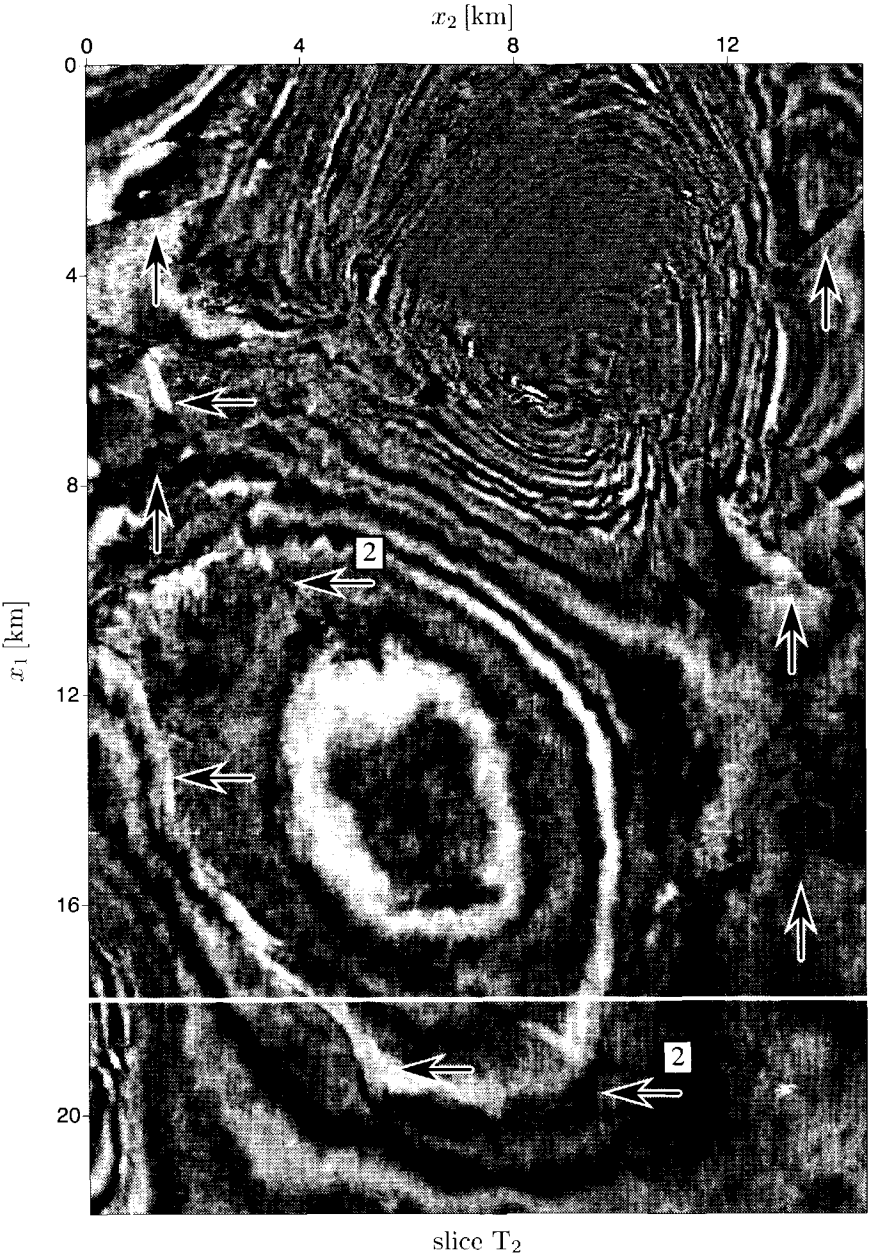


Fig. 6.9 *Time slice at $t = 1292$ ms from the unprocessed migrated volume. To be compared with the same time slice from the processed volume in Figure 6.10.*

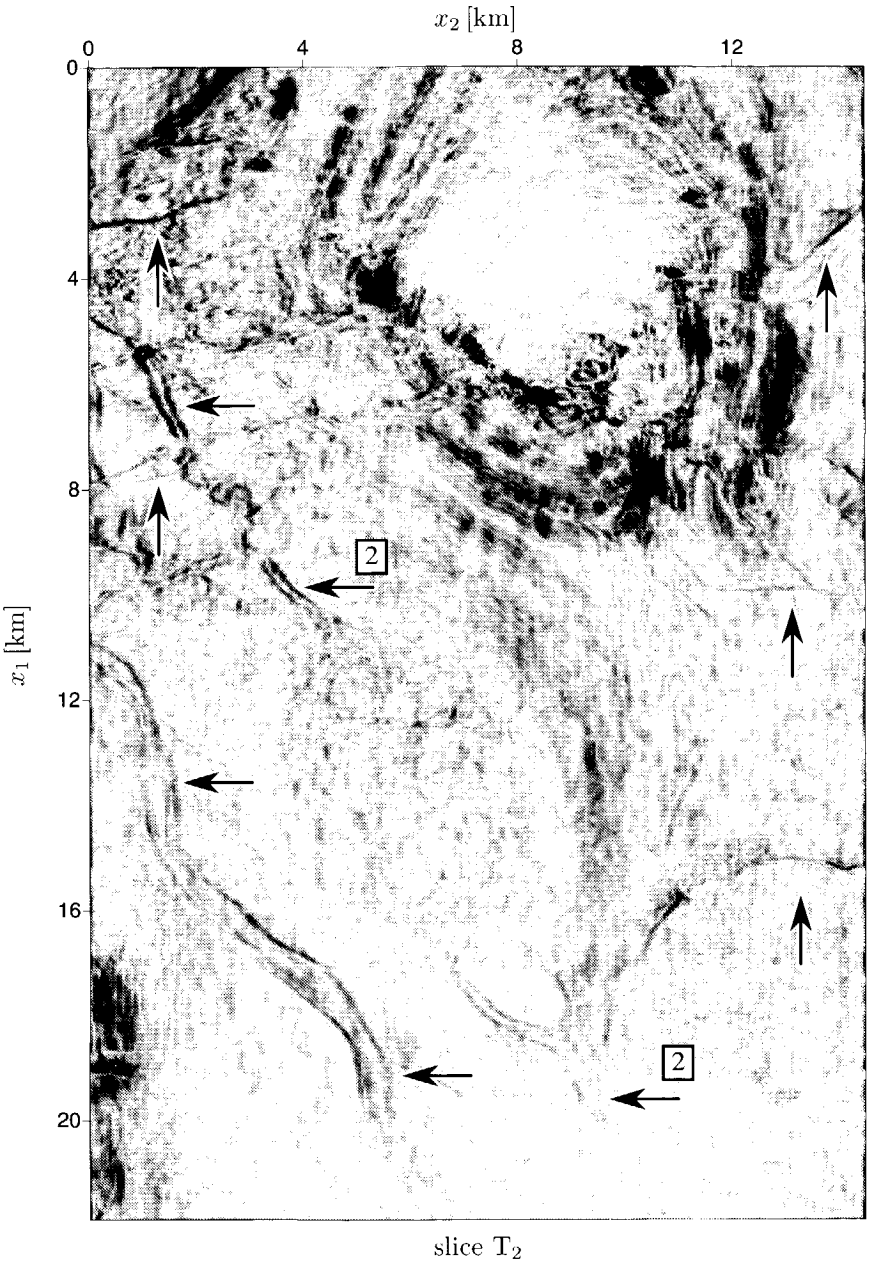


Fig. 6.10 *Time slice at $t = 1292\,ms$ from the processed migrated volume. Note that the channels and faults are at the same positions with respect to their positions in time slice T_1, which is $8\,ms$ above the present time slice. The two channels with ② are part of one big channel crossing the volume. This can be concluded by looking at slices T_3 and T_4 from the processed volume in consecutive figures.*

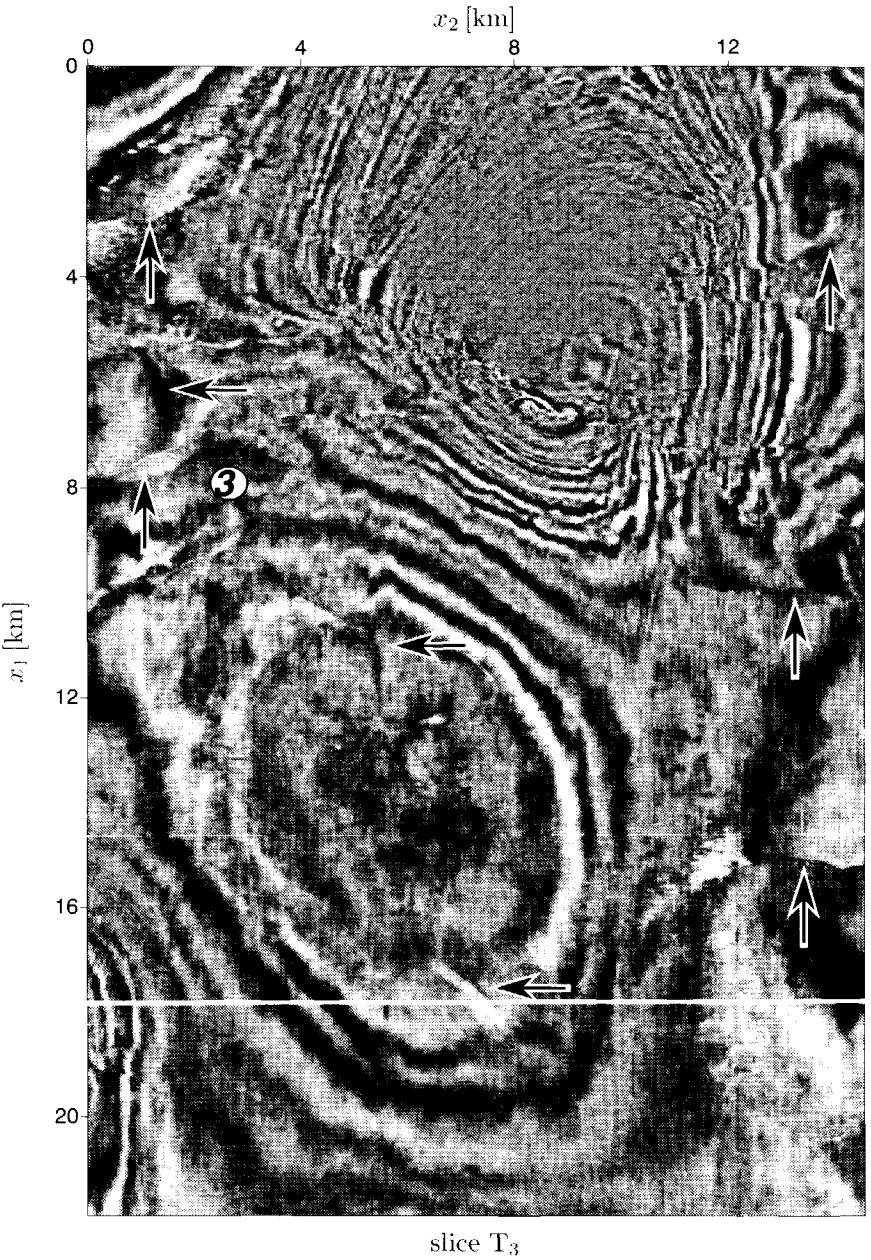


Fig. 6.11 *Time slice at $t = 1316\,ms$ from the unprocessed migrated volume. To be compared with the same time slice from the processed volume in Figure 6.12.*

Fig. 6.12 *Time slice at $t = 1316\,ms$ from the processed migrated volume. Besides the faults, the large channel going from bottom middle to the top left is clearly delineated with respect to its presence in the time slice from the unprocessed volume. At the point denoted by ③ the channel is "cut" away by two faults.*

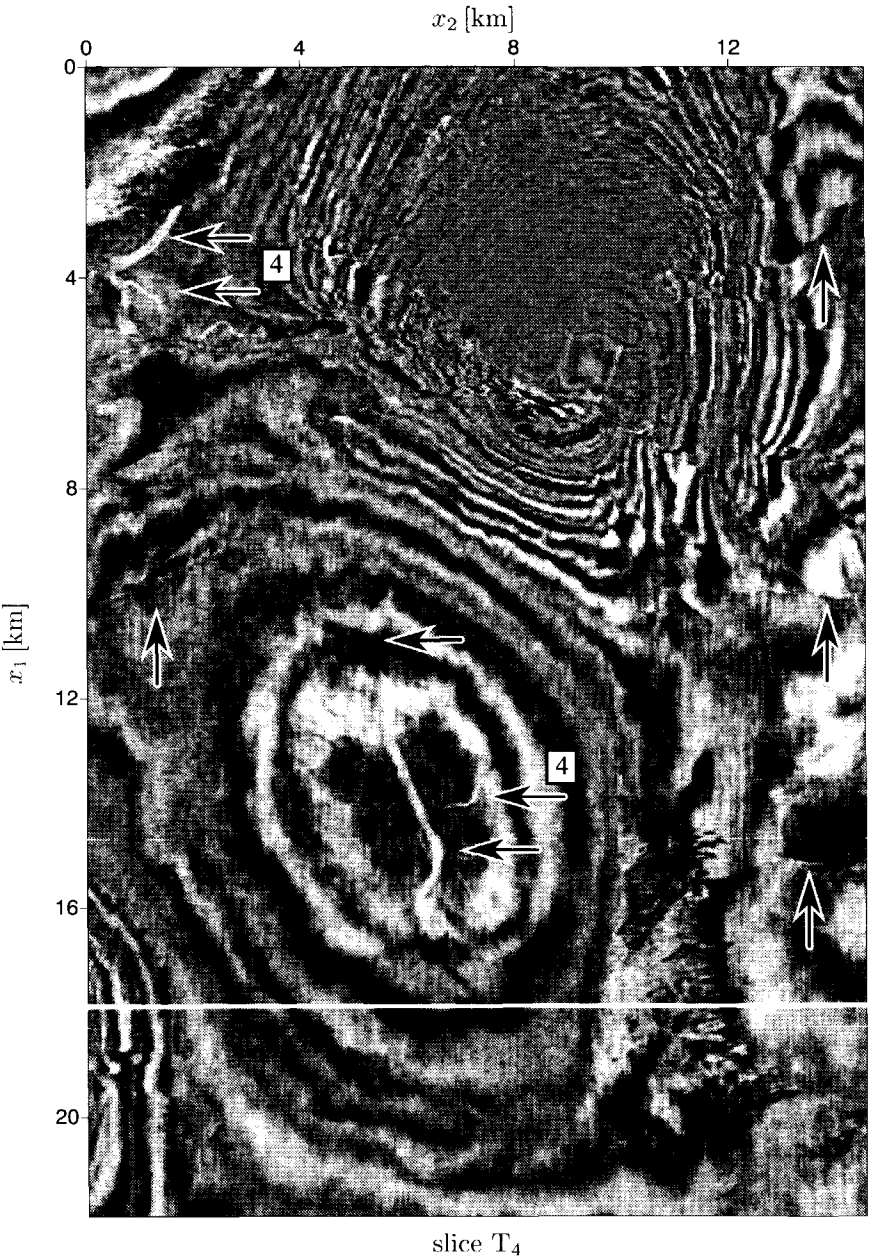


Fig. 6.13 *Time slice at $t = 1348\,ms$ from the unprocessed migrated volume. To be compared with the same time slice from the processed volume in Figure 6.14.*

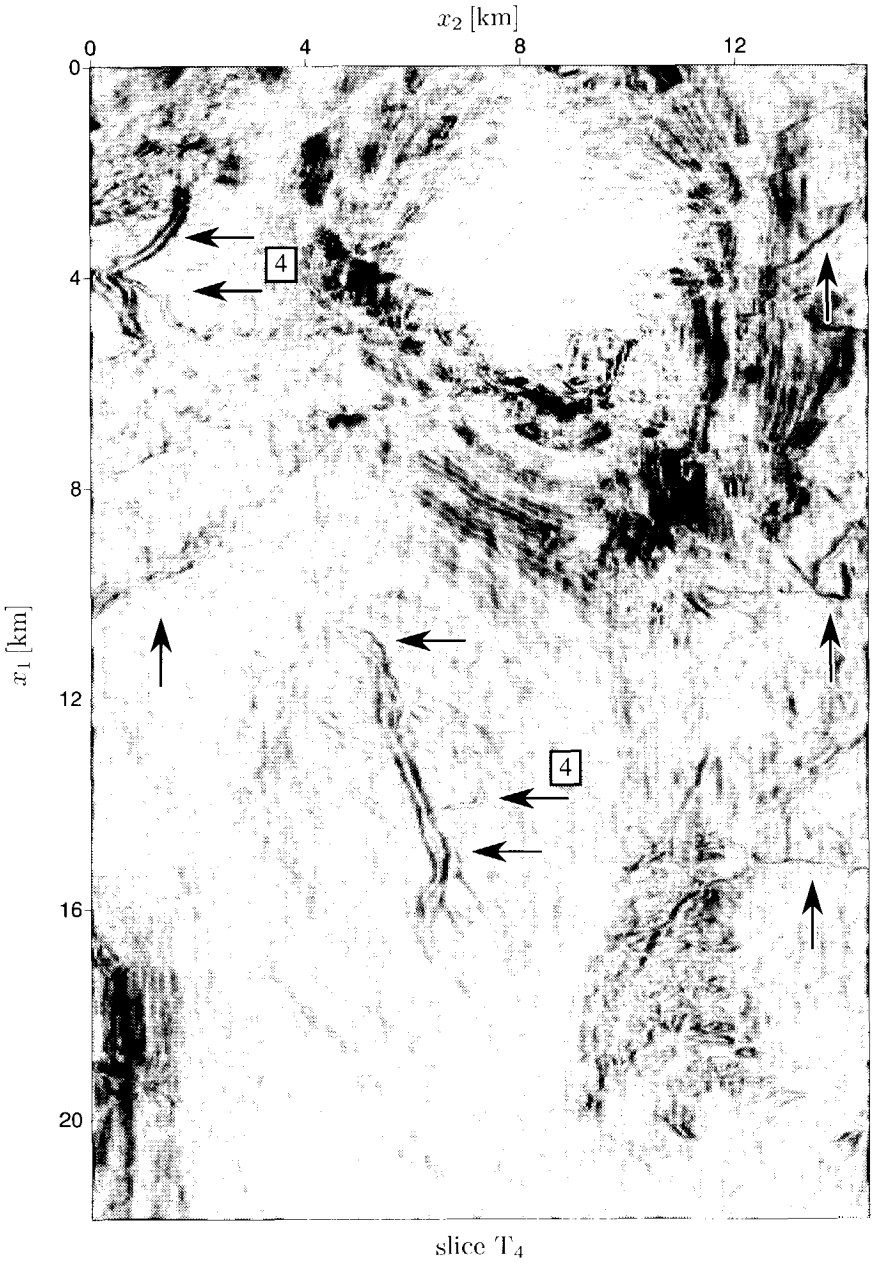


Fig. 6.14 *Time slice at $t = 1348\,ms$ from processed migrated volume. Especially interesting here are the channels. The channel in the middle completes the channel observed in the previous time slices. Note also that even the tiny channels, denoted by ④, are extracted.*

6.4 Discussion

Comparison —How does the wavelet transform approach to 3-D image analysis compare to other methods? Due to the fact that the computational details of competing methods are not known to the author, it is in general difficult to carry out a comparison on a quantitative level. A qualitative comparison however is feasible, because several publications deal with the same data set (Bahorich and Farmer, 1995; Gersztenkorn and Marfurt, 1996; Steeghs, 1997). The method presented in the first reference will be referred to as the three-trace-coherency-algorithm. The method presented by Gersztenkorn and Marfurt (1996) will be referred to as the eigenstructure-algorithm. The last method of Steeghs (1997) will be referred to as the local-Radon-Wigner-algorithm. The algorithm presented in this chapter is referred to as the wavelet-transform-algorithm[7].

Bahorich and Farmer (1995) show the first results of the coherency cube analysis. The algorithm measures the coherency of three traces in an appropriately chosen time window. The left part of figure 6 of Bahorich and Farmer (1995) is roughly the same slice as dealt with in Figure 6.7 and 6.8. The three-trace-coherency-algorithm shows good results for faults, however channels are relatively poorly extracted. The wavelet-transform-algorithm does a better job with respect to the channels. Another difference between the two algorithms is the presence of the salt dome. In the salt dome the coherency is clearly low and consequently it will be clearly visible in the coherency measure. The wavelet-transform-algorithm is more sensitive to amplitudes. Within the salt the amplitudes are low.

Gersztenkorn and Marfurt (1996) compute the eigenstructure of the covariant matrix of a number of subsequent traces. Their eigenstructure-algorithm is very well able to delineate both stratigraphic features and faults on a very detailed level. I consider their results slightly better than the results presented in this chapter, especially with respect to the fault delineation. The wavelet-transform algorithm is at the moment not insensitive enough to "background dip". Furthermore, it is interesting to observe that the sensitivity of the eigenstructure-algorithm is the same in the salt-dome region as the sensitivity of the wavelet-transform-algorithm.

Steeghs (1997) extracts the stratigraphic features and faults by determining the dip via a local 3-D Radon transform. His local-Radon-Wigner-algorithm yields equally good results with respect to delineating faults, however his results are slightly less with respect to rendering channels. The high quality of the wavelet-transform-algorithm with respect to extracting channels is especially due to its sensitivity to amplitude changes. In the Mississippi delta, channels cut through clay layers having clearly different impedance. This is the reason why an amplitude-sensitive algorithm

[7]Remember that a wavelet-transform-algorithm is not necessarily referring to a single unique method. The features that will be extracted depend on the type of differential invariant computed. Here, I used the amplitude of the gradient.

yields good results for such features.

Computational aspects — Let me finalize with mentioning a number of important computational advantages of the presented wavelet-transform-algorithm. First, it is very efficient. It has been pointed out in section 6.2.1 that it can be carried out as an $O(N)$-algorithm in the case that a discrete spatial filter is used as an approximation for a Gaussian function and its derivative. The efficiency allows for an almost real time implementation. Furthermore, the algorithm is conceptually simple. It extracts changes. The changes are combined to compute a differential invariant, such as the amplitude. If a more detailed extraction of faults is required, angle information in terms of the quantities $\Phi\{I\}$ and $\Theta\{I\}$ can be incorporated without making the algorithm more complicated. On the other hand, one can also compute other differential invariants and associated attributes, which might be sensitive to other stratigraphic features or structures. For example, Figure 6.6 shows some slices in a 3-D view for two different differential operators.

Future research —The delineation of stratigraphic features and faults in an otherwise strong-reflector-dominated migrated volume is the first step in facilitating the work of a seismic interpreter. Especially, if consecutive time slices from the migrated volume are inspected, those features become even more manifest, as can be concluded from the view provided by Figure 6.6. Despite the fact that the interpretation simplifies with the help of the time slices from the processed migrated volume, the proposed procedure is just the first step in facilitating the work of an interpreter. The second step consists of the automatic extraction of the stratigraphic features and faults. The result of the presented method might be used as input.

6.5 Summary

In this chapter I have presented a wavelet transform approach to 3-D image analysis of migrated data. The goal is to make stratigraphic features and faults manifest. Since these features are not always present as strong reflection events, identifying them is often a difficult and time consuming task. Image analysis techniques facilitate this task.

The presented method is based upon the realization that useful information does not necessarily come to us at one specific scale. In the field of image analysis this is a well-known fact. A wavelet transform carries out a decomposition in multiple scales, and thus incorporates the multiscale nature of the earth's subsurface in a natural way. More specifically, the wavelet transform has been utilized to extract a particular differential invariant, namely the amplitude of the 3-D gradient. The choice of amplitude of the gradient is suggested by the fact that local incoherences are characteristic of the features that are looked for. The first order spatial derivatives

of the Gaussian function have been used as wavelets.

The method has been applied to a migrated volume of the Gulf of Mexico. Faults and channels are very well delineated. Especially the channels are equally well or even better extracted with the wavelet transform approach than with competing methods. The efficiency of the presented algorithm is an additional advantage with respect to competing methods.

Appendix A

Matrix notation for transformations and representations

The present appendix links up to chapter 2 and serves as a practical guide to chapter 4. In the former, representations and transformations have been introduced, in the latter the transformations have been applied to the imaging problem. Just as I have done in chapter 2, the terminology is engraved onto transformations from the time domain to another domain, and vice versa. Note, however, that the equations are certainly valid for other configurations as well. In chapter 4 the transformations are applied, for example, to the lateral spatial coordinates of the monochromatic representation of primary seismic reflection data. Of course, in such a situation the interpretation has to be adjusted.

A.1 Transformation of vectors

In the realm of multiresolution approximations (see section 2.5.4), regularly sampled discrete data sets have gained a natural position. Discrete data sets can be seen as the coefficients of a projection of a function $f(t)$ in an approximation space V_0. The coefficients $\mathcal{A}_0^d f$ of equation (2.111) will be put in a vector and will be notified here by $\mathbf{f}$. Note again that this discretization is different from the discretization in the coherent state parameters, yielding frames. The time, frequency, Gabor and wavelet transformations discussed in chapter 2 now take the following form.

1. *From time-to-time domain*:

$$\mathbf{f} = \mathbf{U}_\mathcal{T}\mathbf{f}, \tag{A.1}$$

which can be seen as the discrete counterpart of the first of equations (2.37), which I repeat here for convenience

$$f(t') = (\mathrm{U}_{\mathcal{T}} f)(t').$$

Of course, the unitary transformation matrix is equal to the identity matrix $\mathbf{I}$, because $\mathbf{f}$ is already in the time domain.

2. *From time-to-frequency domain*:

$$\tilde{\mathbf{f}} = \mathbf{U}_{\mathcal{F}} \mathbf{f}, \tag{A.2}$$

which is the discrete counterpart of the second of equations (2.37), i.e.

$$\tilde{f}(\omega) = (\mathrm{U}_{\mathcal{F}} f)(\omega).$$

The unitary matrix $\mathbf{U}_{\mathcal{F}}$ contains the discretized Fourier components from $-\omega_{\mathrm{Nyquist}}$ to $+\omega_{\mathrm{Nyquist}}$. It is a unitary matrix which can be efficiently implemented via the Fast Fourier Transform (Bracewell, 1986). Its inverse is just the Hermitian, i.e.

$$\mathbf{f} = (\mathbf{U}_{\mathcal{F}})^H \tilde{\mathbf{f}}. \tag{A.3}$$

Note that the fast Fourier transform assumes a definition of a function on the torus.

3. *From time-to-Gabor domain*:
In section (2.5.2) it was concluded that a Gabor frame with a Gaussian window function will never give rise to an orthogonal decomposition. The discrete counterpart of equation (2.99), which is repeated here

$$\mathfrak{G}\{f, g\}(ma_0, nb_0) = \langle f, g_{mn} \rangle = \int f g^*_{mn}(t)\, \mathrm{d}t,$$

reads

$$\hat{\mathbf{f}} = \mathbf{V}_g \mathbf{f}. \tag{A.4}$$

where, in general, a matrix $\mathbf{V}$ denotes an invertible, but non-unitary matrix and where the subscript g refers to the Gabor transform. The vector $\hat{\mathbf{f}}$ denotes the temporal Gabor domain data. It is generally not of the same length as $\mathbf{f}$, because $\mathbf{V}_g$ is in general not a square matrix. The actual form of the Gabor transform matrix requires the specification of the temporal and frequency shift parameters a_0 and b_0 in equation (2.99) and the specification of the shape of the Gaussian function via the parameter γ in equation (2.65). The inverse transform from the temporal Gabor domain to the temporal domain reads

$$\mathbf{f} = (\mathbf{V}_g)^{-1} \hat{\mathbf{f}}. \tag{A.5}$$

4. *From time-to-wavelet domain*:
The wavelet transform in matrix terms depends on the chosen type of wavelet transform. A distinction is made between a frame decomposition and a multiresolution decomposition:

   (a) The general wavelet frame decomposition, similar to the Gabor frame decomposition, is given by (cf. equation 2.104)

$$\mathcal{W}\{f, \psi\}(\sigma_0^m, nb_0\sigma_0^m) = \langle f, \psi_{mn} \rangle = \int f(t)\psi_{mn}^*(t)\,\mathrm{d}t.$$

   Its discrete equivalent reads

$$\check{\mathbf{f}} = \mathbf{V}_w \mathbf{f}, \tag{A.6}$$

   where $\mathbf{V}_w$ denotes the discrete frame wavelet transformation matrix. Its inverse reads

$$\mathbf{f} = (\mathbf{V}_w)^{-1}\check{\mathbf{f}}. \tag{A.7}$$

   (b) In the case that the wavelets are chosen as a part of a multiresolution approximation, the matrix transformation becomes unitary, according to

$$\check{\mathbf{f}} = \mathbf{U}_w \mathbf{f}, \tag{A.8}$$

   where $\check{\mathbf{f}}$ is a vector containing the elements of $\{\mathcal{D}_1^d f, \cdots, \mathcal{D}_M^d f, \mathcal{A}_M^d f\}$, which is introduced in equation (2.132). A specific multiresolution approximation and a specific value M are required to fix the actual wavelet transform. Equation (A.8) is referred to as the discrete wavelet transform. It can be implemented efficiently[1]. The inverse discrete wavelet transform reads

$$\mathbf{f} = (\mathbf{U}_w)^H \check{\mathbf{f}}. \tag{A.9}$$

   (c) A frequently encountered variation of the discrete wavelet transforms is the biorthogonal wavelet transform. A slight increase in complexity of the forward and inverse wavelet transform, especially due to the fact that different filter pairs are used for the forward and inverse transform, is rewarded by symmetric filters, and thus symmetric scaling functions and (anti-)symmetric wavelets.

[1] The discrete wavelet transform can be carried out in roughly $(N \times L)$ operations, where N is the number of samples in a vector, and L the length of the filter $\mathbf{g}$ or $\mathbf{h}$, introduced in section 2.5.4. With respect to efficiency, but also with respect to flexibility, the recently introduced LiftPack scheme is very promising (Sweldens, 1996). It is more efficient in terms of floating point operations and memory usage.

A.2 Matrix transformations

Equation (2.23) shows how the representation of a kernel in one domain can be transformed to another domain. The discrete equivalent of this scheme can be straightforwardly obtained. The discrete representation of a kernel can be seen as the projection of the kernel in a specific multiresolution subspace, which is fixed here at V_0, i.e.

$$K_{nk} = \langle \mathcal{K}\phi_{0n}, \phi_{0k} \rangle , \tag{A.10}$$

and consequently the matrix $\mathbf{K}$ consists of the matrix elements K_{nk}. The matrix element K_{nk} expresses the response of the physical system of a source at a time[2] represented by k measured at a time represented by n. The transformation of matrices to the domains discussed in the first part of this section reads (for more comments see the previous paragraph) for a unitary transformation matrix $\mathbf{U}$

$$\breve{\mathbf{K}} = \mathbf{U}\mathbf{K}\mathbf{U}^H, \tag{A.11}$$

and its inverse reads

$$\mathbf{K} = \mathbf{U}^H\breve{\mathbf{K}}\mathbf{U}. \tag{A.12}$$

Here, the matrix $\mathbf{U}$ can be any of the previously discussed transformation matrices, and the breve $\breve{}$ can represent any of the aforementioned domains, except the Gabor domain. For the Gabor transform, which is invertible but which can never be orthonormal, a slightly different form has to be used

$$\hat{\mathbf{K}} = \mathbf{V}\mathbf{K}\mathbf{V}^{-1}, \tag{A.13}$$

and its inverse reads

$$\mathbf{K} = \mathbf{V}^{-1}\hat{\mathbf{K}}\mathbf{V}. \tag{A.14}$$

In chapter 4, the interpretation of the matrix representations in different domains is discussed in more detail. For the wavelet transform, I have introduced also a slight variation giving rise to the so-called non-standard wavelet representation (Beylkin et al., 1991).

[2] Note that the actual physical interpretation depends on the functional parameters in the matrix, but the basic structure of the matrix remains for any variable the same.

Appendix B

Homogeneous distributions

This appendix provides background material for chapter 5. It gives a number of properties of homogeneous distributions. These properties are extracted from Gel'fand and Shilow (1960) and Hörmander (1983). For proofs and derivations the reader is referred to the cited references.

The one-sided distributions $|t|_{\pm}^{\alpha}$ are given by

$$|t|_{+}^{\alpha} = \begin{cases} 0 & t \leq 0 \\ t^{\alpha} & t > 0 \end{cases} \qquad |t|_{-}^{\alpha} = \begin{cases} (-t)^{\alpha} & t < 0 \\ 0 & t \geq 0. \end{cases} \tag{B.1}$$

They are locally integrable for $\Re\alpha > -1$. For $\Re\alpha < -1$ and $\alpha \neq -1, -2, \ldots$, $|t|_{\pm}^{\alpha}$ has to be interpreted in the distributional sense.

Let me first consider $|t|_{+}^{\alpha}$ in distributional sense. For a real function $\phi \in C_0^{\infty}(\mathbb{R})$ and any integer $k > 0$, the differential operators in an integral can be transferred according to

$$\langle |t|_{+}^{\alpha}, \phi \rangle = \int_0^{\infty} t^{\alpha} \phi(t)\, \mathrm{d}t \tag{B.2}$$

$$= \frac{(-1)^k}{(\alpha+1)\ldots(\alpha+k)} \int_0^{\infty} t^{\alpha+k} \partial_t^k \phi(t)\, \mathrm{d}t. \tag{B.3}$$

The integrand at the right-hand side of equation (B.3) is analytic for $\Re\alpha > -k-1$, except for $\alpha = -1, -2, \ldots$, where it has simple poles. The residue of $|t|_{+}^{\alpha}$ at those poles is given by $\frac{(-1)^{k-1}}{(k-1)!}\delta^{(k-1)}(t)$ for $\alpha = -k$.

In a similar way one can extend left-sided homogeneous distributions $|t|_{-}^{\alpha}$ to $\alpha \in \mathbb{R}$, $\alpha \neq -1, -2, \ldots$ by inspecting

$$\langle |t|_{-}^{\alpha}, \phi \rangle = \int_{-\infty}^{0} (-t)^{\alpha} \phi(t)\, \mathrm{d}t. \tag{B.4}$$

The residue of $|t|_-^\alpha$ turns out to be given by $\frac{\delta^{(k-1)}(t)}{(k-1)!}$ for $\alpha = -k$.

From the one-sided homogeneous distributions one can construct even and odd combinations according to

$$|t|_{\text{cusp}}^\alpha = |t|_+^\alpha + |t|_-^\alpha, \tag{B.5}$$

and

$$|t|_{\text{diff}}^\alpha = |t|_+^\alpha - |t|_-^\alpha, \tag{B.6}$$

respectively. Since the poles of $|t|_+^\alpha$ and $|t|_-^\alpha$ differ with a factor $(-1)^k$, the distributions $|t|_{\text{cusp}}^\alpha$ have only poles for $\alpha = -1, -3, -5, \ldots$. The residue at the poles is given by $2\frac{\delta^{(2m)}(t)}{(2m)!}$, where α and m are related by $\alpha = -2m - 1$. For $\alpha = -2m$, the distributions $|t|_{\text{cusp}}^\alpha$ exist. Analogous one finds for the odd combination $|t|_{\text{diff}}^\alpha$ of the one-sided distributions, poles for $\alpha = -2m$ with residues $-2\frac{\delta^{(2m-1)}(t)}{(2m-1)!}$. For $\alpha = -2m - 1$, the distributions $|t|_{\text{diff}}^\alpha$ exist.

Normalization of the distributions $|t|_\pm^\alpha$, $|t|_{\text{cusp}}^\alpha$ and $|t|_{\text{diff}}^\alpha$

The poles for negative integers α, can be removed by dividing the discussed distributions by regular functions of α, which have a pole at the point where the distributions have a pole. Moreover, the pole should be of equal strength. A good candidate for such a function is the Gamma function. This can be easily understood if we choose $\phi(t) = \mathrm{e}^{-t}$ in equation (B.2), yielding

$$\langle |t|_+^\alpha, \phi \rangle = \int_0^\infty t^\alpha e^{-t}\,\mathrm{d}t = \Gamma(\alpha + 1). \tag{B.7}$$

Therefore, the appropriate normalization function for $|t|_+^\alpha$ is $\Gamma(\alpha + 1)$. The normalization for $|t|_-^\alpha$ is $\Gamma(\alpha + 1)$, as well. Since $|t|_{\text{cusp}}^\alpha$ and $|t|_{\text{diff}}^\alpha$ do not have poles for all negative integers, here slightly different normalizations have to be chosen: $\Gamma\left(\frac{\alpha+1}{2}\right)$ and $\Gamma\left(\frac{\alpha+2}{2}\right)$, respectively. The normalized distributions defined for $\alpha \in \mathbb{R}$ consequently, are

$$\chi_+^\alpha(t) = \frac{|t|_+^\alpha}{\Gamma(\alpha + 1)} \tag{B.8}$$

$$\chi_-^\alpha(t) = \frac{|t|_-^\alpha}{\Gamma(\alpha + 1)} \tag{B.9}$$

$$\chi_{\text{cusp}}^\alpha(t) = \frac{|t|_{\text{cusp}}^\alpha}{\Gamma\left(\frac{\alpha+1}{2}\right)} \tag{B.10}$$

$$\chi_{\text{diff}}^\alpha(t) = \frac{|t|_{\text{diff}}^\alpha}{\Gamma\left(\frac{\alpha+2}{2}\right)}. \tag{B.11}$$

For the values of α where the numerator has a pole, it is found for an integer $k > 0$

$$\chi_+^\alpha(t)\big|_{\alpha=-k} = \delta^{(k-1)}(t) \tag{B.12}$$

$$\chi_-^\alpha(t)\big|_{\alpha=-k} = (-1)^{(k-1)}\delta^{(k-1)}(t) \tag{B.13}$$

$$\chi_{\text{cusp}}^\alpha(t)\big|_{\alpha=-2k-1} = \frac{(-1)^k\delta^{(2k)}(t)k!}{(2k)!} \tag{B.14}$$

$$\chi_{\text{diff}}^\alpha(t)\big|_{\alpha=-2k} = \frac{(-1)^k\delta^{(2k-1)}(t)(k-1)!}{(2k-1)!}. \tag{B.15}$$

Differentiation of the distributions $\chi_\pm^\alpha$, $|t|_{\text{cusp}}^\alpha$ and $|t|_{\text{diff}}^\alpha$

The derivatives of the normalized distributions are given by

$$\partial_t \chi_+^\alpha(t) = \chi_+^{\alpha-1}(t) \tag{B.16}$$

$$\partial_t \chi_-^\alpha(t) = -\chi_-^{\alpha-1}(t) \tag{B.17}$$

$$\partial_t \chi_{\text{cusp}}^\alpha(t) = \alpha\chi_{\text{diff}}^{\alpha-1}(t) \tag{B.18}$$

$$\partial_t \chi_{\text{diff}}^\alpha(t) = 2\chi_{\text{cusp}}^{\alpha-1}(t), \tag{B.19}$$

which have been used in chapter 5.

Bibliography

Abramowitz, M. and Stegun, I. A. (1970). *Handbook of mathematical functions.* Dover Publications, Inc., New York.

Aki, K. and Richards, P. G. (1980). *Quantitative seismology, Vol. I.* W.H. Freeman and Company, San Fransisco.

Alpert, B., Beylkin, G., Coifman, R., and Rokhlin, V. (1993). Wavelet-like bases for the fast solution of second-kind integral equations. *SIAM J. Sci. Comp.*, **14**(1), 159–184.

Altes, R. A. (1978). The Fourier-Mellin transform and mammalian hearing. *J. Acoust. Soc. Am.*, **63**(1), 174–183.

Arneodo, A., Bacry, E., Jaffard, S., and Muzy, J. F. (1997). Oscillating singularities on Cantor sets. a grand canonical multifractal formalism. *J. Stat. Phys.*, **87**(1/2), 179–209.

Aslaksen, E. W. and Klauder, J. R. (1968). Unitary representation of the affine group. *J. Math. Phys.*, **9**, 206–211.

Aslaksen, E. W. and Klauder, J. R. (1969). Continuous representation theory using the affine group. *J. Math. Phys.*, **10**, 2267–2275.

Bacry, E., Muzy, J. F., and Arneodo, A. (1993). Singularity spectrum of fractal signals from wavelet analysis: Exact results. *J. Stat. Phys.*, **70**(3/4), 635–674.

Bahorich, M. and Farmer, S. (1995). 3-D seismic discontinuity for faults and stratigraphic features: The coherence cube. *The Leading Edge*, **14**(10), 1053–1058.

Baraniuk, R. G. and Jones, D. L. (1995). Unitary equivalence: a new twist on signal processing. *IEEE Trans. Signal Process.*, **43**(11), 2269–2282.

Bastiaans, M. J. (1980). The expansion of an optical signal into a discrete set of Gaussian beams. *Optik*, **57**, 95–102.

Berezin, F. A. and Shubin, M. A. (1991). *The Schrödinger equation.* Kluwer Academic Publishers.

Berkhout, A. J. (1982). *Seismic migration. Imaging of acoustic energy by wave field extrapolation. A. Theoretical aspects*, volume 14A of *Developments in solid earth geophysics.* Elsevier.

Berkhout, A. J. (1984). *Seismic resolution. A quantitave analysis of resolving power of acoustical echo techniques.* Geophysical Press, London.

Berkhout, A. J. (1997a). Pushing the limits of seismic imaging, Part I: Prestack migration in terms of double dynamic focusing. *Geophysics*, **62**(3), 937–953.

Berkhout, A. J. (1997b). Pushing the limits of seismic imaging, Part II: Integration of prestack migration, velocity estimation, and AVO analysis. *Geophysics*, **62**(3), 954–969.

Berkhout, A. J. and Wapenaar, C. P. A. (1990). Delphi: Delft philosophy on acoustic and elastic inversion. *The Leading Edge*, **9**(2), 30–33.

Berkhout, A. J. and Wapenaar, C. P. A. (1993). A unified approach to acoustical reflection imaging. Part II: The inverse problem. *J. Acoust. Soc. Am.*, **93**(4), 2017–2023.

Beylkin, G. (1992). On the representation of operators in bases of compactly supported wavelets. *SIAM J. Numer. Anal.*, **29**, 1716–1740.

Beylkin, G. (1996). Wavelets, multiresolution analysis and fast numerical algorithms. In G. Erlebach, M. Y. Hussaini, and L. M. Jameson, editors, *Wavelets: theory and applications*, ICASE/LaRC Series in Computational Science and Engineering, chapter 4, pages 182–262. Oxford University Press.

Beylkin, G., Coifman, R., and Rokhlin, V. (1991). Fast wavelet transforms and numerical algorithms I. *Comm. Pure Appl. Math.*, **44**, 141–183.

Blacquière, G., Debeye, H. W. J., Wapenaar, C. P. A., and Berkhout, A. J. (1989). 3D table-driven migration. *Geoph. Prosp.*, **37**(8), 925–958.

Bleistein, N. (1987). On the imaging of reflectors in the earth. *Geophysics*, **52**(7), 931–942.

Bleistein, N. and Cohen, J. K. (1979). Direct inversion procedure for Clearbout's equations. *Geophysics*, **44**(6), 1034–1040.

Blok, H. (1995). Theory of electromagnetic waveguides, Volume I. Lecture notes, Delft University of Technology.

Bony, J. (1983). Propagation et interaction des singularit'es pour les solutions des equations au derivées partielles non-lineaires. In *Proc of the Internations Congress of Mathematicians*, pages 1133–1147, Warszawa, Poland.

Bortfeld, R. (1961). Approximation to the reflection and transmission coefficients of plane longitudinal and transverse waves. *Geoph. Prosp.*, **9**, 485–502.

Bosman, C. and Reiter, E. (1993). Seismic data compression using wavelet transforms. In *63th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 1261–1264.

Bracewell, R. N. (1986). *The Fourier transform and its applications*. McGraw-Hill series in electrical engineering. McGraw-Hill, 2nd edition.

Brekhovskikh, L. M. (1960). *Waves in layered media*. Academic Press, New York.

Brekhovskikh, L. M. and Godin, O. A. (1990). *Acoustics of layered media I: Plane and quasi-plane waves*. Springer series on wave phenomena. Springer-Verlag.

Bremmer, H. (1951). The W.K.B. approximation as the first term of a geometric-optical series. *Comm. Pure Appl. Math.*, **4**, 105–115.

Brewster, M. E. and Beylkin, G. (1995). A multiresolution strategy for numerical homogenization. *Appl. Comput. Harmon. Anal.*, **2**, 327–349.

Burt, P. and Adelson, E. (1983). The Laplacian pyramid as a compact image code. *IEEE Trans. Comm.*, **31**, 532–549.

Canny, J. (1986). A computational approach to edge detection. *IEEE Trans. Patt. Anal.*

Mach. Intell., **8**(6), 679–698.

Castagna, J. P. (1993). AVO analysis - tutorial and review. In M. M. Castagna, John P.; Backus, editor, *Offset-dependent reflectivity - Theory and practice of AVO analysis*, pages 3–36. Soc. Expl. Geophys., Tulsa.

Castagna, J. P., Han, D.-H., and Batzle, M. L. (1995). Issues in rock physics and implications for DHI interpretation. *The Leading Edge*, **14**(8), 883–885.

Chen, T. (1995). *Seismic data compression*. Master's thesis, Colorado School of Mines.

Churchill, R. V. and Brown, J. W. (1978). *Fourier series and boundary value problems*. McGraw-Hill, 3rd edition.

Claerbout, J. F. (1971). Towards a unified theory of reflector mapping. *Geophysics*, **36**(3), 467–481.

Claerbout, J. F. (1985). *Imaging the earth's interior*. Blackwell Science Publ., Oxford.

Cohen, A., Daubechies, I., and Vial, P. (1993). Wavelets on the interval and fast wavelet transforms. *Appl. Comput. Harmon. Anal.*, **1**, 54–81.

Cohen, L. (1993). The scale representation. *IEEE Trans. Signal Process.*, **41**(12), 3275–3292.

Cohen, L. (1995). *Time-frequency analysis*. Prentice-Hall.

Coifman, R. R. and Wickerhauser, M. V. (1992). Entropy-based algorithms for best basis selection. *IEEE Trans. Inform. Theory*, **38**(2), 713–718.

Corones, J. (1975). Bremmer series that correct for parabolic approximations. *J. Math. Anal. Appl.*, **50**, 361–372.

Daubechies, I. (1988). Orthonormal bases of compactly supported wavelets. *Comm. Pure Appl. Math.*, **41**, 909–996.

Daubechies, I. (1990). The wavelet transform, time-frequency localization and signal analysis. *IEEE Trans. Inform. Theory*, **36**, 961–1005.

Daubechies, I. (1992). *Ten lectures on wavelets*. SIAM Philadelphia.

Daubechies, I. (1996). Where do wavelets come from? – A personal point of view. *Proc. of the IEEE*, **84**(4), 510–513. In: Special Issue on Wavelets.

Daubechies, I., Grossmann, A., and Meyer, Y. (1986). Painless nonorthogonal expansions. *J. Math. Phys.*, **27**(5), 1271–1283.

Dautray, R. and Lions, J. (1988). *Functional and variational methods*, volume 2 of *Mathematical analysis and numerical methods for science and technology*. Springer-Verlag.

Dautray, R. and Lions, J. (1990). *Spectral theory and applications*, volume 3 of *Mathematical analysis and numerical methods for science and technology*. Springer-Verlag.

Dautray, R. and Lions, J. (1992). *Evolution problems I*, volume 5 of *Mathematical analysis and numerical methods for science and technology*. Springer-Verlag.

de Hoop, M. V. (1992). *Directional decomposition of transient acoustic wave fields*. Ph.D. thesis, Delft University of Technology.

Delprat, N., Escudié, B., Guillemain, P., Kronland-Martinet, R., Tchamitchian, P., and

Torrésani, B. (1992). Asymptotic wavelet and Gabor analysis: Extraction of instantaneous frequencies. *IEEE Trans. Inform. Theory*, **38**(2), 644–664.

Deschamps, G. A. (1971). Gaussian beam as a bundle of complex rays. *Electronic Letters*, **7**, 684–685.

Dessing, F. J. (1995). Seismic migration with one-way operators in the wavelet domain. Technical report, Delft University of Technology. DELPHI, Volume VI, Chapter 22.

Dessing, F. J. and Grimbergen, J. L. T. (1996). Imaging in strongly laterally-varying media by modal decomposition. Technical report, Delft University of Technology. DELPHI, Volume VII, Appendix A.

Dessing, F. J. and Hoekstra, E. V. (1997). Multiscale tools for seismic data compression. Technical report, Delft University of Technology. DELPHI, Volume VIII, Chapter 23.

Dessing, F. J. and Hoekstra, E. V. (1998). Characterization of reflection at singularities. *Geophysics*. Submitted for publication.

Dessing, F. J. and Wapenaar, C. P. A. (1994). Wavefield extrapolation using the wavelet transform. In *64th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 1355–1358.

Dessing, F. J. and Wapenaar, C. P. A. (1995). Efficient migration with one-way operators in the wavelet transform domain. In *65th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 1240–1243.

Dessing, F. J., Hoekstra, E. V., Herrmann, F. J., and Wapenaar, C. P. A. (1996). Multiscale edge detection by means of multiscale migration. In *66th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*.

DeWitt-Morette, C., Maheshwari, A., and Nelson, B. (1979). Path integration in non-relativistic quantum mechanics. *Physics Reports*, **50**(5), 255–372.

Duffey, G. H. (1992). *Applied group theory for physicists and chemists*. Prentice-Hall.

Dunford, N. and Schwartz, J. T. (1963). *Linear operators. Part II: Spectral theory*, volume 7 of *Pure and applied mathematics*. Interscience Publishers.

EC-report (1996). Strategies in oil and gas technology. Study carried out on behalf of the Directorate-General XVII of the European Commission.

Einziger, P. D., Raz, S., and Shapira, M. (1986). Gabor representation and aperture theory. *J. Opt. Soc. Am. A*, **3**(4), 508–522.

Ekern, O. F. (1987). Midgard. In A. M. Spencer, editor, *Geology of the Norwegian oil and gas fields: Norwegian Petroleum Society*, pages 403–410. Graham and Trotman, Londen.

Engelhard, L. (1996). Determination of seismic-wave attenuation by complex trace analysis. *Geophys. J. Int.*, **125**, 608–622.

Ergas, R. A., Polzer, R. S., Donoho, P. L., and Galibert, P. Y. (1996). Pitfalls in compressing land seismic trace data. In *58th EAGE Conference and Technical Exhibition, Extended Abstracts*, page P156.

Ernst, F. E. and Herman, G. C. (1995). Computation of Green's function of laterally varying media by means of a modal expansion. In *65th Ann. Intern. Mtg., Soc. Expl.*

Geophys., Expanded Abstracts, pages 623–626.

Etgen, J. (1994). Stability of explicit depth extrapolation through laterally varying media. In *64th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 1266–1269.

Faris, W. G. (1995). Random waves and localization. *Notices of the AMS*, **42**(8), 848–853.

Feynman, R. P. and Hibbs, A. R. (1965). *Quantum mechanics and path integrals*. McGraw-Hill. Newyork.

Fishman, L. (1992). Exact and operator rational approximate solutions of the Helmholtz, Weyl composition equation in underwater acoustics The quadratic profile. *J. Math. Phys.*, **33**(5), 1887–1914.

Fishman, L. and McCoy, J. J. (1984b). Derivation and application of extended parabolic wave theories. I The factorized Helmholtz equation. *J. Math. Phys.*, **25**(2), 285–296.

Fishman, L. and McCoy, J. J. (1984a). Derivation and application of extended parabolic wave theories. II Path integral representations. *J. Math. Phys.*, **25**(2), 297–308.

Fishman, L., Gautesen, A. K., and Sun, Z. (1997). Uniform high-frequency approximations of the square root Helmholtz operator symbol. *Preprint*.

Florack, L., Saldens, A., ter Haar Romeny, B. M., Koenderink, J., and Viergever, M. (1994). Nonlinear scale-space. In B. M. ter Haar Romeny, editor, *Geometry-driven diffusion*. Kluwer Academic Publishers.

Fokkema, J. T. and van den Berg, P. M. (1993). *Seismic applications of acoustic reciprocity*. Elsevier. Amsterdam.

Foster, D. J., Mosher, C. C., and Hassanzadeh, S. (1994). Wavelet transform methods for geophysical applications. In *64th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 1465–1468.

Gabor, D. (1946). Theory of communication. *J. Inst. Elect. Eng. (London)*, **93**, 429–457.

Gazdag, J. (1978). Wave equation migration with the phase-shift method. *Geophysics*, **43**, 1342–1351.

Gazdag, J. and Sguazzero, P. (1984). Migration of seismic data by phase-shift plus interpolation. *Geophysics*, **49**(2), 124–131.

Gel'fand, I. M. and Shilow, G. (1960). *Verallgemeinerte Funtionen, Vol I*. VEB Deutscher Verlag der Wissenschaften. Also published in English as *Generalized functions* with Academic Press.

Gersho, A. and Gray, R. M. (1992). *Vector quantization and signal compression*. Kluwer Academic Publishers.

Gersztenkorn, A. and Marfurt, K. J. (1996). Eigenstructure based coherence computations. In *66th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 328–331.

Gersztenkorn, A., Sharp, J. A., Scheet, R. M., and Marfurt, K. J. (1996). Delineation of wrench tectonic features using 3-D coherency — offshore Trinidad. In *66th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 344–347.

Gilbert, F. and Backus, G. E. (1966). Propagator matrices in elastic wave and vibration problems. *Geophysics*, **31**(2), 326–332.

Glauber, R. J. (1963). Photon correlations. *Phys. Rev. Lett.*, **10**(3), 84–86.

Gohberg, I. and Goldberg, S. (1981). *Basic operator theory.* Birkhäuser, Boston.

Golub, G. H. and van Loan, C. F. (1984). *Matrix computations.* The John Hopkins University press.

Goswami, J. C., Chan, A. K., and Chui, C. K. (1995). On solving first-kind integral equations using wavelet on a bounded interval. *IEEE Trans. Ant. Prop.*, **43**(6), 614–622.

Grimbergen, J. L. T. (1995). *One-way wave field operators in laterally varying media.* Master's thesis, Delft University of Technology.

Grimbergen, J. L. T., Dessing, F. J., and Wapenaar, C. P. A. (1998). Modal expansion of one-way operators in laterally varying media. *Accepted for publication in Geophysics.*

Grossmann, A., Morlet, J., and Paul, T. (1986). Transforms associated to square integrable group representations II: Examples. *Ann. Inst. Henri Poincaré*, **45**(3), 293–309.

Grossmann, A., Holschneider, M., Kronland-Martinet, R., and Morlet, J. (1987). Detection of abrupt changes in sound signals with the help of wavelet transforms. *Advances in electronic and electron physics, Suppl. 19, Inverse Problems.*

Grubb, H. and Walden, A. (1994). Characterising seismic time series using the discrete wavelet transform. In *56th Mtg. Eur. Assoc. Expl. Geophys., Extended Abstracts*, page H013.

Haar, A. (1910). Zur Theorie der orthogonalen Functionensysteme. *Math. Ann.*, **69**, 331–371.

Hale, D. (1991). Stable explicit depth extrapolation of seismic wavefields. *Geophysics*, **56**(11), 1770–1777.

Hale, D., Hill, N. R., and Stefani, J. (1992). Imaging salt with turning waves. *Geophysics*, **57**, 1453–1462.

Haskell, N. L., Nissen, S. E., Lopez, J. A., and Bahorich, M. S. (1995). 3-D seismic coherency and the imaging of sedimentological features. In *65th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 1532–1534.

Heil, C. E. and Walnut, D. F. (1989). Continuous and discrete wavelet transforms. *SIAM Review*, **31**(1), 628–626.

Herrmann, F. J. (1995). Integration and characterization by scale analysis. Technical report, Delft University of Technology. DELPHI, Volume VI, Chapter 24.

Herrmann, F. J. (1997). *A scaling medium representation: a discussion on well-logs, fractals and waves.* Ph.D. thesis, Delft University of Technology.

Higgins, J. R. (1977). *Completeness and basis properties of sets of special functions.* Cambridge tracts in mathematics 72. Cambridge University Press, Cambridge, UK.

Hill, N. R. (1990). Gaussian beam migration. *Geophysics*, **55**(11), 1416–1428.

Hoekstra, E. V. (1996). *Multiscale analysis of seismic data by the wavelet transform.* Master's thesis, Delft University of Technology.

Hoekstra, E. V. and Dessing, F. J. (1998). Characterization of singularities. *Geophysics.*

Submitted for publication.

Holberg, O. (1988). Towards optimum one-way wave propagation. *Geophysical Prospecting*, **36**, 99–114.

Holschneider, M. (1995). *Wavelets: an analysis tool*. Oxford University Press.

Holschneider, M. and Tchamitchian, P. (1990). Régularité locale de la function 'non-différentiable' de Riemann. In P. G. Lemarié, editor, *Les ondelettes en 1989*, pages 102–124. Springer Verlag.

Hörmander, L. (1983). *The analysis of linear partial differential operators I: Distribution theory and Fourier analysis*. Springer-Verlag.

Jaffard, S. (1989). Exposants de Hölder en points donnés et coefficients d'ondelettes. *C. R. Acad. Sci. Paris Sér. I Math.*, **308**, 79–81.

Jaffard, S. (1991). Pointwise smoothness, two microlocalisation and wavelet coefficients. *Publicacions Mathematiques*, **35**.

Kabir, M. M. N. (1997). *Velocity estimation of the complex subsurface using the common focus point technology*. Ph.D. thesis, Delft University of Technology.

Kaiser, G. (1994). *A friendly guide to wavelets*. Birkhauser.

Kim, H., Ling, H., and Lee, C. (1996). A fast moment method algorithm using spectral domain wavelet concepts. *Radio Science*, **31**(5), 1253–1261.

Kjartansson, E. (1979). Constant Q-wave propagation and attenuation. *Journal of Geophysical Research*, **84**, 4737–4748.

Klauder, J. R. and Skagerstam, B.-S. (1985). *Coherent states*. World scientific.

Koenderink, J. J. (1984). The structure of images. *Biol. Cybern.*, **50**, 363–370.

Kosloff, D. and Kessler, D. (1987). Accurate depth migration by a generalized phase-shift method. *Geophysics*, **52**, 1074–1084.

Lee, D., Mason, I. M., and Jackson, G. M. (1991). Split-step Fourier shot-record migration with deconvolution imaging. *Geophysics*, **56**(11), 1786–1793.

Love, A. E. H. (1927). *A treatise on the mathematical theory of elasticity*. Cambridge University Press.

Luo, Y., Higgs, W. G., and Kowalik, W. S. (1996). Edge detection and stratigraphic analysis using 3D seismic data. In *66th Ann. Inter. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 324–327.

Mallat, S. G. (1989a). Multiresolution approximations and wavelet orthonormal bases of $L^2(\mathbb{R})$. *Trans. Amer. Math. Soc.*, **315**, 69–87.

Mallat, S. G. (1989b). A theory for multiresolution signal decomposition, the wavelet representation. *IEEE Trans. Patt. Anal. Mach. Intell.*, **11**, 674–693.

Mallat, S. G. and Hwang, W. L. (1992). Singularity detection and processing with wavelets. *IEEE Trans. Inform. Theory*, **38**(2), 617–643.

Mallat, S. G. and Zhong, S. (1992). Characterization of signals from multiscale edges. *IEEE Trans. Patt. Anal. Mach. Intell.*, **14**(7), 710–732.

Marr, D. (1982). *Vision.* W.H. Freeman and Company.

Maxwell, J. C. (1882). Science and free will. In L. Campbell and W. Garnett, editors, *The life of James Clerk Maxwell.* Macmillan, London.

Messiah, A. (1958). *Quantum mechanics*, volume I. North-Holland and John Wiley. Translated from the French by G.M. Temmer.

Messiah, A. (1961). *Quantum mechanics*, volume II. North-Holland and John Wiley. Translated from the French by J. Potter.

Meyer, Y. (1985-1986). Principe d'incertitude, bases Hilbertiennes et algebres d'operateurs. Bourbaki Seminar, no. 662.

Meyer, Y. (1993). *Wavelets: algorithms and applications.* SIAM. Translated and revised by Robert D. Ryan.

Morlet, J., Arens, G., Fourgeau, E., and Giard, D. (1982). Wave propagation and sampling theory. *Geophysics*, **47**, 203–236.

Muzy, J. F., Bacry, E., and Arneodo, A. (1993). Multifractal formalism for fractal signals: The structure-function approach versus the wavelet-transform modulus maxima method. *Physical Review E*, **47**(2), 875–884.

Neidell, N. S. and Taner, M. T. (1971). Semblance and other coherency measures for multichannel data. *Geophysics*, **36**(3), 482–497.

Nissen, S. E., Haskell, N. L., Lopez, J. A., Donlon, T. J., and Bahorich, M. S. (1995). 3-D seismic coherency techniques applied to the identification and delineation of slump features. In *65th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 1535–1536.

Nottale, L., Schumacher, G., and Gay, J. (1997). Scale relativity and quantization of the solar system. *Astronomy and Astrophysics*, **322**, 1018–1025.

Pai, D. M. (1985). A new solution for wave equations in inhomogeneous media. *Geophysics*, **50**, 1541–1547.

Pai, D. M. (1988). Generalized f-k (frequency-wavenumber) migration in arbitrarily varying media. *Geophysics*, **53**, 1547–1555.

Pai, D. M. (1991). Wave propagation in inhomogeneous media: a planewave layer interaction method. *Wave Motion*, **13**, 205–209.

Peddy, C. P., Sengupta, M. K., and Fasnacht, T. L. (1995). AVO analysis in high-impedance sandstone reservoirs. *The Leading Edge*, **14**(8), 871–877.

Perelomov, A. M. (1972). Coherent states for arbitrary Lie groups. *Comm. Math. Phys.*, **26**, 222–236.

Picinbono, B. and Martin, W. (1983). Représentation des signaux par amplitude et phase instantanées. *Annales des Télécommunications*, **38**, 179–190.

Rappin, D. and de Bazelaire, E. (1992). Application of the wavelet transform to time-frequency filtering, multiples removal and lower crust frequency studies. In *54th Mtg. Eur. Assoc. Expl. Geophys., Extended Abstracts*, pages 188–189.

Rayleigh, J. W. S. (1877). *The theory of sound*, volume 1. Dover Publications. Reprint

1945.

Rayleigh, J. W. S. (1878). *The theory of sound*, volume 2. Dover Publications. Reprint 1945.

Raz, S. (1987). Beam stacking: A generalized processing technique. *Geophysics*, **52**, 1199–1210.

Reed, M. and Simon, B. (1972). *Functional analysis*, volume I of *Methods of modern mathematical physics*. Academic Press.

Reed, M. and Simon, B. (1975). *Fourier analysis, self-adjointness*, volume II of *Methods of modern mathematical physics*. Academic Press.

Reed, M. and Simon, B. (1978). *Analysis of operators*, volume IV of *Methods of modern mathematical physics*. Academic Press.

Reed, M. and Simon, B. (1979). *Scattering theory*, volume III of *Methods of modern mathematical physics*. Academic Press.

Reiter, E. C. (1996). A quantitative comparison of 1, 2, 3 dimensional wavelet compression methods for seismic data. In *66th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 1630–1633.

Renardy, M. and Rogers, R. C. (1993). *An introduction to partial differential equations*, volume 13 of *Texts in applied mathematics*. Springer-Verlag.

Rietveld, W. E. A. (1995). *Controlled illumination in prestack seismic migration*. Ph.D. thesis, Delft University of Technology.

Robertson, J. D. and Nogami, H. H. (1984). Complex seismic trace analysis of thin beds. *Geophysics*, **49**(4), 344–352.

Rosenfeld, A. and Thurston, M. (1971). Edge and curve detection for visual scene analysis. *IEEE Trans on Computers*, **20**, 562 569.

Saito, N. (1994). *Local feature extraction and its applications using a library of bases*. Ph.D. thesis, Yale University.

Schwartz, L. (1951). *Théorie de distributions*, volume 1. Herrmann, Paris.

Schwartz, L. (1952). *Théorie de distributions*, volume 2. Herrmann, Paris.

Shannon, C. E. (1949). Communication in the presence of noise. In *Proceedings of the I.R.E.*, pages 10–21.

Shubin, M. A. (1987). *Pseudodifferential operators and spectral theory*. Springer-Verlag.

Song, S., Qu, C., and He, J. (1996). Imaging of velocity singularities with multiscale operators. *J. Math. Phys.*, **37**(2).

Souillard, B. (1986). Waves and electrons in inhomogeneous media. In J. Souletie, J. Vannimenus, and R. Stora, editors, *Chance and matter*, pages 305 376. Les Houches session XLVI, North Holland.

Staal, J. J. (1995). *Characterizing the irregularity of measurements by means of the wavelet transform: a preliminary discussion on the implications of scaling/nondifferentiability on the dynamics of waves*. Master's thesis, Delft University of Technology.

Steeghs, T. P. H. (1997). *Local power spectra and seismic interpretation.* Ph.D. thesis, Delft University of Technology.

Steinberg, B. Z. (1994). A multiresolution theory of scattering and diffraction. *Wave Motion*, **19**, 213–232.

Steinberg, B. Z. and Leviatan, Y. (1993). On the use of wavelet expansions in the method of moments. *IEEE Trans. Ant. Prop.*, **41**(5), 610–619.

Steinberg, B. Z. and McCoy, J. J. (1994). Toward local effective parameter theories using multiresolution decomposition. *J. Acoust. Soc. Am.*, **96**(2), 1130–1143.

Stewart, G. W. (1931). Problems suggested by an uncertainty principle in acoustics. *J. Acoust. Soc. Am.*, **2**, 325.

Stoffa, P. L., Fokkema, J. T., de Luna Freire, R. M., and Kessinger, W. P. (1990). Split-step Fourier migration. *Geophysics*, **55**(4), 410–421.

Stolt, R. H. (1978). Migration by Fourier transform. *Geophysics*, **43**(1), 23–48.

Stolt, R. H. and Weglein, A. B. (1985). Migration and inversion of seismic data. *Geophysics*, **50**(12), 2458–2472.

Sweldens, W. (1996). The lifting scheme: A custom-design construction of biorthogonal wavelets. *Appl. Comput. Harmon. Anal.*, **3**(2), 186–200.

Taner, M. T., Koehler, F., and Sheriff, R. E. (1979). Complex seismic trace analysis. *Geophysics*, **44**(6), 1041–1063.

Tarantola, A. (1984). Linearized inversion of seismic reflection data. *Geoph. Prosp.*, **32**, 998–1015.

ter Haar Romeny, B. M. (1994). *Geometry-driven diffusion in computer vision.* Kluwer Academic Publishers.

Thorbecke, J. W. (1997). *Common focus point technology.* Ph.D. thesis, Delft University of Technology.

Ursin, B. (1983). Review of elastic and electromagnetic wave propagation in horizontally layered media. *Geophysics*, **48**(8), 1063–1081.

van Stralen, M. J. N. (1997). *Directional decomposition of electromagnetic and acoustic wave-fields: applications in integrated optics, exploration seismics and underwater acoustics.* Ph.D. thesis, Delft University of Technology.

van Wijngaarden, A. J. (1997). Private communication.

Červený, V., Popov, M. M., and Pšenčik, I. (1982). Computation of wave fields in inhomogeneous media – Gaussian beam approach. *Geoph. J. Roy. Astr. Soc.*, **70**, 109–128.

Verhelst, F. (1997). Private communication.

Verhelst, F. and van Wijngaarden, A. . (1997). Iterative search for a pre-specified scale in seismic time: A scale dependent ray-tracer. Technical report, Delft University of Technology. DELPHI, Volume VIII, Chapter 18.

Verm, R. and Hilterman, F. (1995). Lithology color-coded seismic sections: The calibration of AVO crossplotting to rock properties. *The Leading Edge*, **14**(8), 847–853.

Vetterli, M. and Kovačević, J. (1995). *Wavelets and subband coding.* Prentice Hall.

Vilenkin, N. (1972). *Functional analysis.* Noordhoff.

von Neumann, J. (1932). *Matematische Grundlagen der Quantenmechanik.* Springer-Verlag.

Wagner, R. L. and Chew, W. C. (1995). A study of wavelets for the solution of electromagnetic integral equations. *IEEE Trans. Ant. Prop.*, **43**(8), 802–810.

Wang, B. and Pann, K. (1996). Kirchhoff migration of seismic data compressed by matching pursuit decomposition. In *66th Ann. Intern. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, pages 1642–1645.

Wapenaar, C. P. A. (1996a). Inversion versus migration: a new perspective to an old discussion. *Geophysics*, **61**(3), 804–814.

Wapenaar, C. P. A. (1996b). One-way representations of seismic data. *Geophys. J. Int.*, **127**, 178–188.

Wapenaar, C. P. A. (1997a). Amplitude-versus-angle behaviour of self-similar interfaces. *Geophysics.* Submitted for publication.

Wapenaar, C. P. A. (1997b). Reciprocity properties of one-way propagators. *Geophysics.* Accepted for publication.

Wapenaar, C. P. A. and Berkhout, A. J. (1989). *Elastic wave field extrapolation: redatuming of single- and multi-component seismic data.* Advances in exploration geophysics 2. Elsevier Amsterdam.

Wapenaar, C. P. A. and Grimbergen, J. L. T. (1996). Reciprocity theorems for one-way wavefields. *Geophys. J. Int.*, **127**, 169–177.

Wapenaar, C. P. A., van Geloven, W. J. F., van der Leij, T. S., and van Wijngaarden, A. J. (1997). AVA and the effects of fine-layering. *Geophysics.* Submitted for publication.

Webster (1988). Webster's ninth new collegiate dictionary and Webster's collegiate thesaurus. The NeXT Digital Edition.

Weinberg, H. and Burridge, R. (1974). Horizontal ray theory for ocean acoustics. *J. Acoust. Soc. Am.*, **55**, 63.

Weyl, H. (1931). *Gruppentheorie und Quantummechanik.* Hirzel, Leipzig, 2nd edition.

Wilcox, C. H. (1984). *Sound propagation in stratified fluids.* Springer-Verlag.

Witkin, A. P. (1983). Scale space filtering. In *Proc. Int. Joint. Conf. Artif. Intell.*, pages 1019–1023.

Young, R. M. (1980). *An introduction to nonharmonic Fourier series.* Academic Press, New York.

Zemanian, A. H. (1965). *Distribution theory and transform analysis.* International series in pure and applied mathematics. McGraw-Hill.

Subject Index

Author Index*

*First authors only

Samenvatting

Seismische-dataverwerking door middel van de wavelet-transformatie†

Een exploratie-geofysicus houdt zich bezig met het lokaliseren en karakteriseren van fossiele-brandstofvoorraden in de ondergrond. Door een toenemende vraag naar energie moet deze taak efficiënter en nauwkeuriger volbracht worden. Het onderhavige proefschrift beschrijft de rol van de wavelettransformatie hierin. De nadruk wordt gelegd op de betekenis van de wavelettransformatie voor de bewerking en analyse van seismische data. Bij deze verwerking wordt een onderscheid gemaakt tussen (1) de acquisitie van seismische data, (2) de voorbewerking van de seismische data, (3) de seismische beeldvorming van de ondergrond, (4) de geologische karakterisatie van de ondergrond en (5) de evaluatie van het potentiële reservoir.

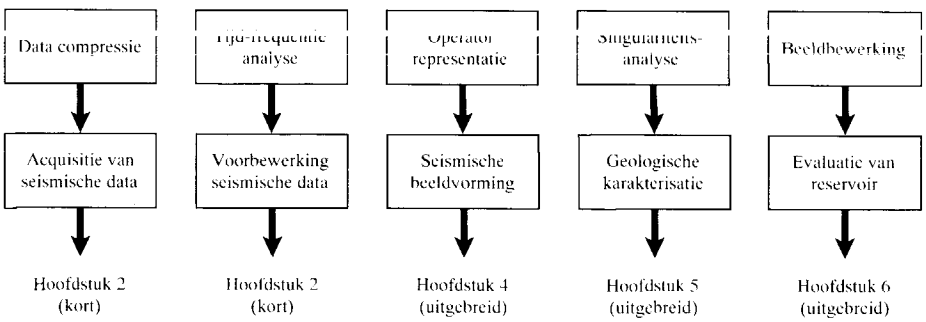
In hoofdstuk 2 worden algemene lineaire transformaties geïntroduceerd. Deze worden onderverdeeld in twee categorieën. De eerste categorie wordt geassocieerd met zelf-geadjungeerde operatoren. Deze operatoren leiden tot orthogonale en omkeerbare transformaties. De tweede categorie wordt gerelateerd aan representaties van het wiskundige begrip groep. Een groep leidt tot een redundante familie van coherente toestanden. Onder deze categorie vallen de Gabor-transformatie en de wavelettransformatie. De wavelettransformatie is een inwendig produkt van een signaal met een set van analysefuncties die een constante vorm hebben, maar een variabele breedte en een variabele lokatie. Het gemiddelde van elk van deze analysefuncties is nul. De breedte wordt aangeduid door het begrip schaal. Na een behandeling van de belangrijkste eigenschappen van de wavelettransformatie wordt hoofdstuk 2 besloten met een korte discussie over de toepassingen van de wavelettransformatie

†Dutch summary of *A wavelet transform approach to seismic processing.*

op de seismische-dataverwerking, zoals in de eerste alinea besproken. Het uitgangspunt van de toepassingen is de eigenschap dat de wavelettransformatie "blind" is voor een signaal dat niet of langzaam verandert, en de eigenschap dat de wavelettransformatie op een natuurlijke wijze een signaal in details en benaderingen op verschillende schalen ontbindt. De onderstaande figuur vat de toepassingen van de wavelettransformatie samen.

In hoofdstuk 3 wordt de golfvergelijking voor één-weg golven afgeleid. Deze dient als basis voor de afleiding van de één-weg representatie van seismische reflectiedata, hetgeen het zogenaamde WRW-model oplevert. Dit is een model in termen van propagatie-operatoren en een reflectie-operator. De propagatie-operatoren zijn functies van de Helmholtz-operator. Aan functies van een operator wordt in hoofdstuk 3 een betekenis gegeven met behulp van een spectrale decompositie. Dit is een decompositie van het golfveld in eigenfuncties. Essentieel hierbij is dat de Helmholtz-operator een zelf-geadjungeerde operator is. Het gebruik van een spectrale decompositie heeft twee voordelen. Ten eerste geeft zij een nauwkeurige, maar niet zeer efficiënte, representatie van de propagatie-operatoren, ook voor een sterk lateraal variërende ondergrond. Een synthetisch voorbeeld laat zien dat deze representatie een accurate beeldvorming van de ondergrond mogelijk maakt. Ten tweede levert zij door middel van het spectrum een beter begrip van golfpropagatie op.

De betekenis van de wavelettransformatie voor de in hoofdstuk 3 afgeleide representatie van seismische reflectiedata wordt in hoofdstuk 4 geanalyseerd. De representatie van seismische reflectiedata kan gegeneraliseerd worden door gebruik te maken van een operatornotatie. De transformaties die in hoofdstuk 2 geïntroduceerd zijn, kunnen nu gebruikt worden om seismische reflectiedata in andere domeinen te representeren. Er wordt in detail een datarepresentatie in termen van wavelets uit-



Een overzicht van de toepassingsgebieden van de wavelettransformatie (bovenste rij), de corresponderende stappen bij de analyse en bewerking van seismische data (middelste rij) en de hoofdstukken waar de verschillende toepassingen worden besproken (onderste rij).

gewerkt. Dit leidt tot een onderverdeling van de seismische data in een set van multi-schaalexperimenten. De datarepresentatie in het wavelet domein wordt gebruikt als uitgangspunt voor een beeldvormingsschema met behulp van wavelets. Een voordeel van de voorgestelde methode is het feit dat de seismische beeldvorming op verschillende schalen kan worden uitgevoerd. Een duidelijk nadeel is het feit dat de propagatie-operatoren geen ijle representatie in het wavelet domein hebben.

De betekenis van de wavelettransformatie voor het karakteriseren van reflectie-eigenschappen van de ondergrond wordt in hoofdstuk 5 besproken. Hier worden snelheidsprofielen beschouwd bestaande uit de som van een constante achtergrondsnelheid en een functie met een lokale geïsoleerde singulariteit. De singulariteiten zijn met name enkelzijdige homogene distributies. Met behulp van een wavelettransformatie-modulus-maxima-analyse (WTMM-analyse) en een fase-analyse is het mogelijk enkelzijdige homogene distributies uniek te karakteriseren in termen van een singulariteitssterkte α en een geïnduceerde fase φ. Dit leidt tot de zogenaamde $(\varphi - \alpha)$-diamant. De reflectie aan dergelijke snelheidsprofielen voor loodrechte inval is in sterke mate afhankelijk van de achtergrondsnelheid. Als de singuliere verstoring veel kleiner is dan de achtergrondsnelheid, dan is de soort van singulariteit uit het karakter van de gereflecteerde respons af te leiden. Anderzijds, als de achtergrondsnelheid veel kleiner is dan de singuliere verstoring, dan geeft het karakter van de gereflecteerde golf geen uitsluitsel over de soort van singulariteit. Op basis van akoestische impedanties, berekend uit snelheids- en dichtheidsmetingen in boorgaten, wordt geconcludeerd dat de singuliere verstoring in het algemeen kleiner is dan de achtergrondsnelheid. De numeriek gevalideerde analyse leidt tot de formulering van een equivalente Zoeppritz-grenslaag en van een singulariteiten-gestuurde inversie (SDI).

Ten slotte wordt in hoofdstuk 6 de rol van de wavelettransformatie bij de interpretatie van 3-D gemigreerde seismische data beschreven. Hierbij is het essentieel dat de wavelettransformatie "blind" is voor iets wat langzaam verandert op de schaal waarop het wavelet "kijkt". Deze eigenschap helpt om uit een 3-D volume stratigrafische gegevens en breuken in de geologische lagen te extraheren die normaal gesproken slecht zichtbaar zijn. De wavelettransformatie wordt hier toegepast in drie onafhankelijke richtingen. Dit geeft de mate van verandering in de drie richtingen. Een kwadratische optelling van de drie veranderingen geeft de totale verandering, op een specifieke schaal en op een specifieke lokatie. Uit het voorbeeld met een gemigreerde dataset uit de Golf van Mexico blijkt dat breuken in de geologische lagen en met name rivierbeddingen uitstekend gelokaliseerd kunnen worden met deze efficiënte methode.

Curriculum vitae

Geboren te Rotterdam op 5 februari 1966, groeide ik, niet al te ver van de Kuip, op in die stad. Vanaf 1978 volgde ik middelbaar onderwijs. In 1984 legde ik het eindexamen gymnasium-β af aan het Sint-Montfort college te Rotterdam. In datzelfde jaar richtte ik mijn blik ten hemel om de studie sterrenkunde aan de Rijks Universiteit Leiden aan te vangen. Aan de eisen voor het propaedeutisch examen voldoed ik een klein jaar later. Daarna schroefde ik het studietempo omlaag en zette mijn zinnen op extracurriculaire activiteiten, waarbij met name het bestuurslidmaatschap van de L.V.V.S. Augustinus en dat van het A.H.G. Aurelius niet onvermeld mogen blijven. Uiteindelijk rondde ik de studie in februari 1991 af. De afstudeerscriptie, geschreven onder begeleiding van dr. J. Lub, handelde over de Baade-Wesselink methode toegepast op RRLyrae-sterren. Als onderdeel van het afstudeerwerk heb ik een maand de Nederlandse telescoop bij de European Southern Observatory in Chili bemand. Na mijn afstuderen bracht ik met veel plezier een jaar door bij Noordwijk Space Expo, het bezoekerscentrum van ESTEC, als muscaal medewerker.

En toen gebeurde het. De blik werd ter aarde geworpen voor een promotieonderzoek in de exploratie-geofysica aan de faculteit der Technische Natuurkunde, onder leiding van prof. dr. ir. A. J. Berkhout. Het onderzoek vond plaats binnen de kaders van het DELPHI-consortium. De originele doelstelling was om een inzicht te krijgen in de mogelijkheden die verschillende wiskundige transformaties bieden voor een efficiënte en nauwkeurige beeldvorming van de ondergrond. In de loop der tijd is deze doelstelling langzaam gemodificeerd tot het de vorm kreeg die heeft geleid tot dit proefschrift.

Gedurende het onderzoek heb ik verschillende internationale congressen in Europa en de Verenigde Staten bezocht en heb ik bij halfjaarlijkse consortiumbijeenkomsten en bij een aantal oliebedrijven afzonderlijk, verslag gedaan van mijn bevindingen. Daarnaast heb ik twee studenten begeleid gedurende hun afstudeerperiode.

Vanaf januari 1998 zal ik de opgedane kennis ten dienste stellen van Shell International Exploration and Production.

Acknowledgment/Dankwoord

To all people who contributed directly or indirectly to the realization of this thesis I am indebted. At the international side, the financial support of the sponsors of the DELPHI-consortium is gratefully acknowledged. The sponsors not only provided the financial support, but they also encouraged me to continue on the wavelet road. Most notably, I would like to mention Cheryl Bosman and Warner Ross with Exxon, Walter Rietveld and Kurt Marfurt with Amoco, Joe Higginbotham with Texaco, for allowing me to visit their offices. Furthermore, I am obligated to SAGA and Schlumberger Geco-Prakla for providing the 2-D and 3-D data set, respectively. Walter Rietveld is acknowledged for the mediation.

Aan de Nederlandse kant ben ik veel dank verschuldigd aan prof. Berkhout en Kees Wapenaar voor hun gezamenlijke begeleiding en voor de ideeën die ze in de loop der jaren hebben aangedragen. Buiten de vakgroep, ben ik dank verschuldigd aan prof. dr. ir. J.T. Fokkema, prof. dr. J.J. Duistermaat en dr. B. de Pagter voor discussies over verschillende onderwerpen behandeld in dit proefschrift. Alle medepromovendi en alle studenten van de afgelopen vijfeneenhalf jaar wil ik bedanken voor de uitstekende tijd die ik met hen in de groep heb gehad. De bijdrage van Joris Grimbergen verdient een speciale vermelding hier. De staf van de vakgroep zorgde voor de infrastructurele ondersteuning.

Speciale dank gaat uit naar de wavelet-mannen van het eerste uur, te weten Felix Herrmann, Joes Staal en Edo Hoekstra. Felix, samen met Etienne hebben we ten minste een jaar van onze tijd verbruikt, maar uiteindelijk was dat het meer dan waard. Joes, jouw constructieve commentaar werd te allen tijde zeer op prijs gesteld, dat dan toch weer wel. Edo, jouw bijdragen waren cruciaal, grote dank.

Dank uiteraard ook aan mijn vrienden, die mij in contact hielden met het reguliere buiten-Delftse leven.

Het thuisfront, te weten de zussen en natuurlijk mijn moeder, hebben mij door de jaren heen gesteund in mijn bezigheden. Aan jullie ben ik dan ook veel dank verschuldigd.

Ten slotte, Joossien, jij mag dit proefschrift zien als een langgerekt huwelijksaanzoek. Zullen we trouwen?