

**Mathematical Signal Analysis:
Wavelets, Wigner Distribution and a
Seismic Application**

P.J. Oonincx

**MATHEMATICAL SIGNAL ANALYSIS:
WAVELETS, WIGNER DISTRIBUTION
AND
A SEISMIC APPLICATION**

**Mathematical Signal Analysis:
Wavelets, Wigner Distribution and a Seismic Application**

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof.dr. J.J.M. Franse
ten overstaan van een door
het college voor promoties ingestelde commissie,
in het openbaar te verdedigen in de Aula der Universiteit
op woensdag 9 februari 2000, te 13.00 uur

door

Peter Jacobus Oonincx
geboren te Kerkrade

Promotor:

Prof.dr. T.H. Koornwinder Universiteit van Amsterdam

Co-promotor:

Dr. N.M. Temme Centrum voor Wiskunde en Informatica

Promotiecommissie:

Prof. dr. J.-P. Antoine	Université Catholique de Louvain
Prof. dr. ir. J. Grasman	Landbouwwuniversiteit Wageningen
Prof. dr. M.S. Keane	Universiteit van Amsterdam
Dr. H.G. ter Morsche	Technische Universiteit Eindhoven
Prof.dr. E.M. Opdam	Universiteit van Amsterdam
Dr. J.J.O.O. Wiegerinck	Universiteit van Amsterdam

Het onderzoek dat tot dit proefschrift heeft geleid werd mede mogelijk gemaakt door de Technologiestichting STW en valt onder de auspiciën van het Thomas Stieltjes Instituut voor Wiskunde.

A mathematician is a machine ...
... for turning coffee into theorems

- - *Paul Erdős* - -

To Birgit and Sacha

Contents

1	Introduction	1
1.1	Time-Frequency Methods	1
1.2	Applications of the Wavelet Transform	13
1.2.1	Denoising	13
1.2.2	Compression	14
1.2.3	Wavelet-Galerkin methods	14
1.3	Some Main Results	15
2	Time-Frequency Analysis	19
2.1	The Fourier Transform	19
2.1.1	The Fourier Transform on L^1	20
2.1.2	The Fourier Transform on L^2	22
2.2	The Windowed Fourier Transform	25
2.3	The Wigner Distribution	33
2.4	Lie Group Theory and the Heisenberg Group	41
3	The Wavelet Transform	45
3.1	The Continuous Wavelet Transform	45
3.2	The Discrete Wavelet Transform	54
3.2.1	Multiresolution Analysis in $L^2(\mathbb{R})$	56
3.2.2	MRA and Filter Banks	58
3.2.3	Implementation of the DWT and its Use for $l^2(\mathbb{Z})$	62
4	A Framework for Multiresolution Analysis in Hilbert Spaces	67
4.1	Frames and Riesz Systems	67
4.2	Multiresolution Analysis in Hilbert Spaces	72
4.3	Riesz Systems Generated by Unitary Operators	76
4.4	Riesz Bases in $l^2(\mathbb{Z}^n)$	78
4.5	MRA and Riesz Bases in Hilbert Spaces	84

5	The FRFT and Affine Transformations in the Wigner Plane	89
5.1	The Fractional Fourier Transform	89
5.1.1	Definition and Properties	90
5.1.2	The FRFT and the Wigner Plane	93
5.2	Affine Transformations in the Wigner Plane	95
5.2.1	A Group Theoretical Approach	99
5.2.2	The FRFT Generalized	106
5.3	A Representation Formula	109
6	Localization Problems in Phase Space	119
6.1	Slepian's Energy Problem	120
6.2	Energy Concentration on a Circle in the Wigner Plane	125
6.3	Localization Problems and the Generalized FRFT	130
6.3.1	The Rectangle/Parallelogram Case and the Rihaczek Distribution	131
6.3.2	The Circle/Ellipse Case	134
7	A Seismic Problem: Automatic S-Phase Detection with the DWT	137
7.1	Introduction and Approach of the Seismic Problem	137
7.2	Seismic methods for S-Phase Picking	139
7.2.1	Characteristic functions based on a cross-power matrix	139
7.2.2	Rotation and the Energy Ratio	141
7.3	On the Use of Characteristic Functions	142
7.3.1	Error Analysis of the Characteristic Functions	142
7.3.2	Problems in Analysing Seismic Data	146
7.4	The Wavelet Method	147
7.4.1	Characteristic Functions and the DWT	148
7.4.2	A Set-Up for the DWT Analysis	152
7.5	Examples and Results	153
	Bibliography	159
	Index	167
	Samenvatting/Dutch Summary	171
	Dankwoord/Acknowledgements	173
	Curriculum Vitae	175

Chapter 1

Introduction

This thesis studies, from a mathematical point of view, a number of transformations that map functions $f \in L^2(\mathbf{R})$, i.e., square integrable functions on \mathbf{R} , to functions $\tilde{f} \in L^2(\mathbf{R}^2)$. In particular we study transformations that are related to the Fourier transform in the following sense.

The Fourier transform maps a function f to a function \hat{f} . For a function of time (a signal), \hat{f} represents the intensity of the fluctuations (frequencies) in the signal f . Analysing a signal in this way is called spectral analysis. Besides the representation in time f and the representation in frequency \hat{f} , there exists transformations $f \mapsto \tilde{f}$ to represent a signal both in time and in frequency. Some of these transformations are discussed extensively in this thesis, namely the windowed Fourier transform, the Wigner distribution, the Rihaczek distribution, the fractional Fourier transform and the wavelet transform. In this introduction we will briefly introduce these transformations.

This introduction also considers applications of the wavelet transform in the field of signal and image processing. We show, that the wavelet transform can be a very useful tool for denoising algorithms, data compression techniques, and numerical analysis.

Finally, this introductory chapter summarizes the contents of this thesis.

1.1 Time-Frequency Methods

In 1822 Fourier published his famous work *Theorie analytique de la Chaleur*, see [31] for an English translation. In this work he stated that a periodic function f could be expressed as

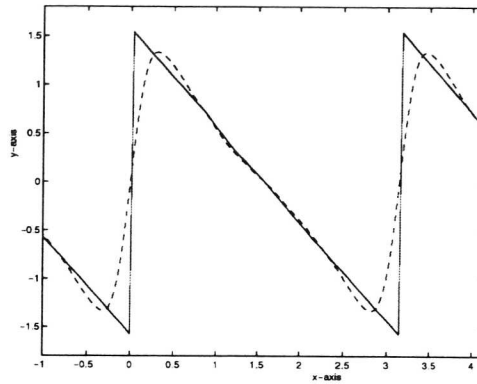


Figure 1.1: Approximation of a discontinuous function by a Fourier series

the sum of trigonometric functions

$$f(x) = a_0 + \sum_{k=1}^{\infty} (a_k \cos(k\omega_0 x) + b_k \sin(k\omega_0 x)), \quad (1.1)$$

for some $a_0, a_k, b_k \in \mathbb{R}$. Here $\omega_0 = 2\pi/T$, with T the period of f . The series in (1.1) is called the Fourier series of f . In the literature the Fourier series of f is mostly given by

$$f(x) = \sum_{k \in \mathbb{Z}} c_k e^{ik\omega_0 x}. \quad (1.2)$$

Not only periodic functions can be expressed in terms of their Fourier series. Also compactly supported functions can be written in this way. To do this, we extend such a function f to a periodic function and compute its Fourier series. This Fourier series, regarded on the support of f , is then said to be the Fourier series of f .

Fourier's idea was that also a discontinuous function f could be expressed in this way, namely as the sum of continuous function. Later, Dirichlet formulated necessary and sufficient conditions such that (1.1) holds pointwise. In Figure 1.1, an example of a discontinuous function and its Fourier series is depicted. As an example we have taken

$$f(x) = \begin{cases} \frac{\pi}{2} - x, & x \in (0, 2\pi), \\ 0, & x = 0, \end{cases}$$

and with $f(x + 2\pi) = f(x)$, for all $x \in \mathbb{R}$. Its Fourier series is given by

$$f(x) = \sum_{k=1}^{\infty} \sin(kx)/k.$$

We observe, that the coefficients a_k in (1.1) all vanish in this example, since f is an odd function. A good approximation of f is already established by a partial sum

$$S_N = \sum_{k=1}^N \sin(kx)/k,$$

with $N > 0$ relatively small. In fact, Figure 1.1 shows f and its approximation by means of the partial sum S_8 .

Analysing signals by means of Fourier series, called spectral analysis, is nowadays a standard technique to obtain additional information of a signal. However, in our era, signals we want to analyse are often not continuous in time. These discrete-time signals are mostly the result of a sampling procedure built into the measurement equipment, that measures an incoming physical signal. To deal with these signals, we can use the discrete Fourier transform (DFT) instead of the Fourier series. The DFT is given by a discretisation (1.2). Particularly, if the signal has been measured during a finite time interval, the number of samples of such a signal is finite and its DFT simplifies to a polynomial on the unit circle.

In this particular case, the DFT can be computed in a fast way by means of the butterfly algorithm, see [19]. This algorithm arranges the Fourier coefficients in such a way, that they can be computed recursively. The complexity of this algorithm is given by $\mathcal{O}(N \log N)$, with N the number of samples in the signal. Computing the DFT in this manner is called the fast Fourier transform (FFT). The existence of such a fast algorithm is an important reason why Fourier analysis has become a standard tool in signal analysis.

For non-periodic functions the Fourier transform provides a tool for spectral analysis. This transform is given by

$$f \mapsto \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x) e^{-ix} dx.$$

The Fourier transform \hat{f} of f can be considered initially for functions that belong to a class of rapidly decreasing functions, called the Schwartz class $\mathcal{S}(\mathbb{R})$, see [88, 89]. Each function $f \in \mathcal{S}(\mathbb{R})$ can also be recovered from its Fourier transform \hat{f} . This means that the Fourier transform indeed offers an alternative way for representing a function $f \in \mathcal{S}(\mathbb{R})$. The Fourier transform can be extended to functions in $L^1(\mathbb{R})$, i.e., absolutely integrable functions on \mathbb{R} , or functions in $L^2(\mathbb{R})$. A unique reconstruction of the original function in $L^1(\mathbb{R})$ or $L^2(\mathbb{R})$ from its Fourier transform is also possible.

After taking the Fourier transform of f , the value $\hat{f}(\omega_0)$ represents the complex-valued amplitude by which a frequency ω_0 appears in the signal f . However, we cannot read off from $\hat{f}(\omega_0)$ at which time intervals the frequency ω_0 appears in f . So, \hat{f} is not localized both in

time and in frequency.

For localizing a signal simultaneously in time and frequency we can use transforms $f \mapsto \tilde{f}$, with $f \in L^2(\mathbb{R})$ and $\tilde{f} \in L^2(\mathbb{R}^2)$, representing the signal both in time and in frequency. However, although such a transform can improve localization compared to the Fourier transform, we are not able to localize time and frequency arbitrarily well. The limitations on simultaneous time-frequency analysis are given by the time-frequency equivalent of Heisenberg's uncertainty relation

$$\int_{\mathbb{R}} x^2 |f(x)|^2 dx \cdot \int_{\mathbb{R}} \omega^2 |\hat{f}(\omega)|^2 d\omega \geq \|f\|_2^4 / 4. \quad (1.3)$$

Equality in (1.3) is established only for the constant multiples of $f(x) = e^{-cx^2}$, for $c > 0$. Originally Heisenberg presented relation (1.3) in quantum mechanics as a relation between the standard deviation of position and the standard deviation of momentum. In his famous paper [34], Gabor translated this relation in terms of time-frequency analysis.

A natural starting point for an overview of time-frequency transformations is the windowed Fourier transform (WFT). The idea of the WFT is to multiply a signal f by a window function h and then to take the Fourier transform of the product function. By translating such a well localized window h along the signal, the WFT is able to analyse the frequency behavior of f during the time interval for which h is localized. Translated into a representation formula the WFT reads

$$\mathcal{F}_h[f](x, \omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) \overline{h(y-x)} e^{-i\omega y} dy. \quad (1.4)$$

Since $\mathcal{F}_h f$ can assume complex values, mostly the spectrogram of a signal is used to analyse the signal's behavior in time and frequency. The spectrogram of f is given by $|\mathcal{F}_h[f](x, \omega)|^2$.

It follows from (1.4), that $h(x)$ should be more or less concentrated around $x = 0$. Furthermore, the behavior of $\mathcal{F}_h f$ in both time and frequency is strongly influenced by the window function h . This suggests that we have to deal with the problem of finding a window function h , that is both well localized and for which $\mathcal{F}_h f$ is a good reproduction of the time-frequency behavior of f .

Since no information of a signal f is thrown away by representing it by means of the WFT, f can also be reconstructed from \mathcal{F}_f . We observe, that $f(x_0)$ contributes to $\mathcal{F}_h(x, \omega)$, for all $x \in \mathbb{R}$, for which

$$f(x_0) \overline{h(x_0 - x)} \neq 0.$$

Consequently, f is represented redundantly by \mathcal{F}_h , which means that no unique reconstruction formula exists to recover f from its WFT, like we have for the Fourier transform.

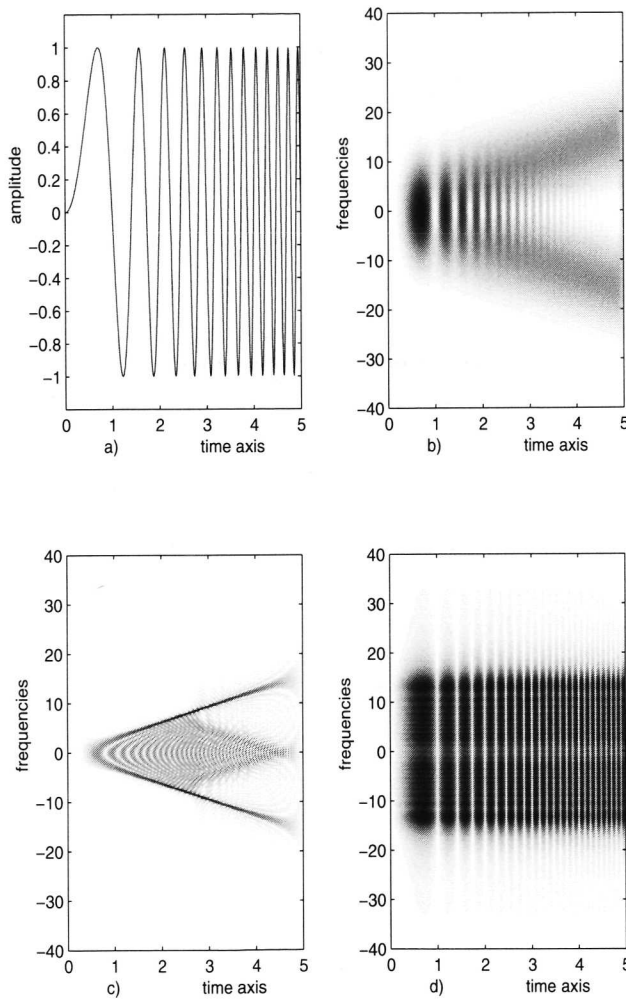


Figure 1.2: Time-frequency representations of a quadratic chirp: fig. a) the original signal, fig. b, c, d) its time-frequency representation by means of the WFT, the Wigner distribution and the Rihaczek distribution respectively.

A window function, that is optimal in the sense, that it gives the best localization of the WFT in time and frequency will strongly depend on the signal f . In 1946, Gabor suggested in his paper [34] to use a window function, that is optimal in the sense that equality in the Heisenberg relation is established. Such a window function is given by

$$h_\sigma(x) = (\pi\sigma^2)^{-1/4} e^{-x^2/2\sigma^2},$$

where $\sigma > 0$. The constant $(\pi\sigma^2)^{-1/4}$ is chosen such that $\|h_\sigma\|_2 = 1$ for all $\sigma > 0$. The WFT that corresponds to this choice for h is given by the Gabor transform

$$\mathcal{G}_\sigma[f](x, \omega) = (2\pi)^{-3/4} \sigma^{-1/2} \int_{\mathbb{R}} f(y) e^{-(y-x)^2/2\sigma^2} e^{-i\omega y} dy. \quad (1.5)$$

In Figure 1.2.b, the spectrogram of the function $\sin(\pi x^2)$ is depicted. In this picture, the dark grey values indicated high amplitudes of the spectrogram at that particular time-frequency point. The function itself is depicted in Figure 1.2.a. For the spectrogram a Gaussian function is used. Obviously, the spectrogram provides information about the signal's behavior both in time and localization. However, energy is spread instead of being perfectly localized.

Choosing an appropriate window is not the only difficulty we have to deal with, when we use the WFT to analyse a signal. An other problem is to choose the 'width' of the window, i.e., its support for compactly supported windows or the parameter σ for the Gaussian function h_σ . If the signal f contains a frequency component with in a very small time interval, the chosen window width can be too large to detect the localization of such a component with high precision. On the other hand, if the window 'width' is chosen too small, the WFT will not detect very low frequency components in f .

A time-frequency representation that only takes the behavior of the signal itself into account is the Wigner distribution

$$\mathcal{WV}[f](x, \omega) = \frac{1}{2\pi} \int_{\mathbb{R}} f(x + t/2) \overline{f(x - t/2)} e^{-it\omega} dt. \quad (1.6)$$

This representation was already introduced in 1932 by Wigner in his paper [104]. He presented this representation in the field of quantum mechanics. In 1948, Ville introduced the representation in the fields of signal analysis in [99]. Therefore, this representation is also known in the literature as the Wigner-Ville distribution. A comprehensive approach of the Wigner distribution as a tool for time-frequency analysis is provided by a paper by Claasen and Mecklenbräuer [15].

We observe, that the Wigner distribution is in fact the Fourier transform of the autocorrelation function $R_{f,x}$, given by

$$R_{f,x}(t) = f(x + t/2) \overline{f(x - t/2)} / \sqrt{2\pi}.$$

This means that the Fourier transform is taken of the product of a signal with a translated version of itself. Consequently, the Wigner distribution is non-linear and it also represents a signal redundantly in time and frequency. Therefore, a signal can be reconstructed from its Wigner distribution, but this cannot be done in a unique way.

The quadratic character of the Wigner distribution is a problem when analysing a sum of signals. Then interference of the two signals appears in the time-frequency analysis. As a result of this interference it can happen that $\mathcal{WV}[f](x, \omega) \neq 0$, while $f(x) = 0$ for a fixed $x \in \mathbf{R}$. This is the case if $f(x) = 0$ for x in a finite interval.

In Figure 1.2.c, the Wigner distribution of the function $\sin(\pi x^2)$ is depicted. We observe, that this function behaves linearly in the Wigner plane. Since we only computed the Wigner distribution for a finite part of the signal, the edges of this time interval cause distortions, that are visualized between the two lines.

In [17], Leon Cohen presented a general class of time-frequency transformations. A general formula for the transformations in his class is given by

$$\tilde{f}(x, \omega) = \frac{1}{4\pi^2} \int_{\mathbf{R}} \int_{\mathbf{R}} \int_{\mathbf{R}} f(u + t/2) \overline{f(u - t/2)} \phi(v, t) e^{-i(xv + \omega t - uv)} du dt dv. \quad (1.7)$$

Starting from this representation formula, all known time-frequency distributions can be derived by choosing an appropriate kernel function ϕ . In his paper, Cohen also showed that properties of the time-frequency representations are reflected by relatively simple constraints on the kernel function ϕ .

Relation (1.7) turns into the Wigner distribution for $\phi = 1$ and it turns into the spectrogram for

$$\phi(v, t) = \int_{\mathbf{R}} h(u + t/2) \overline{h(u - t/2)} e^{-iuv} du.$$

In both cases, Fourier integrals have to be computed for obtaining the Wigner distribution and the spectrogram from (1.7). A third time-frequency representation, that will be used in the sequel of this thesis, is the Rihaczek distribution. This representation is given by

$$\mathcal{R}[f](x, \omega) = f(x) \overline{\hat{f}(\omega)} e^{-i\omega x} / \sqrt{2\pi}. \quad (1.8)$$

It can be obtained from (1.7) by taking $\phi(v, t) = e^{ivt/2}$. For a comprehensive list of time-frequency distributions and its corresponding kernel functions, we refer to [18]. In Figure 1.2.d, the Rihaczek distribution of the function $\sin(\pi x^2)$ is depicted. We observe, that for this particular signal the localization in the phase plane is poor.

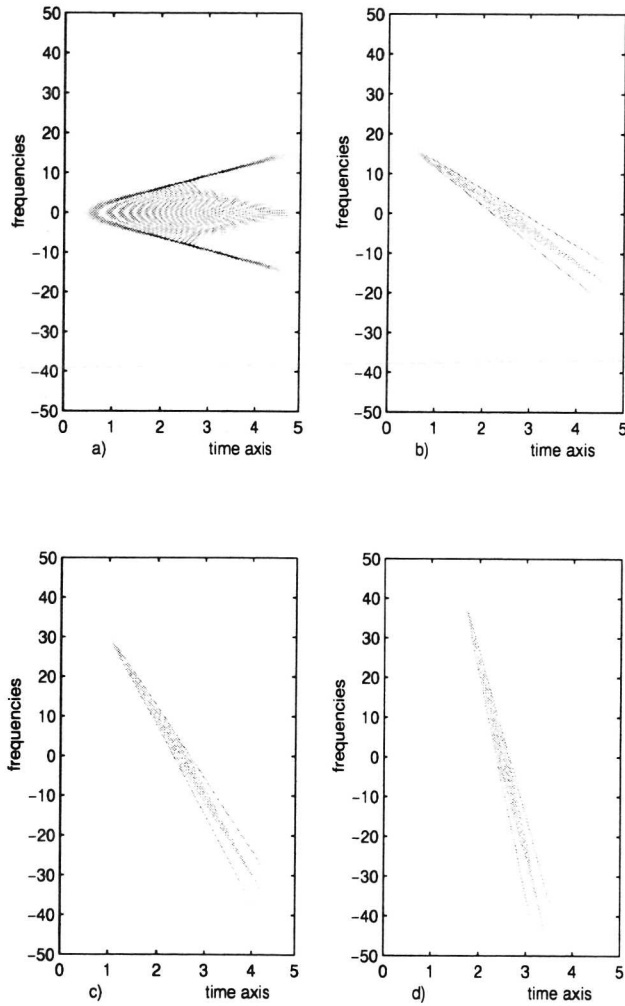


Figure 1.3: The fractional Fourier transform of a quadratic chirp in the Wigner plane: fig. a) $\alpha = 0$, b) $\alpha = \pi/8$, c) $\alpha = \pi/4$ and d) $\alpha = 3\pi/8$.

A representation of a signal in a domain other than the time or frequency domain is given by the fractional Fourier transform (FRFT). This transform is given by

$$\mathcal{F}_\alpha[f](x) = \frac{C_\alpha}{\sqrt{2\pi}|\sin \alpha|} \int_{\mathbb{R}} f(u) e^{i((u^2+x^2) \cdot (\cot \alpha)/2 - ux \csc \alpha)} du, \quad (1.9)$$

for some parameter $\alpha \in \mathbb{R}$ and a constant C_α . This transform was introduced by Namias in 1980. He defined this transform as a fractional power of the Fourier transform,

$$\mathcal{F}_\alpha = \mathcal{F}^\alpha, \quad (1.10)$$

where \mathcal{F} denotes the Fourier transform. Namias derived formula (1.9) starting from his definition (1.10) and using generating functions for Hermite functions, which are eigenfunctions of the Fourier transform. In 1987 Kerr and McBride provided a rigorous mathematical framework for the fractional Fourier transform on $L^2(\mathbb{R})$, see [53, 61].

In 1992 the FRFT became interesting for signal analysis by a paper of Almeida [4]. He showed, that taking the Wigner distribution of $\mathcal{F}_\alpha f$ corresponds to the Wigner distribution of the function f followed by a rotation over an angle α in the Wigner plane, the time-frequency plane that corresponds to the Wigner distribution. In Figure 1.3, this phenomenon is illustrated by taking the FRFT of the signal $\sin(\pi x^2)$ for four different values of α and taking their Wigner distributions. We observe, that Figure 1.3.b, c, d are rotated versions of Figure 1.3.a.

The rotation property of the FRFT inspired mathematicians in the past to study also other transformations in the Wigner plane, that correspond to linear operators on $L^2(\mathbb{R})$. Already before the introduction of the FRFT De Bruijn proposed in [9] a class of operators that are related to linear operators in the Wigner plane. Also we study this problem in this thesis.

The last representation of a signal, that is briefly discussed in this introduction, is the wavelet transform. This transform was introduced in 1984 by Morlet and co-workers, who wanted to analyse geophysical signals with some kind of an adaptive WFT. However, in mathematical circles this transform was not new. It was already known as Caldéron's reproducing formula [10].

The WFT analyses a signal by multiplying it with a sliding window function and then by taking the Fourier transform of this product function. The wavelet transform makes use of the same principle, however the Fourier transform is replaced by a dilation of the window function in $L^2(\mathbb{R})$. In this way the window function can be adapted in a better way to the signal. This transform reads

$$\mathcal{W}_\psi[f](a, b) = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} f(x) \overline{\psi\left(\frac{x-b}{a}\right)} dx, \quad (1.11)$$

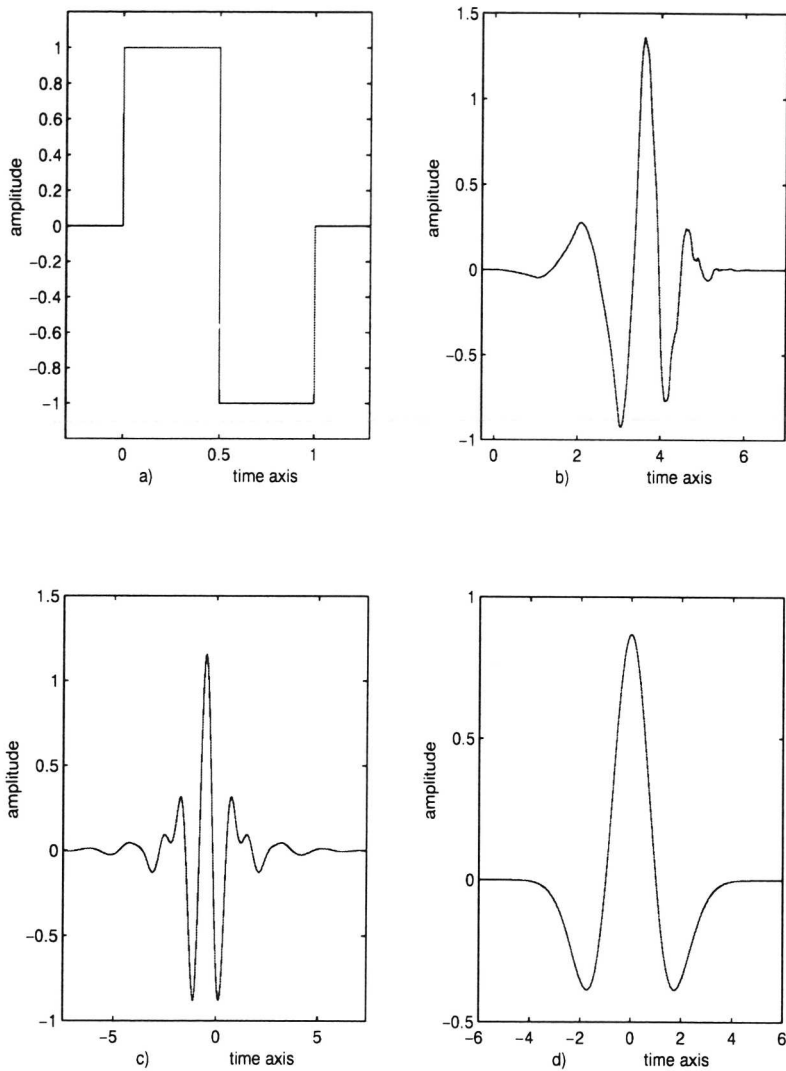


Figure 1.4: Four well-known wavelets: fig. a) the Haar wavelet, b) the Daubechies-4 wavelet, c) the Meyer wavelet, d) the Mexican hat.

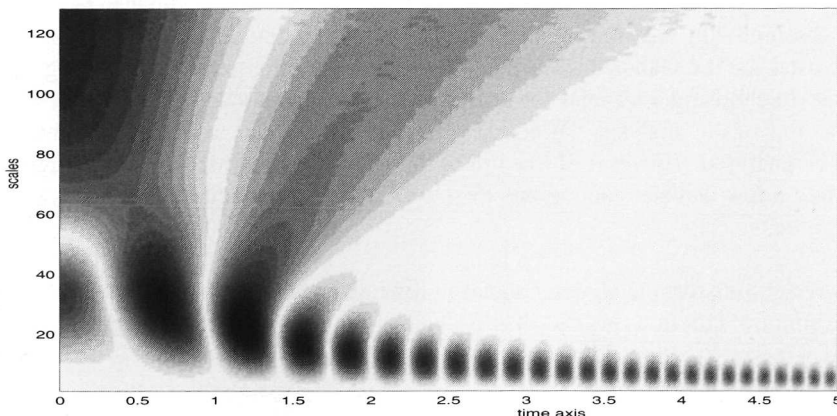


Figure 1.5: The continuous wavelet transform of $\sin(\pi x^2)$.

for some $\psi \in L^2(\mathbb{R})$ and $a \in \mathbb{R}^+$ and $b \in \mathbb{R}$. The window function ψ is called a wavelet if it satisfies the additional condition

$$0 < C_\psi = 2\pi \int_{\mathbb{R}^+} \frac{|\hat{\psi}(a\omega)|^2}{a} da < \infty,$$

for almost all $\omega \in \mathbb{R}$. This condition guarantees the existence of a unique reconstruction formula.

To give an impression, what kind of functions are wavelets, four well known wavelets are depicted in Figure 1.4. The first wavelet we see (fig. a) is the Haar wavelet, which is probably also the oldest known wavelet. In 1910 already Haar used this function for constructing an orthonormal basis in $L^2(\mathbb{R})$ by means of dilations and integer translations of a mother function. We shall recognize this idea later as the discrete wavelet transform. The second wavelet, that is depicted (fig. b), is the Daubechies-4 wavelet. Generally, the Daubechies- N wavelet belongs to a class of wavelets that possess some desirable properties for signal analysis. Also these properties are discussed when we come to the discrete wavelet transform. The index number N denotes the order of regularity of the particular wavelet. We observe, that the Haar wavelet is also a Daubechies wavelet, namely the Daubechies-1 wavelet.

The wavelet in Figure 1.4.c is named after Meyer, who created this Meyer wavelet by starting from some necessary condition on the Fourier transform of the wavelet, e.g. compact support of the Fourier transform. The fourth wavelet (fig. d) is called the Mexican hat, since it looks like a sombrero. In fact, the Mexican hat is the second derivative of the Gaussian function, that is used for the Gabor transform. This wavelet is often used for applying the CWT in physics. In choosing a wavelet for analysing a signal, we are led by the shape of the signal and the aim of the analysis. Of course we want the wavelet to match with the signal, that has to be analysed. However, if the aim of the analysis is to give a time-scale representation with only a few wavelet coefficients $\mathcal{W}_\psi(a, b)$, then a compactly supported wavelet can be appropriate.

The wavelet transform analyses a signal in time and scaling behavior, since it is based on a scaled window function, instead of a frequency representation of the product of the signal and the window function. Therefore, we say that the wavelet transform is a time-scale transformation.

As an example, we analysed the behavior of the quadratic chirp $f(x) = \sin(\pi x^2)$ in time and scale. For this we computed the wavelet transform of this function using the Daubechies-4 wavelet. In Figure 1.5 we have depicted the scalogram $|\mathcal{W}_\psi[f](a, b)|$ for the time period $b \in [0, 5]$ and the scales $a \in (0, 128]$. As in Figure 1.2, the dark grey values indicated high amplitudes of $|\mathcal{W}_\psi[f](a, b)|$ at that particular time-scale point. By comparing the plots of the spectrogram and the Wigner distribution of this function in Figure 1.2 and the scalogram in Figure 1.5 it is obvious, that scale and frequency are reciprocal concepts.

Formula (1.11) is mostly called the continuous wavelet transform (CWT), since there also exists a discrete version of this transform. This discrete wavelet transform (DWT) can be obtained by computing the CWT on a lattice

$$\{(2^m, n2^m) \mid n, m \in \mathbb{Z}\}.$$

Then the DWT of a function f is given by

$$f = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} (f, \psi_{m,n}) \psi_{m,n}, \quad (1.12)$$

with $\psi_{m,n}(x) = \psi(2^m x - n)$ and with $(f, \psi_{m,n})$ denoting the inner product in $L^2(\mathbb{R})$ of f and $\psi_{m,n}$. The breakthrough of the (discrete) wavelet transformation was set by two important contributions by Daubechies and Mallat.

In 1988 Daubechies introduced a method to construct wavelets ψ that are compactly supported and for which $\psi_{m,n}$, $m, n \in \mathbb{Z}$ form an orthonormal basis in $L^2(\mathbb{R})$. Since ψ is compactly supported, also $\psi_{m,n}$ is compactly supported for all $m, n \in \mathbb{Z}$. Consequently, a

compactly supported function f can be written as a finite sum of mutually orthonormal functions with compact support. As we shall see, this property is of great importance if we want to represent a signal or an image by a few coefficients.

An other important contribution for the DWT was given by Mallat in 1989. In his paper [60] he related the DWT to a concept called multiresolution analysis (MRA). This MRA provides an algorithm of complexity $\mathcal{O}(N)$, with N proportional to the number of inner products $(f, \psi_{m,n})$ that are not zero. This algorithm, called the pyramid algorithm, relates the inner products $(f, \psi_{m,n})$ for different values of m and n to each other. In this way only a few inner products have to be computed explicitly. The other ones follow from the computed inner products by means of the algorithm. Computing the DWT in this way is mostly referred to as the fast wavelet transform.

1.2 Applications of the Wavelet Transform

In the past decade the wavelet transform has become of great value not only for signal and image analysis, but also for signal and image processing. Even outside these fields the wavelet transform has shown to be useful in mathematics, e.g. in the field of numerical analysis. Here, we briefly discuss the advantages of using a wavelet transform for these applications.

1.2.1 Denoising

The idea behind denoising signals using a wavelet transform is as follows. Given a measured signal f , that consists of a signal g and some undesired noise signal h , linearity of the wavelet transform yields

$$\mathcal{W}_\psi[f](a, b) = \mathcal{W}_\psi[g](a, b) + \mathcal{W}_\psi[h](a, b).$$

Then, by choosing an appropriate orthogonal wavelet, we expect the wavelet coefficients $\mathcal{W}_\psi[g](a, b)$ to be concentrated within a small set of time-scale points (a, b) . Since the energy of g is then spread over a small set in the time-scale plane, the wavelet coefficients will attain relatively large values at these points. The wavelet coefficients of the noise, $\mathcal{W}_\psi[h](a, b)$, are expected to be spread out over the whole time-scale plane, since in general we cannot indicate dominating frequencies in a noise signal h . Consequently, the wavelet coefficients related to the noise signal h will attain relatively small values at all points in the time-scale plane.

Following these considerations, a natural way to come to a wavelet based denoising algorithm is given by thresholding the wavelet coefficients, i.e., $|\mathcal{W}_\psi[f](a, b)|$ is set to be zero, if $|\mathcal{W}_\psi[f](a, b)| < M$, with M the chosen threshold value. After this thresholding procedure the inverse wavelet transform reconstructs the signal g .

The problem of this method is how to find an appropriate threshold value M . In a series of papers Donoho and co-workers have dealt with this wavelet denoising method if h is a white noise signal, see e.g. [24]. For white noise signals they derived an expression for the threshold value M as a function of the standard deviation σ of the noise. In Chapter 7 we use the same technique for denoising seismic signals. For this application the threshold value is determined by computing the wavelet coefficients of noisy reference signal.

1.2.2 Compression

Using the wavelet transform for analysing signals, we are able to represent fluctuations in a signal with wavelet coefficients at a low scale. If a signal does not fluctuate during a time interval, we can represent the signal on this time interval by a small number coefficients at larger scales. Combining these observations yields that a signal can be represented in a sparse way. This is done by throwing away the wavelet coefficients at high scales if fluctuations in a signal can be observed in the corresponding time interval. Furthermore, the wavelet coefficients at low scales can be put to zero, if the signal only contains low frequencies in the corresponding time interval. This compression procedure is mostly applied to the DWT. We observe, that this compression technique is based on the adaptive character of the wavelet transform, that the WFT does not posses.

For images the same procedure can be applied. The wavelet transform is then replaced by a two-dimensional wavelet transform. For this transform the wavelet is replaced by a wavelet $\psi \in L^2(\mathbb{R}^2)$, the scaling parameter a is replaced by a matrix $A \in \mathbb{R}^{2 \times 2}$ and the translation over b is changed into translation over a two-dimensional lattice. Finally, fluctuations in a signal are translated into edges in an image.

Data compression is an important issue for multimedia applications, where one wants to store a huge set of images in a database or wants to send images to a receiver by the internet in a fast way. In the field of image compression, the JPEG standard does not use the wavelet transform yet, however the new JPEG-2000 standard will be fitted with a wavelet compression technique for data compression.

1.2.3 Wavelet-Galerkin methods

We consider a boundary value problem

$$(\mathcal{L}f)(x) = g(x), \quad x \in [x_1, x_2],$$

with $f(x_1) = 0$, $f(x_2) = 0$, $f \in C^2(\mathbb{R})$, $g \in C(\mathbb{R})$ and \mathcal{L} a second order linear differential operator with continuous coefficients. The differential equation can also be given in a weak formulation, i.e.,

$$(\mathcal{L}f, \phi) = (g, \phi),$$

with $f(x_1) = f(x_2) = 0$, for all $\phi \in C([x_0, x_1])$, that satisfy $\phi(x_0) = 0$, $\phi(x_1) = 0$ and ϕ' piecewise continuous and bounded on $[x_0, x_1]$. The space of such functions ϕ is denoted by V .

Galerkin's method consists of approximating f and ϕ by functions in a finite dimensional subspace V_0 of V , see e.g. [49]. By choosing a set of basis functions ϕ_k , $k = 1, \dots, \dim(V_0)$, the weak formulation can then be written as a linear system of equations

$$A\xi = b,$$

with

$$A(i, j) = (\phi_i, \mathcal{L}\phi_j), \quad b(i) = (g, \phi_i), \quad i, j = 1, \dots, \dim(V_0).$$

Let f_0 denote the approximation of f in V_0 , then ξ is related to f_0 by

$$f_0 = \sum_{i=1}^{\dim(V_0)} \xi(i) \phi_i.$$

Now, assume that $\mathcal{L} = -\frac{d^2}{dx^2}$, then

$$A(i, j) = (\mathcal{L}\phi_i, \phi_j) = (\phi'_i, \phi'_j).$$

The wavelet-Galerkin method uses a wavelet bases for V_0 . If such a wavelet bases is compactly supported, then the stiffness matrix A becomes a sparse matrix. We observe, that if ϕ is a wavelet, then ϕ' is also a wavelet. Mostly, semi-orthogonal wavelet bases with compact support are used to create a sparse stiffness matrix. We observe, that solving the linear system of equations can be established in a fast numerical way if A is sparse. Another way of looking at the wavelet-Galerkin method is to consider the wavelet approach as a pre-conditioning of the matrix A .

A comprehensive view of wavelet based methods in scientific computing is given by a series of papers by Beylkin et al., see e.g. [8].

1.3 Some Main Results

The following chapters can be divided into three parts, with each part consisting of two chapters. The first part consists of two chapters that introduce the Fourier transform, the windowed Fourier transform, the Wigner distribution and the wavelet transform in a rigorous mathematical way. Properties of these transformations and their mutual relations are extensively discussed. Furthermore, we study the relation of the introduced transformation to Lie groups. It is well known, that the wavelet transform is a unitary representation of the linear affine group, see e.g. [55, 62], and that the introduced time-frequency transformations are all

related to a unitary representation of the Heisenberg group, see e.g. [29].

In Chapters 4 and 5 we study two theoretical topics in the fields of the wavelet transform and the Wigner distribution respectively. In relation to the discrete wavelet transform we study the concept of an multiresolution analysis for arbitrary separable Hilbert spaces. By doing this we obtain a functional analytical framework in which the concept of an MRA can be studied. This framework relates the problem of finding wavelet bases in $L^2(\mathbb{R})$ to the problem of finding semi-orthogonal bases in $l^2(\mathbb{Z})$. By means of a discrete Fourier transform the latter problem is translated into the problem of finding invertible matrix functions on the n -fold unit circle. This theoretical approach of the concept of an MRA is illustrated by the example of constructing a spline wavelets, that generate a semi-orthogonal basis in $L^2(\mathbb{R})$. This topic can be found in Chapter 4 and is mainly based on

- [75] P.J. Oonincx and S.J.L. Van Eijndhoven, "Frames, Riesz systems and MRA in Hilbert spaces", *Indag. Math.*, 10 (3), 369-382, 1999.

Chapter 5 concerns affine transformations in the n -dimensional Wigner plane. This topic was inspired by the study of the fractional Fourier transform. This transformation was recently introduced for analysing signals in the Wigner plane, since it acts like a rotation in this plane. We show with relatively elementary results from Lie group theory, that the fractional Fourier transform on $L^2(\mathbb{R}^n)$ can be embedded in a group of unitary operators. Each element in this group corresponds to a symplectic transformation in the Wigner plane. Moreover, linear transformations in the Wigner plane that are related to unitary transformations on $L^2(\mathbb{R}^n)$ can only be symplectic.

We show that the FRFT is a special element of this group, since it is the only transformation that corresponds to an orthogonal symplectic transformation in the one-dimensional case. For the multi-dimensional case we also present other transformations that correspond to orthogonal symplectic transformations. Finally, a representation formula is given for all unitary transformations that correspond to symplectic transformations in the Wigner plane. This topic is mainly based on

- [63] H.G. ter Morsche and P.J. Oonincx, "The fractional Fourier transform and the integral representations of affine transformations in phase space", *to appear*.

The last two chapters of this thesis deal with applications of the studied transformations. The first application deals with energy localization problems and is based on a generalization of the FRFT. Two well-known problems are discussed rigorously, namely maximalization of the energy of time-limited signal within a compact frequency interval and maximalization of a signal's energy within a disc in the Wigner plane. Both problems are already extensively studied in the literature, see [57, 80, 92] and [20, 28, 29, 46]. In this chapter, we give a rigor-

ous proof of Slepian's conjecture, that shows how the eigenvalues of an energy localization operator behave asymptotically. Furthermore, an alternative proof is given for showing that Hermite functions are optimally localized on a disc in the Wigner plane. This proof is based on an argument, that uses the rotation property of the FRFT.

In the second part of Chapter 6 we show how the generalized FRFT can be used to solve a class of localization problems in the phase plane, if the solution of one problem in such a class is known. This procedure is illustrated by using it for the classical localization problems, that we discussed in the first part of this chapter. These problems and their solutions can be found in Chapter 6. This chapter is mainly based on

- [73] P.J. Oonincx, "On time-frequency analysis and time-limitedness", CWI-Report PNA-R9720, 1997.
- [?] P.J. Oonincx and H.G. ter Morsche, "The fractional Fourier transform and its application to energy localisation problems", *to appear*.

Chapter 7 is devoted to an application of the discrete wavelet transform in the field of seismology. Seismic data that are measured when an earthquake has taken place consist of several waves. These waves travel at different velocities from the epicenter of the earthquake towards the earth's surface, where they are measured. Consequently, they do not appear at the same time in a seismogram. Moreover, the several waves are overlapping and are also embedded in different kinds of noise. By detecting the arrival time of the so-called S-waves in a seismogram and relating this arrival time to the first time sample that the earthquake was recognized, an estimate of the distance towards the epicenter can be given. This is done by relating the difference in time between the two measured time samples to the difference in velocity between the S-wave and the P-wave, the first wave to arrive.

We developed an algorithm to detect automatically S-wave arrival times in seismograms. This algorithm is based both on physical properties of the waves and on separation of the waves by means of the discrete wavelet transform. Moreover, the DWT is used to reduce the noise in the waves. By doing this, a better estimate of the S-wave arrival time is established. This research has been done in strong collaboration with the Royal Dutch Meteorological Institute (KNMI). At KNMI the algorithm has been tested for a large set of seismic events. The results of this test are also included in Chapter 7. The mathematical concept of the algorithm has been published in

- [70] P.J. Oonincx, "A Wavelet Method for Detecting S-Waves in Seismic Data", *Computational Geosciences*, 3, 111-134, 1999.

Chapter 2

Time-Frequency Analysis

In this chapter we study the Fourier transform for functions in L^1 and L^2 . Using this transform on a signal one obtains information about its frequency behaviour. To obtain information about a signal's behaviour in both time and frequency we consider the windowed Fourier transform. This transform computes the Fourier transform of a signal within a sliding window. The last transform we discuss in this chapter is the Wigner distribution of a signal. This is a non-linear transform that also gives information about a signal's behaviour in both the time and frequency domain.

In the last section of this chapter we relate the discussed time-frequency representations to the Heisenberg group. Therefore an introduction in Lie group theory is presented. The group theoretical approach will play an important role in Chapter 5.

The contents of this chapter are based on existing literature on Fourier transforms [11, 35, 93, 108], on time-frequency transforms [18, 44] and Lie group theory [97, 101].

2.1 The Fourier Transform

To obtain information on the frequency behaviour of a function we can consider its Fourier transform. This transform computes the frequency spectrum of a given function. However, the Fourier transform neglects information about the function itself. We discuss the Fourier transform first for functions in $L^1(\mathbb{R})$ and next for functions in $L^2(\mathbb{R})$.

2.1.1 The Fourier Transform on L^1

In dealing with a function $f \in L^1(\mathbb{R}^n)$, one can consider the spectrum \hat{f} of f given by its Fourier transform

$$\hat{f}(\omega) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(x) e^{-i(\omega, x)} dx, \quad (2.1)$$

with (\cdot, \cdot) the inner product in \mathbb{R}^n .

In this section, we present some useful properties of the Fourier transform on $L^1(\mathbb{R})$, which can all be generalized straightforwardly to properties for functions in $L^1(\mathbb{R}^n)$. A fundamental property of Fourier transforms on $L^1(\mathbb{R})$ is given by the Riemann-Lebesgue theorem, see e.g. [11].

Theorem 2.1.1 (Riemann-Lebesgue) *Let $f \in L^1(\mathbb{R})$ and \hat{f} be its Fourier transform. Then*

$$|\hat{f}(\omega)| \rightarrow 0 \quad (|\omega| \rightarrow \infty).$$

Also the following fundamental result on \hat{f} can be proved for $f \in L^1(\mathbb{R})$.

Lemma 2.1.2 *Let $f \in L^1(\mathbb{R})$, then $\hat{f} \in C(\mathbb{R})$ and $\|\hat{f}\|_\infty \leq \|f\|_1 / \sqrt{2\pi}$.*

Proof

Let $f \in L^1(\mathbb{R})$. Then

$$\begin{aligned} |\hat{f}(\omega_1) - \hat{f}(\omega_2)| &= \frac{1}{\sqrt{2\pi}} \left| \int_{\mathbb{R}} f(x) (e^{-i\omega_1 x} - e^{-i\omega_2 x}) dx \right| \\ &\leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x)| \cdot |1 - e^{i(\omega_1 - \omega_2)x}| dx \\ &= \sqrt{\frac{2}{\pi}} \int_{\mathbb{R}} |f(x)| \cdot |\sin((\omega_1 - \omega_2)x/2)| dx. \end{aligned}$$

Applying the dominated convergence theorem on the latter result yields

$$|\hat{f}(\omega_1) - \hat{f}(\omega_2)| \rightarrow 0 \quad (\omega_1 \rightarrow \omega_2),$$

which shows that \hat{f} is continuous. Furthermore, we have

$$|\hat{f}(\omega)| \leq \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x) e^{-i\omega x}| dx = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} |f(x)| dx = \|f\|_1 / \sqrt{2\pi}.$$

Taking the supremum over ω establishes the proof. □

Combining Theorem 2.1.1 and Lemma 2.1.2, we arrive at the following corollary.

Corollary 2.1.3 Let $f \in L^1(\mathbb{R})$, then $\hat{f} \in C_0(\mathbb{R})$, the supremum-normed Banach space of all complex continuous functions on \mathbb{R} that vanish at infinity. Furthermore,

$$\|\hat{f}\|_\infty \leq \|f\|_1 / \sqrt{2\pi}.$$

A useful property of the Fourier transform is its action on the convolution product $f * g$ of two function f and g in $L^1(\mathbb{R})$, given by

$$(f * g)(x) = \int_{\mathbb{R}} f(x - u)g(u) du. \quad (2.2)$$

By Young's inequality [108], we have $\|f * g\|_1 \leq \|f\|_1 \cdot \|g\|_1$. Since $f * g \in L^1(\mathbb{R})$, we can compute its Fourier transform

$$\begin{aligned} (f * g)^\sim(\omega) &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} f(x - u)g(u) du e^{-i\omega x} dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \int_{\mathbb{R}} f(x - u)e^{-i\omega(x-u)} g(u)e^{-i\omega u} du dx \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x)e^{-i\omega x} dx \int_{\mathbb{R}} g(u)e^{-i\omega u} du \\ &= \sqrt{2\pi} \hat{f}(\omega) \hat{g}(\omega), \end{aligned} \quad (2.3)$$

using Fubini's theorem. We observe that the iterated integral in the second last line is absolutely convergent, which justifies the change of integration.

At the end of this subsection we study the inverse Fourier transform on $L^1(\mathbb{R})$. Formally, an inverse Fourier transform exists and is given by

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\omega) e^{ix\omega} d\omega. \quad (2.4)$$

The following example shows that \hat{f} is not necessarily in $L^1(\mathbb{R})$ if $f \in L^1(\mathbb{R})$.

Example 2.1.4 Let $f \in L^1(\mathbb{R})$ be given by

$$f(x) = \begin{cases} \sqrt{2\pi} e^{-x}, & x \geq 0, \\ 0, & x < 0. \end{cases}$$

Then its Fourier transform is given by

$$\hat{f}(\omega) = \frac{1}{1 + i\omega},$$

which is not in $L^1(\mathbb{R})$. Note that $\hat{f} \in C_0(\mathbb{R})$, cf. Corollary 2.1.3.

Following [11, 35], we present additional conditions on f and \hat{f} , that are necessary for a well-defined inversion formula.

Theorem 2.1.5 *Let $f \in L^1(\mathbb{R})$ and $\hat{f} \in L^1(\mathbb{R})$. Then*

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \hat{f}(\omega) e^{ix\omega} d\omega \quad \text{a.e. } x \in \mathbb{R}.$$

Moreover, the latter results holds for every $x \in \mathbb{R}$ if also $f \in C(\mathbb{R})$.

2.1.2 The Fourier Transform on L^2

For $f \in L^2(\mathbb{R})$ we want to define its Fourier transform also by (2.1). However, this can only be done if $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, since the Fourier transform of $f \in L^2(\mathbb{R})$ may not be defined everywhere. To come to a definition of the Fourier transform on $L^2(\mathbb{R})$ we will define the Fourier transform first on a dense subspace of both $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$ and then extend it uniquely to $L^2(\mathbb{R})$.

A dense subspace of both $L^1(\mathbb{R})$ and $L^2(\mathbb{R})$ is given by the Schwartz class $S(\mathbb{R})$, see [89, 90].

Definition 2.1.6 *The Schwartz class $S(\mathbb{R}^n)$ is the space of rapidly decreasing C^∞ -functions on \mathbb{R}^n , i.e., for each $k, l \in \mathbb{N}$*

$$\sup_{|\alpha| \leq k, |\beta| \leq l, x \in \mathbb{R}^n} |x_1^{\beta_1} \cdots x_n^{\beta_n} \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} f(x)| < \infty \quad \forall f \in S(\mathbb{R}^n).$$

It can be shown that the Fourier transform \mathcal{F} , given by $\mathcal{F}[f] = \hat{f}$ for $f \in S(\mathbb{R})$, is a bounded linear mapping on $S(\mathbb{R})$ as a subspace of $L^2(\mathbb{R})$. Moreover, \mathcal{F} is an isometry on $S(\mathbb{R})$, with respect to the inner product in $L^2(\mathbb{R})$, see [35, 94]. So, we have Parseval's formula

$$(f, g) = (\mathcal{F}f, \mathcal{F}g),$$

with $(\cdot, \cdot)_2$ the inner product in $L^2(\mathbb{R})$.

Since $S(\mathbb{R})$ is dense in $L^2(\mathbb{R})$, \mathcal{F} can be uniquely extended to a Hilbert space isometry of $L^2(\mathbb{R})$, which yields the definition of the Fourier transform on $L^2(\mathbb{R})$.

Definition 2.1.7 *Let $f \in L^2(\mathbb{R})$. Then its Fourier transform $\hat{f} = \mathcal{F}f$ is given by*

$$\mathcal{F}[f](\omega) = \text{l.i.m.}_{N \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-N}^N f(x) e^{-i\omega x} dx, \quad (2.5)$$

where l.i.m. stands for limit in L^2 mean.

Remark, that this definition coincides with (2.1) if $f \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$. Also we observe that by this definition $\mathcal{F}f$ is a function, defined almost everywhere on \mathbf{R} and belonging to $L^2(\mathbf{R})$. Moreover, with this construction Parseval's formula can be extended to $L^2(\mathbf{R})$

$$\int_{\mathbf{R}} f(x) \overline{g(x)} dx = \int_{\mathbf{R}} \hat{f}(\omega) \overline{\hat{g}(\omega)} d\omega, \quad (2.6)$$

for all $f, g \in L^2(\mathbf{R})$. As a result we also have Plancherel's formula

$$\int_{\mathbf{R}} |f(x)|^2 dx = \int_{\mathbf{R}} |\hat{f}(\omega)|^2 d\omega, \quad (2.7)$$

for all $f \in L^2(\mathbf{R})$. The two equal sides of (2.7) give the energy of $f \in L^2(\mathbf{R})$.

Property (2.3) on convolution products of functions in L^1 can also be proved for $f \in L^2(\mathbf{R})$ and $g \in L^1(\mathbf{R})$. Since $\|f * g\|_2 \leq \|f\|_2 \|g\|_1$ by Young's inequality we can follow (2.3). This yields straightforwardly

$$(f * g)^\sim(\omega) = \sqrt{2\pi} \hat{f}(\omega) \hat{g}(\omega) \text{ a.e. } \forall f \in L^2(\mathbf{R}) \forall g \in L^1(\mathbf{R}).$$

A similar result on convolution products in the Fourier domain can be derived by taking

$$g(x) = \overline{h(x)} e^{i\xi x}$$

in (2.6), with $h \in L^2(\mathbf{R})$. This yields

$$\begin{aligned} \sqrt{2\pi} \mathcal{F}[f \cdot h](\xi) &= \int_{\mathbf{R}} f(x) h(x) e^{-i\xi x} dx = \int_{\mathbf{R}} f(x) \overline{g(x)} dx \\ &= \int_{\mathbf{R}} \hat{f}(\omega) \overline{\hat{g}(\omega)} d\omega \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} \hat{f}(\omega) \left(\int_{\mathbf{R}} h(x) e^{-i(\xi - \omega)x} dx \right) d\omega \\ &= \int_{\mathbf{R}} \hat{f}(\omega) \hat{h}(\xi - \omega) d\omega = (\hat{f} * \hat{h})(\xi). \end{aligned} \quad (2.8)$$

We summarize these results in the following lemma.

Lemma 2.1.8 *Convolution products and the Fourier transform are related by*

1. $(f * g)^\sim(\omega) = \sqrt{2\pi} \hat{f}(\omega) \cdot \hat{g}(\omega)$, for $f \in L^1(\mathbf{R}) \cup L^2(\mathbf{R})$ and $g \in L^1(\mathbf{R})$,

$$2. \sqrt{2\pi} (f \cdot g)(\omega) = (\hat{f} * \hat{g})(\omega), \text{ for } f, g \in L^2(\mathbb{R}).$$

Since $\hat{f} \in L^2(\mathbb{R})$ for $f \in \mathcal{R}$, we can derive an inversion formula using the same construction as for (2.5), i.e.,

$$f(x) = \text{l.i.m.}_{N \rightarrow \infty} \frac{1}{\sqrt{2\pi}} \int_{-N}^N \hat{f}(\omega) e^{i\omega x} d\omega. \quad (2.9)$$

A subspace of $L^2(\mathbb{R})$, which is of special interest in signal analysis is $L^2_{\text{comp}}(\mathbb{R})$, i.e., the space of all functions in $L^2(\mathbb{R})$ with compact support. Related to this space we can define two types of signals.

Definition 2.1.9 A signal $f \in L^2(\mathbb{R})$ is called time-limited if $f \in L^2_{\text{comp}}(\mathbb{R})$.

If $\hat{f} \in L^2_{\text{comp}}(\mathbb{R})$, then f is called band-limited.

Another special class of functions in $L^2(\mathbb{R})$ is the class of functions of exponential type.

Definition 2.1.10 A function $f \in L^2(\mathbb{R})$ is called of exponential type if it extends to a holomorphic function on \mathbb{C} and if there are two positive constants C and Ω such that

$$|f(z)| < C e^{\Omega |\text{Im } z|}, \quad \forall z \in \mathbb{C}.$$

Functions of exponential type can be related to band-limited functions by means of the following lemma.

Lemma 2.1.11 If $f \in L^2(\mathbb{R})$ is band-limited, then f is of exponential type.

Proof

Assume $\hat{f}(\omega) = 0$ for $|\omega| > \Omega$, with $\Omega > 0$. Then

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{i\omega x} d\omega,$$

initially defined for $x \in \mathbb{R}$, remains well-defined for $x \in \mathbb{C}$, and yields a holomorphic function f on \mathbb{C} . Furthermore, we derive

$$\begin{aligned} |f(z)| &= \left| \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} \hat{f}(\omega) e^{i\omega z} d\omega \right| \\ &\leq \frac{1}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} |\hat{f}(\omega) e^{i\omega z}| d\omega \leq \frac{e^{\Omega |\text{Im } z|}}{\sqrt{2\pi}} \int_{-\Omega}^{\Omega} |\hat{f}(\omega)| d\omega \\ &\leq \frac{e^{\Omega |\text{Im } z|}}{\sqrt{2\pi}} \sqrt{\int_{\mathbb{R}} |\hat{f}(\omega)|^2 d\omega} \cdot \sqrt{\int_{-\Omega}^{\Omega} d\omega} = \sqrt{\frac{\Omega \|f\|_2^2}{\pi}} e^{\Omega |\text{Im } z|}, \quad \forall z \in \mathbb{C}. \square \end{aligned}$$

Lemma 2.1.11 can be extended to the Paley-Wiener theorem, a well-known result in Fourier theory; for a proof, see [107].

Theorem 2.1.12 (Paley-Wiener) *If $f \in L^2(\mathbb{R})$ is holomorphic and of exponential type, then f is band-limited. Conversely, if f is band-limited, then f is holomorphic and of exponential type.*

Since a holomorphic function $f \in L^2(\mathbb{R})$, vanishing at a certain interval, has to be identically zero, the Paley-Wiener theorem immediately yields

Corollary 2.1.13 *If $f \in L^2(\mathbb{R})$ is both time-limited and band-limited, then $f = 0$.*

The previous corollary states that there does not exist a non-trivial band-limited signal f , whose energy is contained within a finite interval in the frequency domain, say $[-\Omega, \Omega]$. In Section 6.1 we will deal with this phenomenon. There, we will consider the problem of maximizing the energy of a band-limited signal within a finite interval $[-\Omega, \Omega]$ in the frequency domain.

It is clear, that, when using the Fourier transform to analyse a signal, we hide all information of the signal's behaviour in the time domain. In order to obtain information about a signal simultaneously in the frequency domain and the time domain, we may replace the Fourier transform by one of the integral transforms that we discuss in the following sections.

2.2 The Windowed Fourier Transform

A strategy to obtain information about the frequency contents of a signal $f \in L^2(\mathbb{R})$ locally in time is to first multiply f with a window function $\bar{h} \in L^2(\mathbb{R})$ and then take its Fourier transform, i.e.,

$$\mathcal{F}_h[f](x, \omega) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(y) \overline{h(y-x)} e^{-i\omega y} dy. \quad (2.10)$$

This formula is called the windowed Fourier transform (WFT) or also sometimes in literature [18] the short-time Fourier transform (STFT).

The multi-dimensional WFT is defined as a straightforward generalization of (2.10) by

$$\mathcal{F}_h[f](x, \omega) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(y) \overline{h(y-x)} e^{-i(\omega, y)} dy. \quad (2.11)$$

Properties and results that we derive in this section for the one-dimensional WFT can all be generalized in a direct way for functions in $L^2(\mathbb{R}^n)$.

By introducing the shift operator \mathcal{T}_b on $L^2(\mathbb{R})$ by

$$\mathcal{T}_b[f](x) = f(x - b), \quad (2.12)$$

for some $b \in \mathbb{R}$ and the frequency shift operator operator \mathcal{M}_ω on $L^2(\mathbb{R})$ by

$$\mathcal{M}_\omega[f](x) = e^{i\omega x} f(x) \quad (2.13)$$

for some $\omega \in \mathbb{R}$, we can write (2.10) also as

$$\mathcal{F}_h[f](x, \omega) = (f, \mathcal{M}_\omega \mathcal{T}_x h)_2. \quad (2.14)$$

This notation immediately yields by Schwarz's inequality

$$|\mathcal{F}_h[f](x, \omega)| \leq \|f\|_2 \cdot \|h\|_2 \quad \forall x, \omega \in \mathbb{R}.$$

So $\mathcal{F}_h f \in L^\infty(\mathbb{R}^2)$, for all $f, h \in L^2(\mathbb{R})$.

A desirable property of the WFT is its invariance under the action of a frequency modulation \mathcal{M}_ω and its invariance up to a phase factor under the action of a translation \mathcal{T}_b . Indeed, a straightforward calculation shows

$$\mathcal{F}_h[\mathcal{T}_b f](x, \omega) = e^{-i\omega b} \mathcal{F}_h[f](x - b, \omega) \quad \text{and} \quad \mathcal{F}_h[\mathcal{M}_{\omega_0} f](x, \omega) = \mathcal{F}_h[f](x, \omega - \omega_0).$$

We can also write (2.10) by means of a Fourier transform on $L^1(\mathbb{R})$, namely

$$\mathcal{F}_h[f](x, \omega) = \mathcal{F}[f \cdot \overline{\mathcal{T}_x h}](\omega), \quad (2.15)$$

which can be rewritten as

$$\mathcal{F}_h[f](x, \omega) = (\mathcal{F}[f] * \mathcal{F}[\overline{\mathcal{T}_x h}])(\omega) / \sqrt{2\pi}, \quad (2.16)$$

using Lemma 2.1.8. Using Lemma 2.1.2 on (2.15) immediately yields that $\mathcal{F}_h f$ is continuous in ω for fixed $x \in \mathbb{R}$. Moreover, we have

$$|\mathcal{F}_h[f](x, \omega) - \mathcal{F}_h[f](y, \omega)| = |\mathcal{F}[f \cdot (\overline{\mathcal{T}_x h} - \overline{\mathcal{T}_y h})](\omega)| \leq \|f\|_2 \|\mathcal{T}_x h - \mathcal{T}_y h\|_2 / \sqrt{2\pi},$$

using that the mapping $x \mapsto \mathcal{T}_x h$ is continuous for all $h \in L^2(\mathbb{R})$. This inequality shows that $\mathcal{F}_h f$ is also continuous in x for uniform $\omega \in \mathbb{R}$. Combining this with the fact that $\mathcal{F}_h f$ is continuous in ω for fixed $x \in \mathbb{R}$, we have $\mathcal{F}_h[f] \in C(\mathbb{R}^2)$ for all $f, h \in \mathbb{R}^2$.

In the following example we consider a widely used window function h to compute the WFT of a given function.

Example 2.2.1 We take for h the Gaussian function

$$h_\sigma(x) = (\pi\sigma^2)^{-1/4} e^{-x^2/2\sigma^2},$$

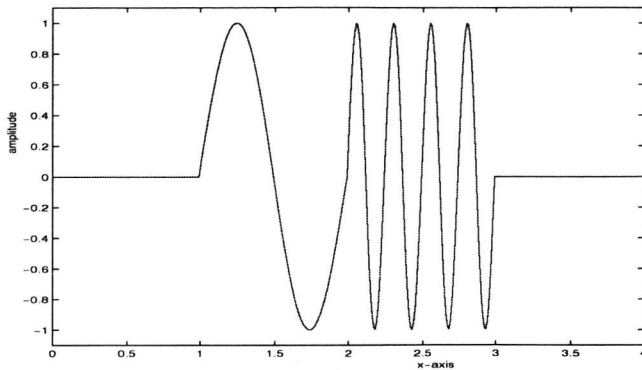


Figure 2.1: A composition of two different oscillations with finite duration.

where $\sigma > 0$. Note that $\|h_\sigma\|_2 = 1$ for all $\sigma > 0$. The corresponding WFT is given by the Gabor transform [34]

$$\mathcal{G}_\sigma[f](x, \omega) = (2\pi)^{-3/4} \sigma^{-1/2} \int_{\mathbb{R}} f(y) e^{-(y-x)^2/2\sigma^2} e^{-i\omega y} dy. \quad (2.17)$$

As an example we use the Gabor transform to obtain both time and frequency information about the function

$$f(x) = \begin{cases} \sin(2\pi x), & x \in [1, 2], \\ \sin(8\pi x), & x \in [2, 3], \\ 0, & \text{otherwise,} \end{cases} \quad (2.18)$$

which is depicted in Figure 2.1.

For the parameter σ three different values have been taken, namely $\sigma^2 = 1/8$, $\sigma^2 = 1/32$ and $\sigma^2 = 1/200$. Figure 2.2 shows the amplitude of $|\mathcal{G}_\sigma[f](x, \omega)|$ for these three choices for σ . We see that in Figure 2.2.a the two frequency components of f are clearly visible and well localized in time. Here maxima of $|\mathcal{G}_{1/2\sqrt{2}}|$ can be observed around $\omega = \pm 2\pi$ for $x \in [1, 2]$ and around $\omega = \pm 8\pi$ for $x \in [2, 3]$.

When decreasing the value of σ to $1/4\sqrt{2}$ only maxima are attained around $\omega = 0$ for $x = 1.25$ and $x = 1.75$, which is depicted in Figure 2.2.b. This means that only information about the function's behavior in time is visible during the interval $[1, 2]$. For $x \in [2, 3]$ the situation remains the same. This phenomenon is explained by the fact that the 'width' of the Gaussian function becomes too small to recognize low frequency behavior.

By decreasing the value of σ even more also higher frequency components are neglected and only maximum amplitudes are measured by the Gabor transform. This situation is visualized in Figure 2.2.c where $\sigma = 1/10\sqrt{2}$ has been chosen.

Parseval's formula (2.6) and notation (2.15) are used to prove a counterpart of Parseval's formula for the WFT.

Theorem 2.2.2 *Let $f, g \in L^2(\mathbf{R})$ and $h \in L^2(\mathbf{R})$, with $h \neq 0$ on a set in \mathbf{R} with positive measure. Then $\mathcal{F}_h[f], \mathcal{F}_h[g] \in L^2(\mathbf{R}^2)$ and*

$$\int_{\mathbf{R}} f(x) \overline{g(x)} dx = \frac{1}{\|h\|_2^2} \int_{\mathbf{R}^2} \mathcal{F}_h[f](x, \omega) \overline{\mathcal{F}_h[g](x, \omega)} d\omega dx. \quad (2.19)$$

Proof

Using notation (2.15) in the right-hand side of (2.19) yields

$$\frac{1}{\|h\|_2^2} \int_{\mathbf{R}^2} \mathcal{F}_h[f](x, \omega) \overline{\mathcal{F}_h[g](x, \omega)} d\omega dx = \frac{1}{\|h\|_2^2} \int_{\mathbf{R}^2} \mathcal{F}[f \cdot \overline{\mathcal{T}_x h}](\omega) \overline{\mathcal{F}[g \cdot \overline{\mathcal{T}_x h}](\omega)} d\omega dx.$$

Now, Parseval's formula (2.6) gives

$$\begin{aligned} & \frac{1}{\|h\|_2^2} \int_{\mathbf{R}^2} \mathcal{F}[f \cdot \overline{\mathcal{T}_x h}](\omega) \overline{\mathcal{F}[g \cdot \overline{\mathcal{T}_x h}](\omega)} d\omega dx \\ &= \frac{1}{\|h\|_2^2} \int_{\mathbf{R}^2} f(y) \overline{g(y)} |h(y-x)|^2 dy dx \\ &= \frac{1}{\|h\|_2^2} \int_{\mathbf{R}} f(y) \overline{g(y)} \int_{\mathbf{R}} |h(y-x)|^2 dx dy = \int_{\mathbf{R}} f(y) \overline{g(y)} dy, \end{aligned}$$

using Fubini's theorem. Since $f \cdot \overline{g} \in L^1(\mathbf{R})$ the last iterated integral is absolutely convergent, which justifies the change of integration. \square

As a result we have Plancherel's formula for the WFT

$$\int_{\mathbf{R}} |f(x)|^2 dx = \frac{1}{\|h\|_2^2} \int_{\mathbf{R}^2} |\mathcal{F}_h[f](x, \omega)|^2 d\omega dx, \quad (2.20)$$

for $f, h \in L^2(\mathbf{R})$, with $h \neq 0$ on a set in \mathbf{R} with positive measure. This shows that the operator $\mathcal{F}_h : L^2(\mathbf{R}) \rightarrow L^2(\mathbf{R}^2)$ is an isometry up to a constant $\|h\|_2$. The term $|\mathcal{F}_h[f](x, \omega)|^2$ in (2.20) is usually called spectrogram [18]. Referring to (2.20), the spectrogram is a measure for the energy density in the time-frequency plane. Note that the spectrogram of a given function $f \in L^2(\mathbf{R})$ strongly depends on the window function $h \in L^2(\mathbf{R})$.

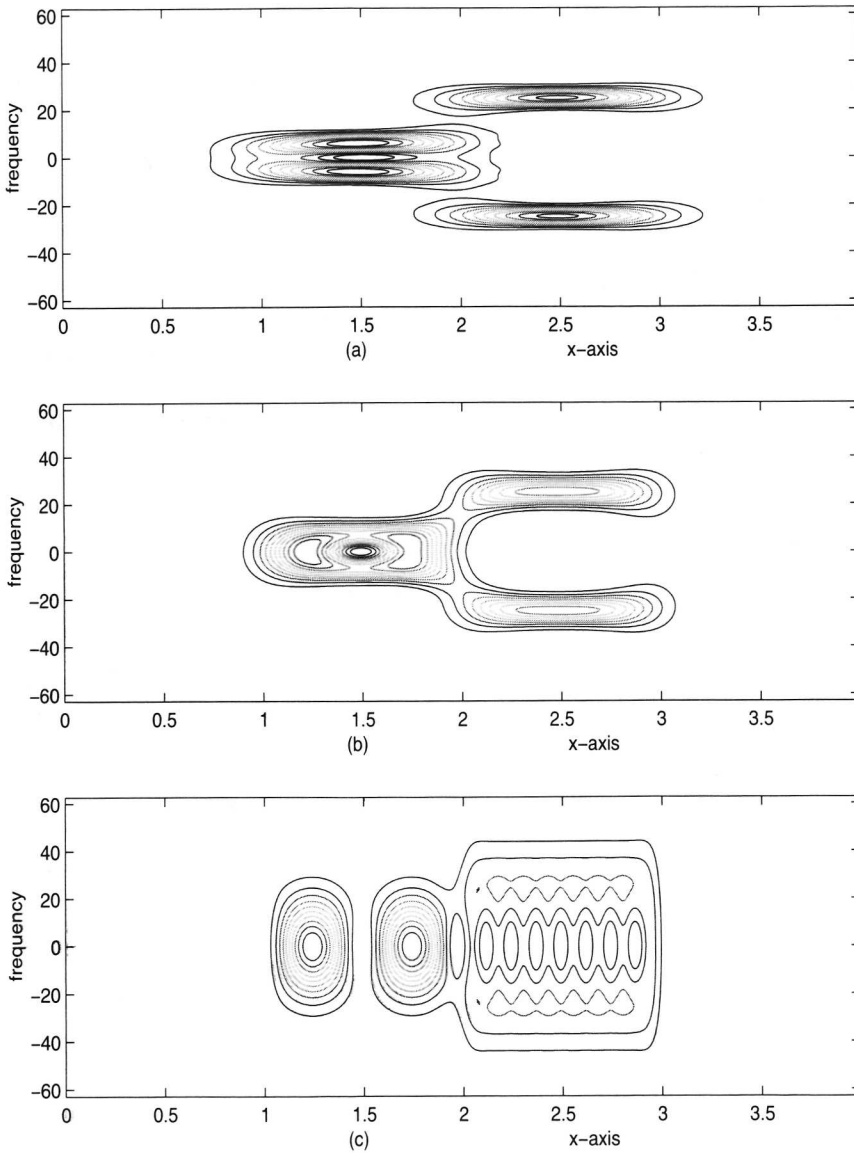


Figure 2.2: A contour plot of the amplitude of a Gabor transform with a) $\sigma^2 = 1/8$, b) $\sigma^2 = 1/32$ and c) $\sigma^2 = 1/200$.

From Parseval's formula (2.19) we derive

$$\begin{aligned}
 & \int_{\mathbb{R}} f(x) \overline{g(x)} dx \\
 &= \frac{1}{\sqrt{2\pi} \|h\|_2^2} \int_{\mathbb{R}^2} \mathcal{F}_h[f](y, \omega) \int_{\mathbb{R}} \overline{g(x)} h(x-y) e^{i\omega x} dx d\omega dy \\
 &= \int_{\mathbb{R}} \left(\frac{1}{\sqrt{2\pi} \|h\|_2^2} \int_{\mathbb{R}^2} \mathcal{F}_h[f](y, \omega) h(x-y) e^{i\omega x} d\omega dy \right) \overline{g(x)} dx, \quad (2.21)
 \end{aligned}$$

which yields the following formal inversion formula.

Theorem 2.2.3 *Let $f, h \in L^2(\mathbb{R})$, with $h \neq 0$ on a set in \mathbb{R} with positive measure.*

$$f(x) = \frac{1}{\sqrt{2\pi} \|h\|_2^2} \int_{\mathbb{R}^2} \mathcal{F}_h[f](y, \omega) h(x-y) e^{i\omega x} d\omega dy, \text{ weakly in } L^2(\mathbb{R}). \quad (2.22)$$

Relation (2.22) should be interpreted by means of (2.21). A stronger result holds if additional conditions on f and h apply.

Theorem 2.2.4 *Let $f, h \in L^2(\mathbb{R})$, with $h \neq 0$ on a set in \mathbb{R} with positive measure. Furthermore, let $\mathcal{F}[f \cdot h] \in L^1(\mathbb{R})$. Then*

$$f(x) = \frac{1}{\sqrt{2\pi} \|h\|_2^2} \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathcal{F}_h[f](y, \omega) h(x-y) e^{i\omega x} d\omega \right) dy, \text{ pointwise.} \quad (2.23)$$

For each $x \in \mathbb{R}$, both the inner and the outer integral are absolutely convergent, but possibly not the double integral.

Proof

Fix $x \in \mathbb{R}$. Then from (2.15) we have

$$\mathcal{F}_h[f](y, \omega) = \mathcal{F}[f \cdot \overline{\tau_y h}](\omega).$$

Since $f \cdot h \in L^1(\mathbb{R})$ and $\mathcal{F}[f \cdot h] \in L^1(\mathbb{R})$ yields $\mathcal{F}[f \cdot \overline{\tau_y h}] \in L^1(\mathbb{R})$, we can take the inverse Fourier transform as given in Theorem 2.1.5. This gives us

$$f(x) \overline{h(x-y)} = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} \mathcal{F}_h[f](y, \omega) e^{i\omega x} d\omega.$$

Multiplying both sides of this equation by $h(x-y)$ followed by integration over y establishes the proof. \square

We observe that a sufficient condition on f and h such that (2.23) holds is given by $\hat{f}, \hat{h} \in L^1(\mathbb{R})$. This follows from Lemma 2.1.8 and Young's inequality.

We have shown that, for a given $h \in L^2(\mathbb{R})$, \mathcal{F}_h maps all $f \in L^2(\mathbb{R})$ to $\mathcal{F}_h f \in L^2(\mathbb{R}^2)$. Moreover, we have proven that f can be recovered from $\mathcal{F}_h f$, if $h \neq 0$ on a set with positive measure. However not every element in $L^2(\mathbb{R}^2)$ is the image of an $f \in L^2(\mathbb{R})$ by means of the WFT, since the range of \mathcal{F}_h , $\text{Ran}(\mathcal{F}_h)$ is only a subset of $L^2(\mathbb{R}^2)$, see [51]. To characterize $\text{Ran}(\mathcal{F}_h)$ we derive from (2.14) and (2.19)

$$\begin{aligned} \mathcal{F}_h[f](x, \omega) &= (f, \mathcal{M}_\omega \mathcal{T}_x h)_2 \\ &= \frac{1}{\|h\|_2^2} \int_{\mathbb{R}^2} \mathcal{F}_h[f](u, v) \overline{\mathcal{F}_h[\mathcal{M}_\omega \mathcal{T}_x h](u, v)} du dv \\ &= \int_{\mathbb{R}^2} k_h(x, \omega; u, v) \mathcal{F}_h[f](u, v) du dv, \end{aligned}$$

with

$$\begin{aligned} k_h(x, \omega; u, v) &= \overline{\mathcal{F}_h[\mathcal{M}_\omega \mathcal{T}_x h](u, v)} = \mathcal{F}[(\mathcal{T}_u h) \cdot \overline{(\mathcal{T}_x h)}](\omega - v) \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} h(y - u) \overline{h(y - x)} e^{-i(\omega - v)y} dy, \end{aligned}$$

the reproducing kernel. Note that by definition $k_h(x, \omega; \cdot, \cdot) \in L^2(\mathbb{R}^2)$ for all $x, \omega \in \mathbb{R}$. From this derivation it follows that $\theta \in L^2(\mathbb{R}^2)$ should necessarily satisfy

$$\theta(x, \omega) = \int_{\mathbb{R}^2} k_h(x, \omega; u, v) \theta(u, v) dv du, \quad (2.24)$$

in order to be an element of $\text{Ran}(\mathcal{F}_h)$. Note that (2.24) is well defined, since $\mathcal{F}_h f$ is continuous for all $f, h \in L^2(\mathbb{R})$.

We can also show that (2.24) is a sufficient condition on θ to be in $\text{Ran}(\mathcal{F}_h)$. This is shown by constructing an $f \in L^2(\mathbb{R})$ such that $\mathcal{F}_h f = \theta$, for some $h \in L^2(\mathbb{R})$ and with θ satisfying (2.24). Before constructing such a function f , we observe that $\theta \in L^2(\mathbb{R}^2)$. As a result of Fubini's theorem, we thus have $\theta(\cdot, v) \in L^2(\mathbb{R})$ for fixed $v \in \mathbb{R}$ and $\theta(u, \cdot) \in L^2(\mathbb{R})$ for fixed $u \in \mathbb{R}$.

We take

$$f(y) = \sqrt{2\pi} \int_{\mathbb{R}^2} \hat{\theta}(\xi, -y) \hat{h}(\xi) e^{i\xi y} d\xi, \quad \text{weakly in } L^2(\mathbb{R}).$$

This means that

$$\int_{\mathbb{R}} f(y) \overline{g(y)} dy = \int_{\mathbb{R}^2} \hat{\theta}(\xi, -y) \hat{h}(\xi) \overline{g(y)} e^{i\xi y} d\xi dy,$$

for all $g \in L^2(\mathbb{R})$. Consequently, we have

$$\left| \int_{\mathbb{R}} f(y) \overline{g(y)} dy \right| \leq \sqrt{2\pi} \|\theta\|_2 \|h\|_2 \|g\|_2,$$

for all $g \in L^2(\mathbb{R})$. To show that $\mathcal{F}_h f = \theta$, we derive for $h \in S(\mathbb{R})$ and $\theta \in S(\mathbb{R}^2)$

$$\begin{aligned} & \int_{\mathbb{R}^2} k_h(x, \omega; u, v) \theta(u, v) du dv \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}^2} \theta(u, v) \left(\int_{\mathbb{R}} h(y-v) \overline{h(y-x)} e^{-iv(\omega-v)} dy \right) du dv \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \left(\int_{\mathbb{R}^2} \theta(u, v) e^{ivv} \left(\int_{\mathbb{R}} \hat{h}(\xi) e^{i(v-u)\xi} d\xi \right) du dv \right) \overline{h(y-x)} e^{-iv\omega} dy \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\int_{\mathbb{R}^2} \theta(u, v) e^{-iu\xi} e^{ivv} du dv \right) \hat{h}(\xi) \overline{h(y-x)} e^{iv\xi} e^{-iv\omega} dy d\xi \\ &= \int_{\mathbb{R}^2} \hat{\theta}(\omega, -y) \hat{h}(\xi) \overline{h(y-x)} e^{i\xi y} e^{-iv\omega} dy d\xi \end{aligned}$$

We observe, that

$$\begin{aligned} \left| \int_{\mathbb{R}^2} k_h(x, \omega; u, v) \theta(u, v) du dv \right| &\leq \|\theta\|_2 \|k_h(x, \omega; \cdot, \cdot)\|_2 \\ &= \|\theta\|_2 \|\mathcal{F}_h \mathcal{M}_\omega \mathcal{T}_x h\|_2 = \|\theta\|_2 \|h\|_2^2 \end{aligned}$$

and

$$\left| \int_{\mathbb{R}^2} \int_{\mathbb{R}^2} \hat{\theta}(\omega, -y) \hat{h}(\omega) \overline{h(y-x)} e^{i\omega y} e^{-iv\omega} dy d\omega \right| \leq \|\theta\|_2 \|h\|_2^2.$$

We conclude, that the previous derivation also holds for $\theta \in L^2(\mathbb{R}^2)$ and $h \in L^2(\mathbb{R})$. From (2.24) and the definition of f it follows directly, that

$$\mathcal{F}_h[f](x, \omega) = \theta(x, \omega).$$

We summarize the previous result in the following theorem.

Theorem 2.2.5 For $h \in L^2(\mathbb{R})$, the range of \mathcal{F}_h , $\text{Ran}(\mathcal{F}_h)$, is given by

$$\text{Ran}(\mathcal{F}_h) = \{\theta \in L^2(\mathbb{R}^2) \mid \theta(x, \omega) = \int_{\mathbb{R}^2} k_h(x, \omega; u, v) \theta(u, v) dv du\}.$$

From Plancherel's formula (2.20) it follows that $\text{Ran}(\mathcal{F}_h)$ is closed and that the transform $f \mapsto \mathcal{F}_h f / \|h\|_2$ is a Hilbert space isometry from $L^2(\mathbb{R})$ onto $\text{Ran}(\mathcal{F}_h)$ as given in Theorem 2.2.5.

2.3 The Wigner Distribution

To obtain both time and frequency information of a signal $f \in L^2(\mathbb{R})$ we may also use the Wigner distribution. This is a bilinear time-frequency representation given by

$$\mathcal{WV}[f](x, \omega) = \frac{1}{2\pi} \int_{\mathbb{R}} f(x + t/2) \overline{f(x - t/2)} e^{-it\omega} dt, \quad (2.25)$$

for all $f \in L^2(\mathbb{R})$. In the sequel we will refer to the domain of the Wigner distribution as the Wigner plane.

In the sequel we will also use the mixed Wigner distribution given by

$$\mathcal{WV}[f, g](x, \omega) = \frac{1}{2\pi} \int_{\mathbb{R}} f(x + t/2) \overline{g(x - t/2)} e^{-it\omega} dt, \quad (2.26)$$

for all $f, g \in L^2(\mathbb{R})$. Obviously, this representation coincides with the Wigner distribution if $f = g$.

The mixed Wigner distribution can also be written as a windowed Fourier transform. We derive

$$\begin{aligned} \mathcal{WV}[f, g](x, \omega) &= \frac{1}{\pi} \int_{\mathbb{R}} f(x + t) \overline{g(x - t)} e^{-2it\omega} dt, \\ &= \frac{e^{2ix\omega}}{\pi} \int_{\mathbb{R}} f(y) \overline{g(2x - y)} e^{-2iy\omega} dy, \\ &= \sqrt{\frac{2}{\pi}} e^{2ix\omega} \mathcal{F}_{g(\dots)}[f](2x, 2\omega). \end{aligned} \quad (2.27)$$

Particularly, we have

$$\mathcal{WV}[f](x, \omega) = \sqrt{\frac{2}{\pi}} e^{2iz\omega} \mathcal{F}_{f(-\cdot)}[f](2x, 2\omega). \quad (2.28)$$

The relation between the WFT and the Wigner distribution will be used to derive elementary properties of the Wigner distribution, that also hold for the WFT.

The multi-dimensional Wigner distribution is defined from a straightforward generalization of (2.25) by

$$\mathcal{WV}[f](x, \omega) = (2\pi)^{-n} \int_{\mathbb{R}^n} f(x + t/2) \overline{f(x - t/2)} e^{-i(t, \omega)} dt, \quad (2.29)$$

for all $f \in L^2(\mathbb{R}^n)$ and with (\cdot, \cdot) the inner product in \mathbb{R}^n . As for the Fourier transform and the WFT we only discuss properties of the Wigner distribution for $f \in L^2(\mathbb{R})$. Generalizations of these results for functions in $L^2(\mathbb{R}^n)$ can be made in a rather direct way.

The Wigner distribution is, like the WFT, invariant under the action of both translation \mathcal{T}_b and frequency modulation \mathcal{M}_{ω_0} . A straightforward calculation shows

$$\mathcal{WV}[\mathcal{T}_b f](x, \omega) = \mathcal{WV}[f](x - b, \omega) \text{ and } \mathcal{WV}[\mathcal{M}_{\omega_0} f](x, \omega) = \mathcal{WV}[f](x, \omega - \omega_0).$$

Furthermore, by a change of variables in (2.25) it follows immediately that the Wigner distribution is real-valued, i.e., $\mathcal{WV}[f] = \mathcal{WV}[\overline{f}]$, and that

$$\mathcal{WV}[\overline{f}](x, \omega) = \mathcal{WV}[f](x, -\omega), \quad (2.30)$$

for all $f \in L^2(\mathbb{R})$. In particular (2.30) yields $\mathcal{WV}[f](x, \omega) = \mathcal{WV}[f](x, -\omega)$ for all real-valued $f \in L^2(\mathbb{R})$. Rewriting (2.25) enables us to derive more useful properties of the Wigner distribution.

By defining $h_{x, \omega}(t) = f(x + t/2) e^{-it\omega/2} / \sqrt{2\pi}$, for $f \in L^2(\mathbb{R})$, we can also write (2.25) as

$$\mathcal{WV}[f](x, \omega) = \int_{\mathbb{R}} h_{x, \omega}(t) \overline{h_{x, \omega}(-t)} dt.$$

Now, Parseval's formula (2.6) yields

$$\begin{aligned} \mathcal{WV}[f](x, \omega) &= \int_{\mathbb{R}} \hat{h}_{x, \omega}(\theta) \overline{\hat{h}_{x, \omega}(-\theta)} d\theta \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \hat{f}(\omega + \theta/2) \overline{\hat{f}(\omega - \theta/2)} e^{i\theta x} d\theta, \end{aligned} \quad (2.31)$$

for all $f \in L^2(\mathbb{R})$. Relation (2.31) shows that $\mathcal{WV}[f](\cdot, \omega)$ is the Fourier transform of a function in $L^1(\mathbb{R})$. So, Lemma 2.1.2 can be applied. This yields that $\mathcal{WV}[f](\cdot, \omega)$ is bounded and continuous for fixed $\omega \in \mathbb{R}$. In the same manner it follows from (2.25) that $\mathcal{WV}[f](x, \cdot)$ is bounded and continuous for fixed $x \in \mathbb{R}$. Continuity of the Wigner distribution in both variables follows immediately from (2.28) and the fact that $\mathcal{F}_h[f] \in C(\mathbb{R}^2)$, for all $f, h \in L^2(\mathbb{R})$. Concluding, we have $\mathcal{WV}[f] \in L^\infty(\mathbb{R}^2) \cap C(\mathbb{R}^2)$, for all $f \in L^2(\mathbb{R})$.

Also Relation (2.31) yields immediately

$$\mathcal{WV}[\mathcal{F}f](x, \omega) = \mathcal{WV}[f](-\omega, x), \quad (2.32)$$

for all $f \in L^2(\mathbb{R})$.

As we have seen in Example 2.2.1, the energy of a time-limited signal f is spread by means of the WFT over a larger interval than $\text{supp}(f)$, the support of f . An advantage of the Wigner distribution is that this phenomenon does not appear for the Wigner distribution. In fact, the following properties hold

$$f(x) = 0, x \notin [T_1, T_2] \implies \mathcal{WV}[f](x, \omega) = 0, x \notin [T_1, T_2], \omega \in \mathbb{R}, \quad (2.33)$$

$$\hat{f}(\omega) = 0, \omega \notin [\Omega_1, \Omega_2] \implies \mathcal{WV}[f](x, \omega) = 0, \omega \notin [\Omega_1, \Omega_2], x \in \mathbb{R}, \quad (2.34)$$

for $f \in L^2(\mathbb{R})$ and $T_1, T_2, \Omega_1, \Omega_2 \in \mathbb{R}$. Property (2.33) states that the energy of a time-limited signal is not spread over a larger interval than its support. Furthermore, (2.34) states that the energy of a band-limited signal is not spread over a larger interval than the support of its Fourier transform. We observe, that the WFT does not have these properties.

Properties (2.33) and (2.34) follow straightforwardly from (2.25) and (2.31). The converse of (2.33) and (2.34) also holds for $\hat{f} \in L^1(\mathbb{R})$ and $f \in L^1(\mathbb{R})$ respectively. This is shown in the sequel of this section.

By rewriting the integrand in (2.31) we get

$$\begin{aligned} & \hat{f}(\omega + \theta/2) \overline{\hat{f}(\omega - \theta/2)} \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} f(x) \overline{f(y)} e^{-ix(\omega + \theta/2)} e^{iy(\omega - \theta/2)} dx dy \\ &= \frac{1}{2\pi} \int_{\mathbb{R}} \int_{\mathbb{R}} f(u + t/2) \overline{f(u - t/2)} e^{-iu\theta} e^{-it\omega} du dt \\ &= \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} M[f](-\theta, t) e^{-it\omega} dt, \end{aligned}$$

with

$$M[f](\theta, t) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(u + t/2) \overline{f(u - t/2)} e^{i u \theta} du. \quad (2.35)$$

The function $M[f]$ is called the characteristic function of the Wigner distribution, see e.g. [18]. Note that $M[f](\cdot, t)$ is the Fourier transform of $f(\cdot + t/2) \overline{f(\cdot - t/2)}$, which is in $L^1(\mathbb{R})$ for all $t \in \mathbb{R}$. Using this characteristic function we obtain

$$\mathcal{WV}[f](x, \omega) = (2\pi)^{-3/2} \int_{\mathbb{R}} \int_{\mathbb{R}} M[f](\theta, t) e^{-i\theta x} e^{-it\omega} d\theta dt. \quad (2.36)$$

Introducing the autocorrelation function $R_{f,x}$ of $f \in L^2(\mathbb{R})$ by

$$R_{f,x}(t) = f(x + t/2) \overline{f(x - t/2)} / \sqrt{2\pi}$$

gives the last representation of the Wigner distribution which we discuss here. We have

$$\mathcal{WV}[f](x, \omega) = \mathcal{F}[R_{f,x}](\omega). \quad (2.37)$$

In the following example we compute the Wigner distribution of two signals and of their sum.

Example 2.3.1 We introduce the function

$$f_{p,q}(x) = \begin{cases} p \sin(2\pi x), & x \in [1, 2), \\ q \sin(8\pi x), & x \in [2, 3), \\ 0, & \text{otherwise.} \end{cases} \quad (2.38)$$

Note that $f_{1,1}$ coincides with (2.18). Furthermore, $f_{p,q}$ is the sum of two waves with finite duration and different frequencies if $p \cdot q \neq 0$.

We compute the Wigner distribution for $f_{1,0}$, $f_{0,1}$ and $f_{1,1}$. Contour plots of these distributions are depicted in Figure 2.3.a, 2.3.b and 2.3.c respectively.

Since $f_{p,q}$ is time-limited for all $p, q \in \mathbb{R}$, Relation (2.33) holds for all $f_{p,q}$. This is also clearly visible in the Figure 2.3. Besides, in Figure 2.3.a and 2.3.b maxima of $\mathcal{WV}[f_{p,q}]$ appear around frequency $\omega = 0$ and extremal points of $f_{p,q}$. Minima of $\mathcal{WV}[f_{p,q}]$ appear in these figures around $\omega = 0$ and extremal points of $f'_{p,q}$. Moreover, at these points we have $\mathcal{WV}[f_{p,q}] < 0$. Also maxima of $\mathcal{WV}[f_{p,q}]$ appear around the characteristic frequencies of the two waves.

Comparing Figure 2.3.c with Figure 2.3.a and 2.3.b, it is clear that \mathcal{WV} is a non-linear transformation. This non-linearity introduces cross-terms in the computation of the Wigner distribution of a sum of functions. The cross-terms are clearly visible in the time interval $[1.5, 2.5]$.

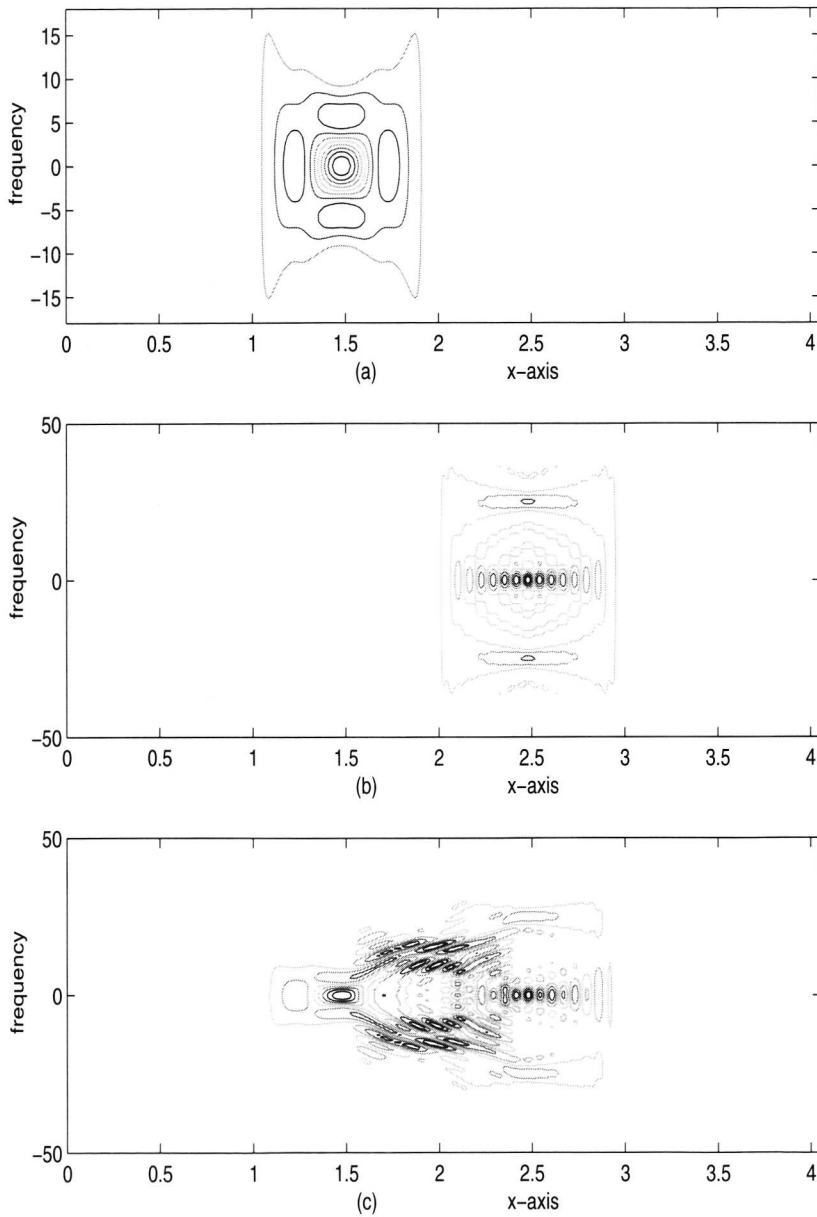


Figure 2.3: A contour plot of the amplitude of the Wigner distribution of the function a) $f_{1,0}$, b) $f_{0,1}$ and c) $f_{1,1}$.

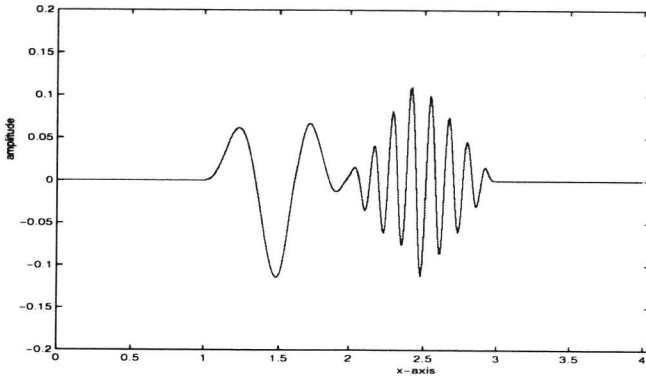


Figure 2.4: The Wigner distribution $\mathcal{WV}[f_{1,1}]$ plotted for frequency $\omega = 0$.

A more detailed discussion on these cross-terms and their effect on time-frequency analysis can be found in e.g. [18].

In Figure 2.4, $\mathcal{WV}[f_{1,1}](x, 0)$ is plotted as a function of x . As we observed before local maxima and minima can be noticed for this frequency at the values for which respectively $f_{1,1}$ and $f'_{1,1}$ attain their extremal values. It is also visible that the minimal values of $\mathcal{WV}[f_{1,1}](x, 0)$ are indeed negative.

We proceed our discussion of the Wigner distribution with a counterpart of Plancherel's formula. To deduce such a formula for the Wigner distribution we use relation (2.37).

Lemma 2.3.2 *Let $f \in L^2(\mathbb{R})$. Then*

$$\left(\int_{\mathbb{R}} |f(x)|^2 dx \right)^2 = 2\pi \int_{\mathbb{R}} \int_{\mathbb{R}} |\mathcal{WV}[f](x, \omega)|^2 d\omega dx.$$

Proof

We derive

$$\begin{aligned} \left(\int_{\mathbb{R}} |f(x)|^2 dx \right)^2 &= \int_{\mathbb{R}} |f(u)|^2 du \int_{\mathbb{R}} |f(v)|^2 dv \\ &= \int_{\mathbb{R}} \int_{\mathbb{R}} |f(x + t/2)|^2 |f(x - t/2)|^2 dt dx \end{aligned}$$

$$\begin{aligned}
&= \int_{\mathbf{R}} \int_{\mathbf{R}} |f(x+t/2) f(x-t/2)|^2 dt dx \\
&= 2\pi \int_{\mathbf{R}} \int_{\mathbf{R}} |R_{f,x}(t)|^2 dt dx.
\end{aligned}$$

Applying Plancherel's formula on the inner integral of the latter result yields

$$\begin{aligned}
\left(\int_{\mathbf{R}} |f(x)|^2 dx \right)^2 &= 2\pi \int_{\mathbf{R}} \int_{\mathbf{R}} |\mathcal{F}[R_{f,x}](\omega)|^2 d\omega dx \\
&= 2\pi \int_{\mathbf{R}} \int_{\mathbf{R}} |\mathcal{WV}[f](x, \omega)|^2 d\omega dx,
\end{aligned}$$

which follows from (2.37). \square

Lemma 2.3.2 also follow from combining (2.20) and (2.28). We observe that this lemma also yields $\mathcal{WV}[f] \in L^2(\mathbf{R}^2)$ for all $f \in L^2(\mathbf{R})$. A counterpart of Parseval's formula also exists. This is given by Moyal's formula, which is derived in the following theorem.

Theorem 2.3.3 (Moyal) *Let $f, g \in L^2(\mathbf{R})$. Then*

$$|(f, g)|^2 = 2\pi \int_{\mathbf{R}} \int_{\mathbf{R}} \mathcal{WV}[f](x, \omega) \mathcal{WV}[g](x, \omega) d\omega dx.$$

Proof

First we observe that $\mathcal{WV}[f](x, \omega) \mathcal{WV}[g](x, \omega) \in L^1(\mathbf{R}^2)$. This follows from Schwarz's inequality and Lemma 2.3.2

$$\begin{aligned}
\int_{\mathbf{R}} \int_{\mathbf{R}} |\mathcal{WV}[f](x, \omega) \mathcal{WV}[g](x, \omega)| d\omega dx &\leq \|\mathcal{WV}[f]\|_2 \|\mathcal{WV}[g]\|_2 \\
&= \|f\|_2^2 \|g\|_2^2 / 2\pi.
\end{aligned}$$

We derive as a corollary of Fubini's theorem

$$\begin{aligned}
\int_{\mathbf{R}} \mathcal{WV}[f](x, \omega) \mathcal{WV}[g](x, \omega) d\omega &= \int_{\mathbf{R}} \mathcal{F}[R_{f,x}](\omega) \overline{\mathcal{F}[R_{g,x}](\omega)} d\omega \\
&= \int_{\mathbf{R}} R_{f,x}(t) \overline{R_{g,x}(t)} dt,
\end{aligned}$$

using Parseval's formula. Integrating the latter result over x yields

$$\begin{aligned}
 & 2\pi \int_{\mathbf{R}} \int_{\mathbf{R}} \mathcal{WV}[f](x, \omega) \overline{\mathcal{WV}[g](x, \omega)} d\omega dx \\
 &= \int_{\mathbf{R}} \int_{\mathbf{R}} f(x+t/2) \overline{g(x+t/2)} \overline{f(x-t/2)} g(x-t/2) dt dx \\
 &= \int_{\mathbf{R}} \int_{\mathbf{R}} f(u) \overline{g(u)} \overline{f(v)} g(v) du dv = |(f, g)|^2.
 \end{aligned}$$

□

A further desirable property of the Wigner distribution is given in the following theorem.

Theorem 2.3.4 *Let $f \in L^2(\mathbf{R})$. Then*

$$|f(x)|^2 = \int_{\mathbf{R}} \mathcal{WV}[f](x, \omega) d\omega, \text{ if } \hat{f} \in L^1(\mathbf{R}) \quad (2.39)$$

$$|\hat{f}(\omega)|^2 = \int_{\mathbf{R}} \mathcal{WV}[f](x, \omega) dx, \text{ if } f \in L^1(\mathbf{R}). \quad (2.40)$$

Proof

We derive from (2.31)

$$\begin{aligned}
 \int_{\mathbf{R}} |\mathcal{WV}[f](x, \omega)| d\omega &\leq \frac{1}{2\pi} \int_{\mathbf{R}} \int_{\mathbf{R}} |\hat{f}(\omega + \theta/2)| |\hat{f}(\omega - \theta/2)| d\theta d\omega \\
 &= \frac{1}{2\pi} \int_{\mathbf{R}} \int_{\mathbf{R}} |\hat{f}(u)| |\hat{f}(v)| du dv = \|\hat{f}\|_1^2 / 2\pi.
 \end{aligned}$$

Fix $x \in \mathbf{R}$. Then $\mathcal{WV}[f](x, \cdot) \in L^1(\mathbf{R})$ if $\hat{f} \in L^1(\mathbf{R})$. Equivalently, $\mathcal{F}R_{f,x} \in L^1(\mathbf{R})$ if $\hat{f} \in L^1(\mathbf{R})$, cf. (2.37). Also $R_{f,x} \in C(\mathbf{R})$, since f is continuous, which follows from applying Theorem 2.1.2 on \hat{f} . Finally, $R_{f,x} \in L^1(\mathbf{R})$ since $f \in L^2(\mathbf{R})$. Now, we can apply Theorem 2.1.5, which yields

$$|f(x)|^2 = \sqrt{2\pi} R_{f,x}(0) = \int_{\mathbf{R}} \mathcal{F}[R_{f,x}](\omega) d\omega = \int_{\mathbf{R}} \mathcal{WV}[f](x, \omega) d\omega.$$

This proves (2.39). Relation (2.40) is proved in the same manner by replacing \hat{f} by f . □

Relations (2.39) and (2.40) are called the time-frequency marginals, see also [18]. As a corollary of Theorem 2.3.4 we have for $\hat{f} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$

$$\mathcal{WV}[f](x, \omega) = 0, x \notin [T_1, T_2], \omega \in \mathbb{R} \implies f(x) = 0, x \notin [T_1, T_2]$$

and for $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$

$$\mathcal{WV}[f](x, \omega) = 0, \omega \notin [\Omega_1, \Omega_2], x \in \mathbb{R} \implies \hat{f}(\omega) = 0, \omega \notin [\Omega_1, \Omega_2],$$

with $T_1, T_2, \Omega_1, \Omega_2 \in \mathbb{R}$.

A last result on the energy density of the Wigner distribution is obtained by integrating (2.40) over ω . This yields

$$\|f\|_2^2 = \int_{\mathbb{R}} \mathcal{WV}[f](x, \omega) dx d\omega, \quad (2.41)$$

for $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ or $\hat{f} \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$.

For a comprehensive list of other properties of the one-dimensional Wigner distribution we refer to [15, 44].

At the end of this discussion of the Wigner distribution, we consider its range. We have already observed that the range of the Wigner distribution, $\text{Ran}(\mathcal{WV})$, is in $L^2(\mathbb{R}^2)$. However, $\mathcal{WV} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}^2)$ is not surjective. So $\text{Ran}(\mathcal{WV}) \neq L^2(\mathbb{R}^2)$. Moreover, in [105] Wigner showed that a bilinear time-frequency distribution satisfying time-frequency marginals as in Theorem 2.3.4 cannot be non-negative everywhere in the time-frequency plane. In [47], Janssen showed that for one class of signals and one class of signals only, the Wigner distribution is positive, namely for

$$f(x) = \left(\frac{a}{\pi}\right)^{1/4} e^{-ax^2/2 + ibx^2/2 + iz\omega_0}$$

we obtain

$$\mathcal{WV}[f](x, \omega) = \frac{1}{\pi} e^{-ax^2/2 - (\omega - bx - \omega_0)^2/a},$$

with $a, b, \omega_0 \in \mathbb{R}$.

2.4 Lie Group Theory and the Heisenberg Group

In this section we will discuss the Heisenberg group and its relation to the multi-dimensional WFT and mainly to the multi-dimensional Wigner distribution. In Chapter 5 we will recall

these properties to construct time-frequency operators using a group theoretical approach.

We start with some standard definitions on Lie group theory, that can be found in e.g. [98, 102].

Definition 2.4.1 A set G with both a topological and a group structure is called a topological group if the mapping

$$(x, y) \mapsto xy^{-1} \quad (2.42)$$

is a continuous mapping from $G \times G$ onto G . A topological group G is called a Lie group if there is a differentiable structure on G , compatible with its topology, such that G converts into a C^∞ -manifold and for which the mapping (2.42) is C^∞ .

Related to a Lie group G we can also look for a Lie subgroup G' defined as a Lie group that is a subgroup of the group G and a C^∞ -submanifold of the C^∞ -manifold G . In the following example we shall consider a well-known Lie group and some of its Lie subgroups.

Example 2.4.2 Consider the group $GL(n) = \{M \in \mathbb{R}^{n \times n} \mid \det M \neq 0\}$. It can be verified rather easily that $GL(n)$ is a Lie group using the fact that the mapping $M \mapsto \det M$ is continuous. Some well-known Lie subgroups of $GL(n)$ are given by

1. $SL(n) = \{M \in GL(n) \mid \det M = 1\}$,
2. $O(n) = \{M \in GL(n) \mid M^T M = I\}$,
3. $SO(n) = \{M \in O(n) \mid \det M = 1\}$.

Another example of a well-known Lie group is the Heisenberg group, which is defined as follows.

Definition 2.4.3 The $2n+1$ -dimensional Heisenberg group H_n is identified with $\mathbb{R}^n \times \mathbb{R}^n \times \mathbb{R}$ with the multiplication law

$$(p_1, q_1, t_1)(p_2, q_2, t_2) = (p_1 + p_2, q_1 + q_2, t_1 + t_2 + ((q_1, p_2) - (p_1, q_2))/2). \quad (2.43)$$

To relate a topological group to an operator on a separable Hilbert space, we use the concept of topological group representations.

Definition 2.4.4 Let G be a topological group, H be a Hilbert space and $B(H)$ be the space of all bounded operators on H . Then a representation of G in H is a mapping $\mu : G \rightarrow B(H)$ for which

1. $\mu(x)\mu(y) = \mu(xy)$, for all $x, y \in G$,

2. $\mu(e) = \mathcal{I}$, with e the identity of G and \mathcal{I} the identity operator on \mathcal{H} ,
3. $x \mapsto \mu(x)f$ is a continuous mapping from G to H , for all $f \in H$.

Note, that Definition 2.4.4 yields that μ is a group homomorphism, which is continuous in the strong operator topology of $B(H)$.

Topological group representations may satisfy several important properties. A first desirable property of a representation is that it is unitary, i.e., $\mu(x) \in U(H)$, for all $x \in G$, where $U(H)$ denotes the space of all unitary operators on H . Furthermore, μ is said to be irreducible if $\{0\}$ and H are the only closed subspaces of H that are invariant under the action of $\mu(x)$, for all $x \in G$. A last property concerns the equivalence of two representations. A representation μ is said to be equivalent with a representation $\rho : G \rightarrow B(H)$ if there exists an operator $V \in U(H)$, such that

$$\rho = V^* \mu V. \quad (2.44)$$

Note that a unitary representation μ is a group homomorphism, which is continuous in the strong operator topology of $U(H)$. Also we observe, that for unitary representations it can be proved, see e.g. [54], that μ is irreducible if and only if for $\rho = \mu$, (2.44) only holds for $V = C\mathcal{I}$, with $|C| = 1$.

An irreducible unitary representation of H_n in the space $L^2(\mathbb{R}^n)$ is given by the Schrödinger representation

$$\mu(p, q, t)[f](x) = e^{i(p, x)} e^{i(t + (p, q)/2)} f(x + q). \quad (2.45)$$

In the sequel of this section the representation μ will denote the Schrödinger representation.

The WFT of a function $f \in L^2(\mathbb{R}^n)$ can be given in terms of this representation. According to (2.11) we have

$$\mathcal{F}_h[f](x, \omega) = (2\pi)^{-n/2} e^{-i(\omega, x)/2} (f, \mu(\omega, -x, 0)h)_2, \quad (2.46)$$

for all $f, h \in L^2(\mathbb{R}^n)$. The spectrogram $|\mathcal{F}_h[f](x, \omega)|^2$ can be given by

$$|\mathcal{F}_h[f](x, \omega)|^2 = (2\pi)^{-n} |(f, \mu(\omega, -x, t)h)_2|^2 \quad \forall t \in \mathbb{R}. \quad (2.47)$$

A relation between the Heisenberg group and the Wigner distribution can be derived using the characteristic function $M[f]$ for the multi-dimensional Wigner distribution. We derive

$$\begin{aligned} M[f](p, q) &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(u + q/2) \overline{f(u - q/2)} e^{i(p, u)} du \\ &= (2\pi)^{-n/2} e^{i(p, q)/2} \int_{\mathbb{R}^n} f(u + q) \overline{f(u)} e^{i(p, u)} du \\ &= (2\pi)^{-n/2} (\mu(p, q, 0)f, f)_2. \end{aligned} \quad (2.48)$$

This yields

$$\mathcal{WV}[f](x, \omega) = (2\pi)^{-n/2} \mathcal{F}[M[f]](x, \omega) = (2\pi)^{-n} \mathcal{F}[(\mu(\cdot, \cdot, 0)f, f)_2](x, \omega), \quad (2.49)$$

with \mathcal{WV} the multi-dimensional Wigner distribution and \mathcal{F} the n -dimensional Fourier transform. By polarization, we get that (2.49) also holds for the mixed Wigner distribution, i.e.,

$$\mathcal{WV}[f, g](x, \omega) = (2\pi)^{-n} \mathcal{F}[(\mu(\cdot, \cdot, 0)f, g)_2](x, \omega). \quad (2.50)$$

Since μ is irreducible, we have for unitary operators \mathcal{V} , as a corollary of (2.52),

$$\mathcal{WV}[f] = \mathcal{WV}[\mathcal{V}f] \iff \mathcal{V} = CI, \quad |C| = 1. \quad (2.51)$$

We have seen that the Wigner distribution is related to the Schrödinger representation by means of the characteristic function. Assume that there exists a unitary representation ρ of H_n in $U(L^2(\mathbb{R}^n))$, for which $\mu = \mathcal{V}^* \rho \mathcal{V}$, for some $\mathcal{V} \in U(L^2(\mathbb{R}^n))$. Then

$$\begin{aligned} \mathcal{WV}[\mathcal{V}f](x, \omega) &= (2\pi)^{-n} \mathcal{F}[(\mu(\cdot, \cdot, 0)\mathcal{V}f, \mathcal{V}f)_2](x, \omega) \\ &= (2\pi)^{-n} \mathcal{F}[(\mathcal{V}^* \mu(\cdot, \cdot, 0)\mathcal{V}f, f)_2](x, \omega) \\ &= (2\pi)^{-n} \mathcal{F}[(\rho(\cdot, \cdot, 0)f, f)_2](x, \omega), \end{aligned}$$

for all $f \in L^2(\mathbb{R}^n)$. This yields

$$\mathcal{WV}[\mathcal{V}f](x, \omega) = (2\pi)^{-2n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (\rho(p, q, 0)f, f)_2 e^{-i(p, x)} e^{-i(q, \omega)} d\theta dv. \quad (2.52)$$

We will return to these results in Chapter 5.

Chapter 3

The Wavelet Transform

In this chapter we study a transform that analyses signals in time/space and scale, the wavelet transform. In the first section we introduce the continuous wavelet transform (CWT) on $L^2(\mathbb{R}^n)$, particularly on $L^2(\mathbb{R})$. Definitions and properties of this transform are given. Furthermore, we discuss a group theoretical approach for the CWT. In the second section we consider the discrete wavelet transform (DWT). After an introduction by means of a sampled CWT, we consider its relation with the concepts of Multiresolution Analysis (MRA) and filter bank theory. A fast algorithm to compute the DWT and its use on discrete-time functions in $l^2(\mathbb{Z})$ are discussed.

This chapter is mainly based on existing literature on the Wavelet transform [23, 45, 51, 55, 59, 62] and its relation to filter banks by means of an MRA [16, 43, 60, 94]. Therefore, we will refer several times to the existing literature throughout this chapter. However, we also add some new ideas.

3.1 The Continuous Wavelet Transform

The CWT of $f \in L^2(\mathbb{R})$ is a linear operator defined by

$$\mathcal{W}_\psi[f](a, b) = \frac{1}{\sqrt{a}} \int_{\mathbb{R}} f(x) \overline{\psi\left(\frac{x-b}{a}\right)} dx, \quad (3.1)$$

for some $\psi \in L^2(\mathbb{R})$ and $a \in \mathbb{R}^+$ and $b \in \mathbb{R}$. By introducing the dilation operator \mathcal{D}_a on $L^2(\mathbb{R})$ by

$$\mathcal{D}_a[f](x) = \frac{1}{\sqrt{|a|}} f\left(\frac{x}{a}\right), \quad (3.2)$$

for some $a \in \mathbb{R}$ we can write (3.1) also as

$$\mathcal{W}_\psi[f](a, b) = (f, \mathcal{T}_b \mathcal{D}_a \psi), \quad (3.3)$$

with \mathcal{T}_b the shift operator as introduced in (2.12). Using this notation we prove that $\mathcal{W}_\psi[f]$ is continuous on \mathbb{H} for all $f \in L^2(\mathbb{R})$, with $\mathbb{H} = \mathbb{R}^+ \times \mathbb{R}$.

Lemma 3.1.1 $\mathcal{W}_\psi[f]$ is continuous on \mathbb{H} for all $f \in L^2(\mathbb{R})$.

Proof

Using the notation $\psi_{a,b} = \mathcal{T}_b \mathcal{D}_a \psi$, with $\psi \in L^2(\mathbb{R})$, we prove this lemma in two parts. First we give a proof for continuous functions ψ with compact support and next for arbitrary $\psi \in L^2(\mathbb{R})$. Assume $\psi \in L^2_{\text{comp}}(\mathbb{R})$, then ψ is uniformly continuous. So

$$\|\psi_{a,b} - \psi_{a',b'}\|_2 \rightarrow 0, \quad (a, b) \rightarrow (a', b').$$

Now, choose $\varepsilon > 0$ and take $g \in L^2(\mathbb{R})$ such that $\|g - \psi\| < \varepsilon/3$, with ψ continuous and compactly supported. Then

$$\begin{aligned} \|g_{a,b} - g_{a',b'}\|_2 &= \|g_{a,b} - \psi_{a,b} + \psi_{a,b} - \psi_{a',b'} + \psi_{a',b'} - g_{a',b'}\|_2 \\ &\leq \|g_{a,b} - \psi_{a,b}\|_2 + \|\psi_{a,b} - \psi_{a',b'}\|_2 + \|\psi_{a',b'} - g_{a',b'}\|_2 \\ &\leq 2\|g - \psi\|_2 + \|\psi_{a,b} - \psi_{a',b'}\|_2. \end{aligned}$$

To establish the proof, we take $(a, b) - (a', b')$ small such that $\|\psi_{a,b} - \psi_{a',b'}\|_2 < \varepsilon/3$. \square

Furthermore, from (3.3) we obtain by Schwarz's inequality

$$|\mathcal{W}_\psi[f](a, b)| \leq \|f\|_2 \cdot \|\psi\|_2 \quad \forall a \in \mathbb{R}^+, \forall b \in \mathbb{R},$$

which yields $\mathcal{W}_\psi f \in L^\infty(\mathbb{H})$, for all $f \in L^2(\mathbb{R})$. Later on we shall show that under certain conditions on ψ , $\mathcal{W}_\psi : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{H}, a^{-2} db da)$ is an isometry up to a constant that depends on ψ . Here $L^2(\mathbb{H}, a^{-2} db da)$ denotes the space of all Euclidean square integrable functions on $L^2(\mathbb{H})$ with respect to the measure $a^{-2} db da$. Besides, by Parseval's theorem we can write (3.3) in terms of the Fourier transforms of f and ψ

$$\begin{aligned} \mathcal{W}_\psi[f](a, b) &= (\mathcal{F}f, \mathcal{F}\mathcal{T}_b \mathcal{D}_a \psi) \\ &= \sqrt{a} \int_{\mathbb{R}} \hat{f}(\omega) \overline{\hat{\psi}(a\omega)} e^{ib\omega} d\omega. \end{aligned} \quad (3.4)$$

A third way to write (3.1) is given by means of a convolution product. Take $\check{\psi}(x) = \psi(-x)$, then

$$\mathcal{W}_\psi[f](a, b) = (f * \mathcal{D}_a \check{\psi})(b). \quad (3.5)$$

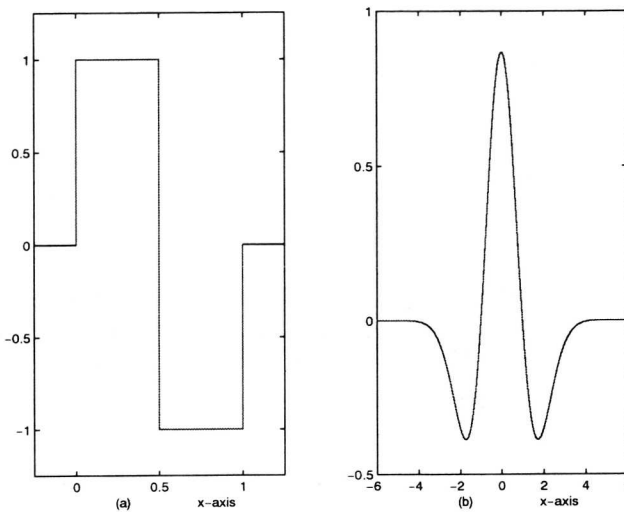


Figure 3.1: Two admissible wavelets: a) the Haar wavelet, b) the Mexican hat.

The CWT can be generalized into a transform on $L^2(\mathbb{R}^n)$ by replacing the scalar a by a non-singular matrix $A \in \mathbb{R}^{n \times n}$ and the scalar b by a vector $b \in \mathbb{R}^n$, i.e.,

$$W_\psi[s](A, b) = \frac{1}{\sqrt{|\det(A)|}} \int_{\mathbb{R}^n} s(x) \overline{\psi(A^{-1}(x - b))} dx.$$

In the literature mostly $A = aI$ is taken, see e.g. [23, 59]. For this choice, properties for the multi-dimensional CWT can be proved in a rather straightforward way using properties of the one-dimensional CWT. Murenzi followed a different approach in [66]. He introduced a multi-dimensional CWT based on a dilation operator that involves a non-singular matrix $A \in \mathbb{R}^{n \times n}$, translations in \mathbb{R}^n and a rotation operator. Also properties of this CWT resemble the properties we will deduce for the one dimensional CWT in the sequel of this chapter.

Definition 3.1.2 A function $\psi \in L^2(\mathbb{R})$ which satisfies the admissibility condition

$$0 < C_\psi = 2\pi \int_{\mathbb{R}^+} \frac{|\hat{\psi}(a\omega)|^2}{a} da < \infty, \quad (3.6)$$

for almost all $\omega \in \mathbb{R}$ is called an (admissible) wavelet.

Note that all $\psi \in L^2(\mathbb{R})$ are admissible wavelets if $\psi \neq 0$, $\hat{\psi}$ differentiable in 0 and $\hat{\psi}(0) = 0$. Furthermore, the set of admissible wavelets is dense in $L^2(\mathbb{R})$, which is not

very hard to prove, see e.g. [59].

We introduce two functions that satisfy the admissibility condition, namely the Haar wavelet and the Mexican hat.

Example 3.1.3 The Haar wavelet is defined by

$$\psi(x) = \begin{cases} 1, & x \in [0, 1/2), \\ -1, & x \in [1/2, 1), \\ 0, & \text{otherwise.} \end{cases} \quad (3.7)$$

The Haar wavelet is depicted in Figure 3.1.a. Later we will see that the Haar wavelet is admissible, since it is compactly supported and $\int_{\mathbf{R}} \psi(x) dx = 0$. These two conditions are

sufficient to guarantee that ψ is admissible. However, here we show that the Haar wavelet is an admissible wavelet by computing

$$\hat{\psi}(\omega) = \frac{1}{\sqrt{2\pi}} \left(\int_0^{1/2} e^{-i\omega x} dx - \int_{1/2}^1 e^{-i\omega x} dx \right) = \frac{1}{\sqrt{2\pi}} \left(\frac{1 + e^{-i\omega} - 2e^{-i\omega/2}}{i\omega} \right),$$

and so

$$\begin{aligned} \frac{|\hat{\psi}(a\omega)|^2}{a} &= \frac{|1 + e^{-ia\omega} - 2e^{-ia\omega/2}|^2}{a^3\omega^2} = \frac{|e^{-ia\omega/2}|^2 \cdot |e^{ia\omega/4} - e^{-ia\omega/4}|^4}{a^3\omega^2} \\ &= 16 \frac{\sin^4(a\omega/4)}{a^3\omega^2}. \end{aligned}$$

Integrating by parts yields

$$\begin{aligned} C_\psi &= \lim_{N \rightarrow \infty} \int_0^{N\omega/4} \frac{\sin^4(x)}{x^3} dx \\ &= \lim_{N \rightarrow \infty} \left. \frac{-\sin^4(x)}{2x^2} \right|_{x=0}^{N\omega/4} + \lim_{N \rightarrow \infty} \frac{1}{4} \int_0^{N\omega/4} \frac{2\sin(2x) - \sin(4x)}{x^2} dx \\ &= \lim_{N \rightarrow \infty} \left. \frac{\sin(4x) - 2\sin(2x)}{4x} \right|_{x=0}^{N\omega/4} + \lim_{N \rightarrow \infty} \int_0^{N\omega/4} \frac{\cos(2x) - \cos(4x)}{x} dx \\ &= \lim_{N \rightarrow \infty} \left(\int_0^{N\omega/4} \frac{\cos(2x) - 1}{x} dx - \int_0^{N\omega/2} \frac{\cos(4x) - 1}{x} dx \right) \\ &= \lim_{N \rightarrow \infty} Ci(N\omega/4) - \ln(N\omega/4) - Ci(N\omega/2) + \ln(N\omega/2) = \ln 2, \end{aligned}$$

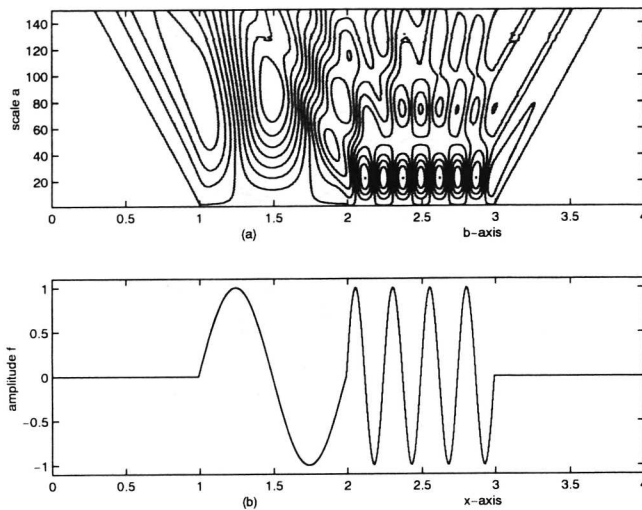


Figure 3.2: The CWT using the Haar wavelet: a) a contour plot of $\mathcal{W}_\psi[f](a, b)$ with ψ the Haar wavelet, b) the original signal f .

where Ci denotes the cosine integral, see [95].

We used the Haar wavelet to compute the CWT of the function f as given in (2.18). In Figure 3.2.a the contour plot of this CWT is depicted. In this plot maxima of $\mathcal{W}_\psi[f]$ can be observed around scale $a = 80$ for $b \in [1, 2]$ and around scale $a = 20$ for $b \in [2, 3]$. This difference in scaling behaviour is due to the difference in frequency at the corresponding intervals $x \in [1, 2]$ and $x \in [2, 3]$. Note that the frequency of f increases by a factor of 4 going from one interval to the other and that the scale corresponds to reciprocal values in frequency. Finally, we observe that the energy is also spread outside the interval $b \in [1, 3]$. Moreover, energy is spread more widely for increasing scales, which is due to the convolution product (3.5).

Example 3.1.4 The Mexican hat ψ is defined by

$$\psi(x) = -\frac{d^2}{dx^2}e^{-x^2/2} = (1 - x^2)e^{-x^2/2}. \quad (3.8)$$

The Mexican hat is depicted in Figure 3.1.b. Since $\mathcal{F}[f'](\omega) = i\omega\hat{f}(\omega)$ and

$$\int_{\mathbb{R}} e^{-x^2/2} e^{-ix\omega} d\omega = e^{-\omega^2/2},$$

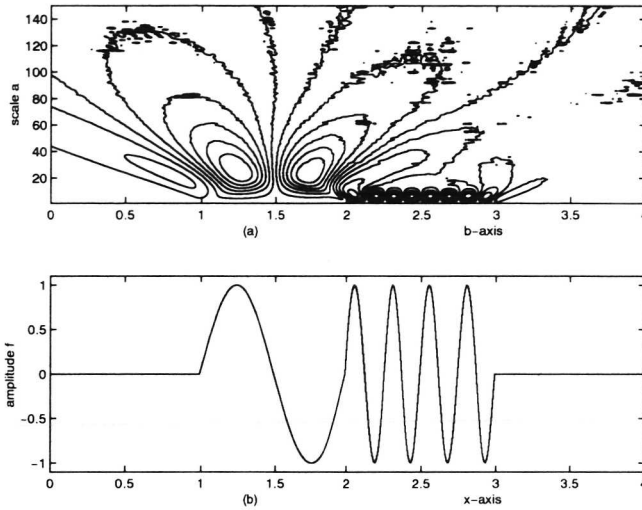


Figure 3.3: The CWT using the Mexican hat: a) a contour plot of $\mathcal{W}_\psi[f](a, b)$ with ψ the Mexican hat, b) the original signal f .

we get $\hat{\psi}(\omega) = \frac{1}{\sqrt{2\pi}}\omega^2 e^{-\omega^2/2}$. So

$$C_\psi = 2\pi \int_{\mathbf{R}^+} \frac{|\hat{\psi}(a\omega)|^2}{a} da = \int_{\mathbf{R}^+} a^3 \omega^4 e^{-a^2 \omega^2} da = 1/2 \int_{\mathbf{R}^+} y e^{-y} dy = 1/2.$$

Also with the Mexican hat we have computed the CWT of f , as defined in (2.18). A contour plot of this CWT is shown in Figure 3.3.a. As in Figure 3.2.a maxima of $W_\psi[f]$ can be observed. However, here maxima are located around scale $a = 20$ for $b \in [1, 2]$ and around scale $a = 5$ for $b \in [2, 3]$. The difference in scaling behaviour compared to the CWT using the Haar wavelet is due to a difference in frequency behaviour of both wavelets. It can be seen in Figure 3.1 that frequencies of the Haar wavelet that contain most energy are located around frequencies that are about 4 times higher than the corresponding frequencies of the Mexican hat.

To understand the admissibility condition we give some necessary and sufficient conditions on ψ such that it is a wavelet. A necessary condition on $\psi \in L^1(\mathbf{R}) \cap L^2(\mathbf{R})$ such that it satisfies (3.6) can be derived by applying Lemma 2.1.2 on C_ψ . Since $\psi \in L^1(\mathbf{R})$ this

theorem yields $\hat{\psi} \in C(\mathbb{R})$, particularly $\hat{\psi}$ is continuous in 0, which yields with (3.6)

$$\int_{\mathbb{R}} \psi(x) dx = \hat{\psi}(0) = 0.$$

A sufficient condition on $\psi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that it is a wavelet is given by

$$\int_{\mathbb{R}} \psi(x) dx = 0 \quad \text{and} \quad \int_{\mathbb{R}} |x|^l |\psi(x)| dx < \infty, \quad (3.9)$$

for $l > 1/2$. A proof of this result can be found in e.g. [59]. From this sufficient condition it follows immediately that a compactly supported $\psi \in L^2(\mathbb{R})$ is a wavelet if and only if $\int_{\mathbb{R}} \psi(x) dx = 0$, since $\int_{\mathbb{R}} |x|^l |\psi(x)| dx < \infty$, for $l \geq 0$ if ψ is compactly supported. For more sufficient conditions on ψ such that it satisfies the admissibility condition the reader may consult [23, 45, 55].

The following theorem shows that for admissible ψ , $\mathcal{W}_\psi : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{H}, a^{-2} db da)$ is an isometry up to the constant $\sqrt{C_\psi}$.

Theorem 3.1.5 *Let $f \in L^2(\mathbb{R})$ and let $\psi \in L^2(\mathbb{R})$ be an admissible wavelet. Then*

$$\int_{\mathbb{R}} |f(x)|^2 dx = 1/C_\psi \int_{\mathbb{H}} |\mathcal{W}_\psi[f](a, b)|^2 db \frac{da}{a^2}. \quad (3.10)$$

Proof

From (3.5) we get

$$|\mathcal{W}_\psi[f](a, b)|^2 = |(f * \mathcal{D}_a \check{\psi})(b)|^2.$$

We integrate both the left-hand side and the right-hand side of this relation over \mathbb{H} and apply Plancherel's formula (2.7) on the right-hand side. Now, we arrive cf. [55] at

$$\int_{\mathbb{H}} |\mathcal{W}_\psi[f](a, b)|^2 db \frac{da}{a^2} = 2\pi \int_{\mathbb{H}} |\hat{f}(\omega)|^2 |\hat{\psi}(a\omega)|^2 d\omega \frac{da}{a}.$$

So

$$\begin{aligned} \int_{\mathbb{H}} |\mathcal{W}_\psi[f](a, b)|^2 db \frac{da}{a^2} &= \int_{\mathbb{R}} |\hat{f}(\omega)|^2 \left(2\pi \int_{\mathbb{R}^+} \frac{|\hat{\psi}(a\omega)|^2}{a} da \right) d\omega \\ &= C_\psi \int_{\mathbb{R}} |\hat{f}(\omega)|^2 d\omega = C_\psi \|f\|_2^2, \end{aligned}$$

using Fubini's theorem and Plancherel's formula. \square

Relation (3.10) can be seen as Plancherel's formula for the CWT. As a result of this relation we get

$$f \in L^2(\mathbb{R}) \implies \mathcal{W}_\psi f \in L^2(\mathbb{H}, a^{-2} db da),$$

for any ψ for which $C_\psi < \infty$. Moreover, by polarization, (3.10) yields immediately Parseval's formula for the CWT

$$\int_{\mathbb{R}} f(x) \overline{g(x)} dx = \frac{1}{C_\psi} \int_{\mathbb{H}} \mathcal{W}_\psi[f](a, b) \overline{\mathcal{W}_\psi[g](a, b)} db \frac{da}{a^2}, \quad (3.11)$$

for $f, g \in L^2(\mathbb{R})$, if ψ is admissible. Formula (3.11) gives

$$\begin{aligned} \int_{\mathbb{R}} f(x) \overline{g(x)} dx &= \frac{1}{C_\psi} \int_{\mathbb{H}} \mathcal{W}_\psi[f](a, b) \int_{\mathbb{R}} 1/\sqrt{a} \overline{g(x)} \psi\left(\frac{x-b}{a}\right) dx db \frac{da}{a^2}, \\ &= \int_{\mathbb{R}} \left(\frac{1}{C_\psi} \int_{\mathbb{H}} \mathcal{W}_\psi[f](a, b) \psi\left(\frac{x-b}{a}\right) db \frac{da}{a^2 \sqrt{a}} \right) \overline{g(x)} dx, \end{aligned}$$

for all $g \in L^2(\mathbb{R})$. This proves the following formal reconstruction theorem.

Theorem 3.1.6 *Let $f \in L^2(\mathbb{R})$ and let $\psi \in L^2(\mathbb{R})$ be an admissible wavelet. Then*

$$f(x) = 1/C_\psi \int_{\mathbb{H}} \mathcal{W}_\psi[f](a, b) \psi\left(\frac{x-b}{a}\right) db \frac{da}{a^2 \sqrt{a}}, \text{ weakly in } L^2(\mathbb{R}). \quad (3.12)$$

Relation (3.12) should be interpreted by means of the first identity derived above before Theorem 3.1.6. A stronger result holds for $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ and $\hat{f} \in L^1(\mathbb{R})$. With these assumptions on f it can be proved, see e.g. [55], that (3.12) holds pointwise. Furthermore, under these assumptions, for each $x \in \mathbb{R}$, both the inner integral (over b) and the outer integral (over a) are necessarily absolutely convergent, while the double integral in (3.12) is not necessarily absolutely convergent. Note that $f(x)$ is well-defined for all $x \in \mathbb{R}$, since f is continuous due to Lemma 2.1.2.

We have shown that \mathcal{W}_ψ maps all $f \in L^2(\mathbb{R})$ into some $\mathcal{W}_\psi f \in L^2(\mathbb{H}, a^{-2} db da)$. Moreover, we have shown that we can reconstruct f from $\mathcal{W}_\psi f$, if ψ is an admissible wavelet. However the range of \mathcal{W}_ψ , $\text{Ran}(\mathcal{W}_\psi)$ is only a subset of $L^2(\mathbb{H}, a^{-2} db da)$. To characterize $\text{Ran}(\mathcal{W}_\psi)$ for admissible wavelets ψ we derive from (3.11) using Fubini's theorem

$$\begin{aligned} \mathcal{W}_\psi[f](a, b) &= (f, \mathcal{T}_b \mathcal{D}_a \psi) \\ &= \frac{1}{C_\psi \sqrt{a}} \int_{\mathbb{R}} \int_{\mathbb{H}} \mathcal{W}_\psi[f](u, v) \psi\left(\frac{x-v}{u}\right) \overline{\psi\left(\frac{x-b}{a}\right)} dv \frac{du}{u^2 \sqrt{u}} dx \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{C_\psi \sqrt{a}} \int_{\mathbf{H}} \int_{\mathbf{R}} \mathcal{W}_\psi[f](u, v) \psi\left(\frac{x-v}{u}\right) \overline{\psi\left(\frac{x-b}{a}\right)} dx dv \frac{du}{u^2 \sqrt{u}} \\
&= \int_{\mathbf{H}} k_\psi(a, b; u, v) \mathcal{W}_\psi[f](u, v) dv \frac{du}{u^2},
\end{aligned}$$

with $k_\psi(a, b; u, v) = (\mathcal{T}_v \mathcal{D}_u \psi, \mathcal{T}_b \mathcal{D}_a \psi) / C_\psi$ the reproducing kernel. Remark, that

$$k_\psi(a, b; u, v) = \overline{W_\psi[\mathcal{T}_b \mathcal{D}_a \psi](u, v)}.$$

Therefore, $k_\psi(a, b; u, v) \in L^2(\mathbf{H}, u^{-2} dv du)$, for all $(a, b) \in \mathbf{H}$. A necessary condition on $h \in L^2(\mathbf{H}, a^{-2} db da)$ such that it is in $\text{Ran}(\mathcal{W}_\psi)$ is given by

$$h(a, b) = \int_{\mathbf{H}} k_\psi(a, b; u, v) h(u, v) dv \frac{du}{u^2}. \quad (3.13)$$

Note that h is continuous on \mathbf{H} by Lemma 3.1.1. Therefore (3.13) is well defined.

Moreover, it can be shown [51], that (3.13) is also a sufficient condition on h to be in $\text{Ran}(\mathcal{W}_\psi)$. Resuming,

$$\text{Ran}(\mathcal{W}_\psi) = \{h \in L^2(\mathbf{H}, a^{-2} db da) \mid h(a, b) = \int_{\mathbf{H}} k_\psi(a, b; u, v) h(u, v) dv \frac{du}{u^2}\}.$$

Obviously, $\text{Ran}(\mathcal{W}_\psi)$ is closed due to Theorem 3.1.5. Combining this result with Theorem 3.1.5 yields that the transform $f \mapsto W_\psi f / \sqrt{C_\psi}$ is a Hilbert space isometry from $L^2(\mathbf{R})$ onto $\text{Ran}(\mathcal{W}_\psi)$ as given above.

We observe that the results on the reproducing kernel and the range of \mathcal{W}_ψ as presented here are similar to the results we presented in Section 2.2, where we considered the WFT.

In Chapter 2 we have seen that the WFT and the Wigner distribution are related to the Heisenberg group by means of the Schrödinger representation. Also the CWT is related to a group, namely the Lie group G , which is identified with \mathbf{H} with the multiplication law

$$(a_1, b_1)(a_2, b_2) = (a_1 a_2, a_1 b_2 + b_1). \quad (3.14)$$

We observe that the affine-linear group G_{af} identified with \mathbf{R}^2 with multiplication law (3.14) is isomorphic to $\mathbf{Z}_2 \ltimes G$. The left and right Haar measures (μ_L and μ_R) of G are Borel measures for which $\mu_L(gE) = \mu_L(E)$ and $\mu_R(Eg) = \mu_R(E)$ for all $g \in G$ and all Borel sets $E \subset G$. A straightforward calculation yields $d\mu_L(a, b) = a^{-2} db da$ and $d\mu_R(a, b) = a^{-1} db da$.

We introduce a representation of G in $U(L^2(\mathbb{R}))$ by

$$\pi(a, b) = \mathcal{T}_b \mathcal{D}_a, \quad \forall (a, b) \in G.$$

Obviously π is a group homomorphism. Furthermore, it is continuous in the strong operator topology of $U(L^2(\mathbb{R}))$ and $\pi(a, b)$ is unitary for all $(a, b) \in G$. In this setting a function $h \in L^2(\mathbb{R})$ is called an admissible vector if

$$(a, b) \mapsto (f, \pi(a, b)h)_2 \in L^2(\mathbb{H}, \mu_L) \quad \forall f \in L^2(\mathbb{R}). \quad (3.15)$$

Since $(f, \pi(a, b)h) = \mathcal{W}_h[f](a, b)$ Theorem 3.1.5 states that h is an admissible vector if and only if $C_h < \infty$. We already observed that the set of admissible wavelets is dense in $L^2(\mathbb{R})$ and therefore the set of all $h \in L^2(\mathbb{R})$ that satisfy (3.15) is dense in $L^2(\mathbb{R})$ as well. The following theorem shows that π is irreducible.

Theorem 3.1.7 *Let π be the unitary representation of G in $U(L^2(\mathbb{R}))$ by*

$$\pi(a, b) = \mathcal{T}_b \mathcal{D}_a, \quad \forall (a, b) \in G.$$

Then π is irreducible.

Proof

We assume π is reducible. Then there exists a closed linear subspace $V \subset L^2(\mathbb{R})$, with $V \neq \{0\}$ and $V \neq L^2(\mathbb{R})$, such that

$$\pi(a, b)V \subset V \quad \text{for all } (a, b) \in G.$$

Then there exists non-trivial vectors $g \in V$ and $f \in V^\perp$ such that $(f, \pi(a, b)g)_2 = 0$ for all $(a, b) \in G$. Now

$$\|f\|_2^2 C_g = (f, \pi(a, b)g)_2 = 0,$$

yields $f = 0$ or $g = 0$, which is in contradiction with $f \neq 0$ and $g \neq 0$. \square

To conclude, the continuous wavelet transform with a wavelet ψ is a unitary irreducible representation of the Lie group G with admissible vector ψ . More detailed studies on the wavelet transform and group theory can be found in [38, 39, 66].

3.2 The Discrete Wavelet Transform

In the previous section we considered the CWT. We showed that this integral transform is able to analyse signals both in time/space and scale. Moreover, it turned out that such signals can be recovered from their CWT. In this section we consider the problem of calculating efficiently the wavelet transform of a function and reconstructing it efficiently from its transform.

A first approach is to compute the wavelet transform only for a discrete subset $L \subset \mathbf{H}$, e.g.

$$L = \{(a_0^m, nb_0a_0^m) \mid n, m \in \mathbf{Z}\},$$

for some $a_0 > 1$ and $b_0 > 0$ and to replace the double integral in (3.12) by a double sum over L . Using such an approach we have to show that the integral can indeed be replaced by a double sum without loss of information. Furthermore, reconstruction of a function $f \in L^2(\mathbf{R})$ by means of this double sum should depend continuously on f . This kind of stability is guaranteed if the tuple (ψ, a_0, b_0) generates a wavelet frame, i.e.,

$$m_F \|f\|_2^2 \leq \sum_{(a,b) \in L} |\mathcal{W}_\psi[f](a,b)|^2 \leq M_F \|f\|_2^2, \quad \forall f \in L^2(\mathbf{R}) \quad (3.16)$$

for some constants $m_F, M_F > 0$ called the frame bounds. We observe that if (ψ, a_0, b_0) generates a wavelet frame with $m_F = M_F = 1$ and if $\|\psi\|_2 = 1$ then

$$\{\mathcal{T}_{nb_0a_0^m} \mathcal{D}_{a_0^m} \psi \mid n, m \in \mathbf{Z}\} \quad (3.17)$$

is an orthonormal basis in $L^2(\mathbf{R})$ and conversely. In that case we can transform and reconstruct any $f \in L^2(\mathbf{R})$ with respect to the lattice L using a transformation called the discrete wavelet transform (DWT). This DWT provides an efficient algorithm, that does not need to compute $\mathcal{W}_\psi[f](a,b)$ for all $(a,b) \in L$ by means of inner products. In Section 3.2.2 we will discuss this algorithm, called the pyramid algorithm.

In [22] Daubechies has given necessary and sufficient conditions on (ψ, a_0, b_0) such that (3.17) is a wavelet frame. We summarize some of these conditions in the following theorem.

Theorem 3.2.1 Assume that

1. $\text{ess} \inf_{|\omega| \in [1, a_0]} \sum_{m \in \mathbf{Z}} |\hat{\psi}(a_0^m \omega)|^2 > 0$,
2. $\text{ess} \sup_{|\omega| \in [1, a_0]} \sum_{m \in \mathbf{Z}} |\hat{\psi}(a_0^m \omega)|^2 < \infty$,
3. $\sup_{x \in \mathbf{R}} (1 + x^2)^{(1+\delta)/2} h(x) < \infty$, for some $\delta > 0$ with

$$h(x) = \sup_{|\omega| \in [1, a_0]} \sum_{m \in \mathbf{Z}} |\hat{\psi}(a_0^m \omega)| |\hat{\psi}(a_0^m \omega + x)|.$$

Then there exists an $N > 0$ such that

- a. (ψ, a_0, b_0) generates a wavelet frame for $\psi \in L^2(\mathbf{R})$, $a_0 > 1$ and $0 < b_0 < N$,
- b. $\forall \varepsilon > 0 \exists b_0 \in [N, N + \varepsilon] : (\psi, a_0, b_0)$ does not generate a wavelet frame.

Moreover, if (ψ, a_0, b_0) generates a wavelet frame, then

$$m_F \leq \frac{C_\psi}{2b_0 \ln a_0} \leq M_F. \quad (3.18)$$

Proof

Cf. [22]. □

From (3.18) it follows immediately that ψ is an admissible wavelet if (ψ, a_0, b_0) generates a wavelet frame or an orthonormal basis in $L^2(\mathbb{R})$. A more elegant and fast way to come to a DWT by means of orthonormal wavelet bases in $L^2(\mathbb{R})$ is given by the concept of a multiresolution analysis (MRA), which is considered in the sequel of this section. In Chapter 4 we shall return to the notion of frames related to MRA.

3.2.1 Multiresolution Analysis in $L^2(\mathbb{R})$

The concept of an MRA is due to Mallat [60] and Meyer [62] and was originally used as a signal processing tool by means of perfect reconstruction filter banks [43, 95], which is discussed in Section 3.2.2. We start with the definition of an MRA, following [16, 23, 60].

Definition 3.2.2 An MRA in $L^2(\mathbb{R})$ is an increasing sequence of closed subspaces V_j , $j \in \mathbb{Z}$, in $L^2(\mathbb{R})$,

$$\cdots \subset V_2 \subset V_1 \subset V_0 \subset V_{-1} \subset V_{-2} \cdots,$$

such that

1. $\bigcup_{j \in \mathbb{Z}} V_j$ is dense in $L^2(\mathbb{R})$,
2. $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$,
3. $f \in V_j \iff D^{-1}f \in V_{j-1}, \quad \forall j \in \mathbb{Z}$,
4. $\exists_{\phi \in L^2(\mathbb{R})} : \{\mathcal{T}^k \phi \mid k \in \mathbb{Z}\}$ is an orthonormal basis for V_0 ,

with $\mathcal{D} := \mathcal{D}_2$ and $\mathcal{T} := \mathcal{T}_1$ and ϕ real-valued.

For a more general concept of an MRA we can replace Condition 4 by

$$\exists_{\phi \in L^2(\mathbb{R})} : \{\mathcal{T}^k \phi \mid k \in \mathbb{Z}\} \text{ is a Riesz basis for } V_0.$$

Characterizations of a Riesz basis are given in [108]. Also the concept of an MRA for $L^2(\mathbb{R}^n)$ has been described in the literature thoroughly, see [23, 62]. In Chapter 4 we consider this general concept of an MRA in a functional analytical setting. Here we stick at Definition 3.2.2.

By Definition 3.2.2 an orthonormal basis for V_j , with $j \in \mathbb{Z}$ fixed, is given by

$$\{\mathcal{D}^j \mathcal{T}^k \phi \mid k \in \mathbb{Z}\},$$

once such function ϕ has been found. Such ϕ is called scaling function. Since \mathcal{D} is a unitary operator on $L^2(\mathbb{R})$ and V_0 is invariant under the action of \mathcal{T} , the collection

$$\{\mathcal{D}^{-1} \mathcal{T}^k \phi \mid k \in \mathbb{Z}\}$$

is an orthonormal basis for V_{-1} . As we also have $\phi \in V_{-1}$, we get

$$\phi = \sum_{k \in \mathbb{Z}} p(k) \mathcal{D}^{-1} \mathcal{T}^k \phi, \quad (3.19)$$

for some real-valued $p \in l^2(\mathbb{Z})$. In Section 3.2.2 we also want p to generate a bounded convolution operator on $l^2(\mathbb{Z})$. Therefore we require $p \in l^1(\mathbb{Z})$. Relation (3.19) is referred to as scale relation and p as scale sequence.

We consider again the inclusion $V_0 \subset V_{-1}$. Obviously we can define a subspace W_0 such that $W_0 \simeq V_{-1}/V_0$. For a unique definition of W_0 , we take $W_0 = V_{-1} \cap V_0^\perp$. Using the invariance of the subspaces V_j under the action of the unitary operator \mathcal{D}^{-1} we arrive in a natural way at the definition of the closed subspaces $W_j \subset L^2(\mathbb{R})$ by putting $W_j = V_{j-1} \cap V_j^\perp$. Recursively repeating the orthonormal decomposition of some V_{-J} into V_{-J+1} and W_{-J+1} yields

$$\begin{aligned} V_{-J} &= V_{-J+1} \oplus W_{-J+1} = V_{-J+2} \oplus W_{-J+2} \oplus W_{-J+1} \\ &= \dots = V_J \oplus \left(\bigoplus_{j=J}^{-J+1} W_j \right). \end{aligned}$$

Taking $J \rightarrow \infty$ and applying Conditions 1 and 2 from Definition 3.2.2 leads to

$$\bigoplus_{j \in \mathbb{Z}} W_j = L^2(\mathbb{R}).$$

Now assume that we can find a real-valued function $\psi \in V_{-1}$, such that $\{\mathcal{T}^k \psi \mid k \in \mathbb{Z}\}$ is an orthonormal basis for W_0 . Then $\{\mathcal{D}^j \mathcal{T}^k \psi \mid k \in \mathbb{Z}\}$ is an orthonormal basis for W_j , $j \in \mathbb{Z}$. Since the subspaces W_j are chosen to be mutually orthogonal, we then have an orthonormal basis in $L^2(\mathbb{R})$ given by $\{\mathcal{D}^j \mathcal{T}^k \psi \mid j, k \in \mathbb{Z}\}$. By (3.18) the function ψ is a wavelet and

$$\{\mathcal{D}^j \mathcal{T}^k \psi \mid k \in \mathbb{Z}\}$$

is a wavelet basis for W_j , for fixed $j \in \mathbb{Z}$. So, using these wavelet bases we are able to decompose any $f \in L^2(\mathbb{R})$ into functions at several scales.

Since the wavelet function ψ should be in V_{-1} , there exists also a scale relation for ψ

$$\psi = \sum_{k \in \mathbb{Z}} q(k) \mathcal{D}^{-1} \mathcal{T}^k \phi, \quad (3.20)$$

for some real-valued $q \in l^2(\mathbb{Z})$, the scaling sequence for ψ . As in (3.19) we will also require $q \in l^1(\mathbb{Z})$. By this relation, the problem of finding ψ can be replaced by the problem of finding q if ϕ , and therefore also p , is known. A well-known choice [23] for q is given by

$$q(k) = (-1)^k p(1 - k). \quad (3.21)$$

Note that we have introduced the concept of an MRA for $L^2(\mathbb{R})$ to constitute orthonormal wavelet bases in $L^2(\mathbb{R})$ by means of dilations and translations. These bases are equal to the wavelet frames (3.17) defined on the lattice

$$L_d = \{(2^m, n2^m) \mid m, n \in \mathbb{Z}\} \quad (3.22)$$

with frame bounds $m_F = M_F = 1$ and $\|\psi\|_2 = 1$.

An MRA for $L^2(\mathbb{R}^n)$ can be defined in a similar way, see e.g. [23, 62]. In Chapter 4 we also consider a framework of such an MRA as well as a setting for an MRA which generates wavelet bases defined on other lattices than (3.22).

3.2.2 MRA and Filter Banks

An MRA can be related to filter banks by looking at a one-level decomposition and reconstruction $V_{j-1} = V_j \oplus W_j$. At the same time this relation yields a scheme to calculate the wavelet transform of a function in $L^2(\mathbb{R})$ and to reconstruct it from its transform on L_d in a fast way. For this decomposition and reconstruction we introduce the orthoprojectors \mathcal{P}_j and \mathcal{Q}_j on V_j and W_j respectively. By definition we have

$$\mathcal{P}_{j-1} = \mathcal{P}_j + \mathcal{Q}_j, \quad (3.23)$$

for all $j \in \mathbb{Z}$. Note that Conditions 1 and 2 from Definition 3.2.2 yield

$$\lim_{j \rightarrow -\infty} \mathcal{P}_j f = f, \quad (3.24)$$

and

$$\lim_{j \rightarrow \infty} \mathcal{P}_j f = 0, \quad (3.25)$$

for all $f \in L^2(\mathbb{R})$.

The decomposition algorithm:

We assume $\mathcal{P}_{j-1}f \in V_{j-1}$ is known for a certain $j \in \mathbf{Z}$. Consequently there exists a sequence $c_{j-1} \in l^2(\mathbf{Z})$ such that

$$\mathcal{P}_{j-1}f = \sum_{k \in \mathbf{Z}} c_{j-1}(k) \mathcal{D}^{j-1} \mathcal{T}^k \phi.$$

Moreover, the sequence c_{j-1} is given by $c_{j-1}(k) = (\mathcal{D}^{j-1} \mathcal{T}^k \phi, f)_2$. Following decomposition (3.23) we have

$$\mathcal{P}_{j-1}f = \sum_{k \in \mathbf{Z}} c_j(k) \mathcal{D}^j \mathcal{T}^k \phi + \sum_{k \in \mathbf{Z}} d_j(k) \mathcal{D}^j \mathcal{T}^k \psi, \quad (3.26)$$

with the sequence d_j given by

$$\mathcal{Q}_j f = \sum_{k \in \mathbf{Z}} d_j(k) \mathcal{D}^j \mathcal{T}^k \psi.$$

Note that the sequences c_j and d_j can be calculated respectively by $c_j(k) = (\mathcal{D}^j \mathcal{T}^k \phi, f)_2$ and $d_j(k) = (\mathcal{D}^j \mathcal{T}^k \psi, f)_2$. However with the known sequence c_{j-1} we can also derive using (3.19) and (3.26)

$$\begin{aligned} c_j(k) &= (\mathcal{P}_{j-1}f, \mathcal{D}^j \mathcal{T}^k \phi)_2 = \sum_{n \in \mathbf{Z}} p(n) (\mathcal{P}_{j-1}f, \mathcal{D}^{j-1} \mathcal{T}^{2k+n} \phi)_2 \\ &= \sum_{m, n \in \mathbf{Z}} c_{j-1}(m) p(n) (\mathcal{D}^{j-1} \mathcal{T}^m \phi, \mathcal{D}^{j-1} \mathcal{T}^{2k+n} \phi)_2 \\ &= \sum_{n \in \mathbf{Z}} c_{j-1}(2k+n) p(n) = ((\downarrow 2)[c_{j-1} * \check{p}])(k), \end{aligned} \quad (3.27)$$

with $(c_{j-1} * \check{p})(k) = \sum_{n \in \mathbf{Z}} c_{j-1}(n-k) \check{p}(k)$, $\check{p}(n) = p(-n)$ and with $(\downarrow 2)$ the downsampling operator given by

$$((\downarrow 2)[u])(k) = u(2k),$$

for all $u \in l^2(\mathbf{Z})$. In the same manner we get

$$d_j = (\downarrow 2)(c_{j-1} * \check{q}). \quad (3.28)$$

So c_j and d_j are obtained from c_{j-1} by taking c_{j-1} as input for the linear time-invariant filters \check{p} and \check{q} respectively. After this filtering operation the sequences are downsampled by a factor 2. This can be visualized by means of an analysis part of a two channel filter bank as depicted in Figure 3.4.

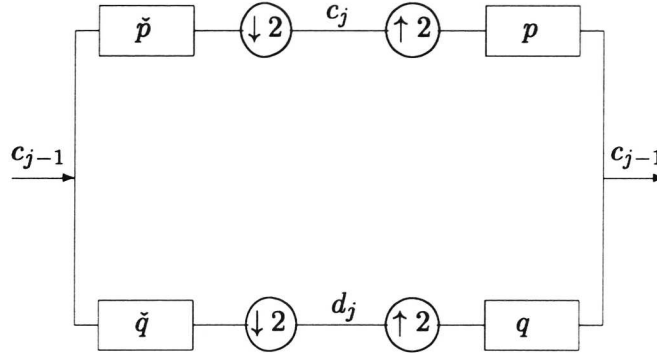


Figure 3.4: A two channel filter bank related to the wavelet filters p and q

Recursively we get expressions for c_{j+n} and d_{j+n} for $n \geq 1$, namely

$$c_{j+n} = ((\downarrow 2)C_{\check{p}})^n c_j \quad \text{and} \quad d_{j+n} = (\downarrow 2)C_{\check{q}}((\downarrow 2)C_{\check{p}})^{n-1} c_j, \quad (3.29)$$

where C_u denotes convolution with $u \in l^1(\mathbb{Z})$, i.e.

$$C_u[c](n) = \sum_{k \in \mathbb{Z}} u(n-k)c(k),$$

for all $c \in l^2(\mathbb{Z})$. In terms of filter banks we can say that c_{j+n} and d_{j+n} are obtained by iterative use of the analysis part of two channel filter bank of Figure 3.4.

The reconstruction algorithm:

Once we have computed a decomposition of $\mathcal{P}_{j-1}f$ into $\mathcal{P}_j f$ and $\mathcal{Q}_j f$ by means of the coefficients c_j and d_j , we can also recover c_{j-1} out of c_j and d_j in an efficient way. In order to come to such a reconstruction formula we will represent $\mathcal{P}_j f$ and $\mathcal{Q}_j f$ in terms of $\mathcal{D}^{j-1}\mathcal{T}^k\phi$, $k \in \mathbb{Z}$, the basis functions of V_{j-1} . For this we introduce for $f \in L^2(\mathbb{R})$ the l^2 -sequences

$$\alpha_{n,m}(k) = (\mathcal{P}_n f, \mathcal{D}^m \mathcal{T}^k \phi) \quad \text{and} \quad \beta_{n,m}(k) = (\mathcal{Q}_n f, \mathcal{D}^m \mathcal{T}^k \phi) \quad (3.30)$$

Since $\mathcal{P}_j f \in V_{j-1}$, we can write

$$\mathcal{P}_j f = \sum_{k \in \mathbb{Z}} \alpha_{j,j-1}(k) \mathcal{D}^{j-1} \mathcal{T}^k \phi.$$

An expression for $\alpha_{j,j-1}$ is found by taking the inner product with $\mathcal{D}^{j-1}\mathcal{T}^k\phi$ in both the left-hand side and the right-hand side of this equation. This yields in combination with (3.26) and (3.19)

$$\begin{aligned}\alpha_{j,j-1}(k) &= (\mathcal{P}_j f, \mathcal{D}^{j-1}\mathcal{T}^k\phi)_2 = \sum_{n \in \mathbb{Z}} c_j(n) (\mathcal{D}^j \mathcal{T}^n \phi, \mathcal{D}^{j-1}\mathcal{T}^k\phi)_2 \\ &= \sum_{m, n \in \mathbb{Z}} c_j(n) p(m) (\mathcal{D}^{j-1}\mathcal{T}^{m+2n}\phi, \mathcal{D}^{j-1}\mathcal{T}^k\phi)_2 \\ &= \sum_{n \in \mathbb{Z}} c_j(n) p(k-2n) = (((\uparrow 2)c_j) * p)(k),\end{aligned}\quad (3.31)$$

with $(\uparrow 2)$ the upsampling operator given by

$$(\uparrow 2)[u](k) = \begin{cases} u(k/2), & k \bmod 2 = 0, \\ 0, & \text{otherwise,} \end{cases}$$

for all $u \in l^2(\mathbb{Z})$. We observe that $(\downarrow 2)(\uparrow 2) = \mathcal{I}$ and $(\downarrow 2)^* = (\uparrow 2)$. In the same manner we have

$$\mathcal{Q}_j f = \sum_{k \in \mathbb{Z}} \beta_{j,j-1}(k) \mathcal{D}^{j-1}\mathcal{T}^k\phi,$$

with

$$\beta_{j,j-1} = (((\uparrow 2)d_j) * q). \quad (3.32)$$

So we derived the reconstruction formula

$$c_{j-1} = \alpha_{j,j-1} + \beta_{j,j-1} = (((\uparrow 2)c_j) * p + ((\uparrow 2)d_j) * q).$$

Hence, c_{j-1} can be recovered from $\alpha_{j,j-1}$ and $\beta_{j,j-1}$. Here $\alpha_{j,j-1}$ and $\beta_{j,j-1}$ can be seen as the output sequences of the linear time-invariant filters p and q respectively with input sequences c_j and d_j upsampled by a factor 2. This can be visualized by means of a synthesis part of a two channel filter bank as has been depicted also in Figure 3.4.

Recursively we get expressions for $\alpha_{j+n,j}$ and $\beta_{j+n,j}$ for $n \geq 1$, namely

$$\alpha_{j+n,j} = (\mathcal{C}_p(\uparrow 2))^n c_{j+n} \quad \text{and} \quad \beta_{j+n,j} = (\mathcal{C}_p(\uparrow 2))^{n-1} \mathcal{C}_q(\uparrow 2) d_{j+n}. \quad (3.33)$$

In terms of filter banks we can say that $\alpha_{j+n,j}$ and $\beta_{j+n,j}$ are obtained by iterative use of the synthesis part of two channel filter bank of Figure 3.4. This recursive approach to obtain $\alpha_{j+n,j}$ and $\beta_{j+n,j}$ is depicted in Figure 3.5 for $n = 2$.

The algorithm of decomposing and reconstructing functions by means of filter banks related to an MRA is known as the pyramid algorithm.

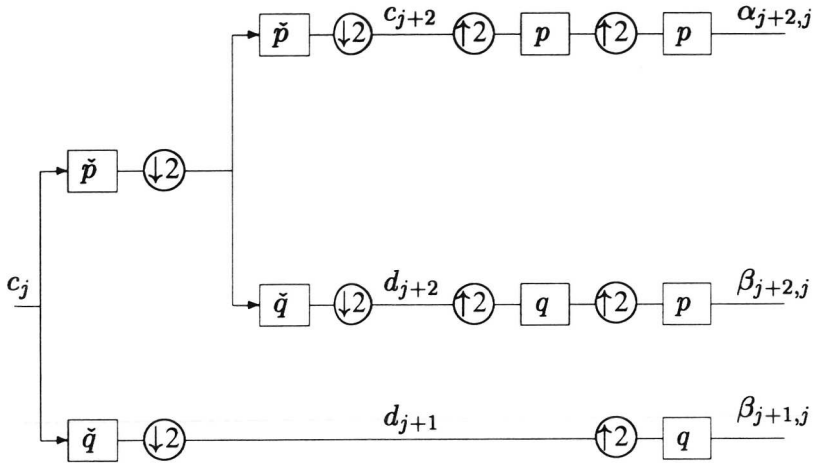


Figure 3.5: Decomposition and reconstruction by means of a filter bank at one resolution level

3.2.3 Implementation of the DWT and its Use for $l^2(\mathbb{Z})$

A problem that appears in decomposing a function $f \in L^2(\mathbb{R})$ at several scales by means of the pyramid algorithm is that the coefficients c_j should be known in order to compute $\mathcal{P}_{j+m}f$, $m \in \mathbb{N}$, in a fast way. Computing c_j does not only slow down the algorithm, it can also be impossible if only samples of f are given. The last problem appears if f is given by discrete-time measurements.

An approximation of c_j can be given by

$$c_j(k) = f(2^j k).$$

Note that if only measurements of f are available, we can identify these measurements with c_j assuming that the samples are taken from some $f \in L^2(\mathbb{R})$ at sampling rate 2^j . The following theorem, which is a generalization of an exercise in [59], shows that this is a good approximation, under certain conditions on f and the scaling function ϕ .

Theorem 3.2.3 Let $f \in L^2(\mathbb{R})$ be Hölder continuous of order $\alpha \in (0, 1]$, i.e.,

$$|f(x) - f(y)| \leq C \cdot |x - y|^\alpha, \quad \forall x, y \in \mathbb{R}, \quad (3.34)$$

for a constant $C > 0$ and let $s_j(k) = f(2^j k)$ for some $j \in \mathbb{Z}$. Let the scaling function ϕ be

continuous in \mathbb{Z} and let ϕ satisfy

$$\sum_{k \in \mathbb{Z}} \phi(x - k) = 1 \text{ a.e. } x \in \mathbb{R}, \quad (3.35)$$

where the sum converges absolutely almost everywhere, and

$$\sum_{k \in \mathbb{Z}} |k|^\alpha |\phi(k)| < \infty. \quad (3.36)$$

Then

$$\forall \varepsilon > 0 \exists J \in \mathbb{Z} \forall j < J \forall n \in \mathbb{Z} |\mathcal{P}_j[f](2^j n) - S_j[f](2^j n)| < \varepsilon,$$

with $S_j : L^2(\mathbb{R}) \rightarrow V_j$ given by

$$S_j[f](x) = \sum_{k \in \mathbb{Z}} s_j(k) \phi(2^{-j}x - k).$$

Proof

For all $j \in \mathbb{Z}$ we can write

$$\begin{aligned} |\mathcal{P}_j[f](x) - S_j[f](x)| &= |\mathcal{P}_j[f](x) - f(x) - S_j[f](x) + f(x)| \\ &\leq |\mathcal{P}_j[f](x) - f(x)| + |S_j[f](x) - f(x)|. \end{aligned}$$

From the Hölder continuity of f and Jackson's inequality, see e.g. [105], it follows that

$$\begin{aligned} \|\mathcal{P}_j f - f\|_\infty &\leq C \sup_{0 < |h| < 2^j} \|f - \mathcal{T}_h f\|_\infty \\ &\leq C_0 2^{\alpha j} \rightarrow 0 \quad (j \rightarrow -\infty), \end{aligned}$$

for some positive constants C and C_0 . The proof is established by showing that

$$|S_j[f](2^j n) - f(2^j n)| \rightarrow 0 \quad (j \rightarrow -\infty).$$

We derive

$$\begin{aligned} |S_j[f](2^j n) - f(2^j n)| &= \left| \sum_{k \in \mathbb{Z}} s_j(k) \phi(n - k) - s_j(n) \right| \\ &= \left| \sum_{k \in \mathbb{Z}} (s_j(k) - s_j(n)) \phi(n - k) \right| \\ &\leq \sum_{k \in \mathbb{Z}} |f(2^j n) - f(2^j(n - k))| |\phi(k)| \\ &\leq C 2^{\alpha j} \sum_{k \in \mathbb{Z}} |k|^\alpha |\phi(k)| \rightarrow 0 \quad (j \rightarrow -\infty). \end{aligned}$$

□

In the previous theorem we used two conditions on ϕ , which might seem strong and a bit unfamiliar as well. However, Condition (3.36) is already satisfied if ϕ is compactly supported, which is the case for the well-known Daubechies functions [23] and the spline scaling functions, i.e. B-splines [12]. For ϕ continuous in \mathbb{Z} , sufficient conditions on ϕ such that Condition (3.35) is satisfied are given in the following lemma.

Lemma 3.2.4 *Let $\phi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ such that $\{\mathcal{T}^k \phi \mid k \in \mathbb{Z}\}$ is an orthonormal set in $L^2(\mathbb{R})$ and $\hat{\phi}(0) = 1/\sqrt{2\pi}$. Then*

$$\sum_{k \in \mathbb{Z}} \phi(x - k) = 1 \text{ a.e. } x \in \mathbb{R},$$

where the sum converges absolutely almost everywhere.

Proof

Take $g(\omega) = \sum_{l \in \mathbb{Z}} |\hat{\phi}(\omega + 2\pi l)|^2$. Then g is 2π -periodic and L^1 . Its Fourier coefficients are given by

$$\begin{aligned} c_k &= 1/\sqrt{2\pi} \int_0^{2\pi} g(\omega) e^{-i\omega k} d\omega = 1/\sqrt{2\pi} \sum_{l \in \mathbb{Z}} \int_0^{2\pi} |\hat{\phi}(\omega + 2\pi l)|^2 e^{-i\omega k} d\omega \\ &= 1/\sqrt{2\pi} \int_{\mathbb{R}} |\hat{\phi}(\omega)|^2 e^{-i\omega k} d\omega = 1/\sqrt{2\pi} (\phi, \mathcal{T}^{-k} \phi) = 1/\sqrt{2\pi} \delta_{0,k}. \end{aligned}$$

This yields $g(\omega) = 1/\sqrt{2\pi}$ for all $\omega \in \mathbb{R}$, since $\hat{\phi}$ is continuous. In particular we have $g(0) = 1/\sqrt{2\pi}$, which leads in combination with $\hat{\phi}(0) = 1/\sqrt{2\pi}$ to $\hat{\phi}(2\pi l) = 0$ for all $l \in \mathbb{Z} \setminus \{0\}$. Now, put $h(x) = \sum_{k \in \mathbb{Z}} \phi(x - k)$. Since $\phi \in L^1(\mathbb{R})$ this sum converges abso-

lutely. This means that h is well defined. Furthermore, h is 1-periodic and L^1 . Its Fourier coefficients can be computed in the same manner as for g . We get $c_l = \sqrt{2\pi} \hat{\phi}(2\pi l) = \delta_{0,l}$. Writing down the Fourier series of h establishes the proof. \square

The property of ϕ as considered in Lemma 3.2.4 is called the partition of the unity. If ϕ is also continuous in $k \in \mathbb{Z}$, then the preceding lemma gives sufficient conditions on ϕ such that (3.35) holds. The conditions on ϕ such that this partition is guaranteed are quite natural. In fact, Wojtaszczyk showed in [106] that every scaling function $\phi \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$ of an MRA as defined in Definition 3.2.2 satisfies the conditions in Lemma 3.2.4. Moreover, also scaling functions ϕ for which $\{\mathcal{T}^k \phi \mid k \in \mathbb{Z}\}$ is a Riesz basis may satisfy Conditions (3.35) and (3.36). A classical example of such scaling functions ϕ are the cardinal B-spline functions [12], which will be discussed briefly in Section 4.5.

Starting with a discrete-time function, obtained from measurements, we would like to get a decomposition at various discrete-time resolution levels instead of a DWT for $L^2(\mathbb{R})$. A natural way to obtain such a decomposition is to identify a given $s \in l^2(\mathbb{Z})$ with a sequence of coefficients c_j for some $j \in \mathbb{Z}$. So, we construct a function in $L^2(\mathbb{R})$ with $\mathcal{U}_j : l^2(\mathbb{Z}) \rightarrow V_j$ by

$$\mathcal{U}_j[s](x) = 2^{-j/2} \sum_{k \in \mathbb{Z}} s(k) \phi(2^{-j}x - k),$$

for all $x \in \mathbb{R}$. Note that $\mathcal{P}_j \mathcal{U}_j = \mathcal{U}_j$ and $s(k) = (\mathcal{U}_j s, \mathcal{D}^j \mathcal{T}^k \phi)$.

A decomposition of s at level m , denoted by $s^{(m)}$ should satisfy

$$\mathcal{U}_j s^{(m)} = \mathcal{Q}_{j+m} \mathcal{U}_j s.$$

Its approximation at level m , denoted by $s_{ap}^{(m)}$ should satisfy

$$\mathcal{U}_j s_{ap}^{(m)} = \mathcal{P}_{j+m} \mathcal{U}_j s.$$

These relations hold if and only if $s^{(m)} = \beta_{j+m,j}$, $s_{ap}^{(m)} = \alpha_{j+m,j}$ and $c_j = s$, with α and β as in (3.33). Combining this result with (3.29) we arrive at the definition of the DWT for $l^2(\mathbb{Z})$.

Definition 3.2.5 Let p and q be the scale sequences as given in (3.19) and (3.20). Furthermore, let $\mathcal{G}_u = \mathcal{C}_u(\uparrow 2)$ for $u \in l^1(\mathbb{Z})$. The l^2 -DWT of a sequence $s \in l^2(\mathbb{Z})$ at scale $m \in \mathbb{N}$ is given by

$$s^{(m)} = \mathcal{G}_p^{m-1} \mathcal{G}_q \mathcal{G}_q^* (\mathcal{G}_p^*)^{m-1} s. \quad (3.37)$$

Its approximation at scale $m \in \mathbb{N}$ is given by

$$s_{ap}^{(m)} = \mathcal{G}_p^m (\mathcal{G}_p^*)^m s. \quad (3.38)$$

In the following lemma we come to some useful properties of the l^2 -DWT, which we already met in the case of the DWT for $L^2(\mathbb{R})$.

Lemma 3.2.6 For $s \in l^2(\mathbb{Z})$, let $s^{(m)}$ and $s_{ap}^{(m)}$ denote the l^2 -DWT at level m and its approximation at level m respectively for $m \in \mathbb{N}$. Then

1. $\|s\|_2^2 = \|s_a^{(M)}\|_2^2 + \sum_{m=1}^M \|s^{(m)}\|_2^2$, for all $M \in \mathbb{N}$,
2. $\lim_{m \rightarrow \infty} \|s^{(m)}\|_2 = 0$.

Proof

First we observe that \mathcal{U}_j is unitary for all $j \in \mathbb{Z}$, which yields

$$\begin{aligned}
 \|s\|_2^2 &= \|\mathcal{U}_j s\|_2^2 = \|\mathcal{P}_j \mathcal{U}_j s\|_2^2 \\
 &= \|(\mathcal{P}_{j+M} + \sum_{m=1}^M \mathcal{Q}_{j+m}) \mathcal{U}_j s\|_2^2 = \|\mathcal{P}_{j+M} \mathcal{U}_j s\|_2^2 + \sum_{m=1}^M \|\mathcal{Q}_{j+m} \mathcal{U}_j s\|_2^2 \\
 &= \|\mathcal{U}_j s_{ap}^{(M)}\|_2^2 + \sum_{m=1}^M \|\mathcal{U}_j s^{(m)}\|_2^2 = \|s_{ap}^{(M)}\|_2^2 + \sum_{m=1}^M \|s^{(m)}\|_2^2.
 \end{aligned}$$

Using (3.25) we get

$$\begin{aligned}
 \|s^{(m)}\|_2^2 &= \|\mathcal{U}_j s^{(m)}\|_2^2 = \|\mathcal{Q}_{j+m} \mathcal{U}_j s\|_2^2 \\
 &= \|\mathcal{P}_{j+m-1} \mathcal{U}_j s\|_2^2 - \|\mathcal{P}_{j+m} \mathcal{U}_j s\|_2^2 \rightarrow 0 \quad (m \rightarrow -\infty),
 \end{aligned}$$

with $j \in \mathbb{Z}$ fixed. □

Chapter 4

A Framework for Multiresolution Analysis in Hilbert Spaces

In Section 3.2 we already introduced the DWT for functions in $L^2(\mathbf{R})$ by means of a multiresolution analysis. In this chapter we consider a functional analytical setting of an MRA for separable Hilbert spaces H . Using this framework necessary and sufficient conditions on operators on H and functions in H are derived such that they constitute wavelet bases in H .

Meyer already gave strong hints for a generalization of MRA in [62]. A more general concept, like we present here, was also investigated by Goodman, Lee and Tang in [36]. However they use a different approach to construct bases in H using MRA.

This chapter is based on [75].

4.1 Frames and Riesz Systems

In the previous chapter we already mentioned the concept of frames and Riesz bases, related to the DWT in $L^2(\mathbf{R})$. Here we introduce these concept for arbitrary separable Hilbert spaces H with inner product (\cdot, \cdot) . Furthermore, we come to some results that relate these concepts.

In the sequel we denote for a countable index set \mathbf{D} , the Hilbert space of all square summable functions from \mathbf{D} into \mathbb{C} by $l^2(\mathbf{D})$ and its inner product by $(\cdot, \cdot)_{\mathbf{D}}$. The standard orthonormal basis in $l^2(\mathbf{D})$ is denoted by $\{e_j\}_{j \in \mathbf{D}}$, so $e_j(i) = \delta_{i,j}$ for $i, j \in \mathbf{D}$. The expression $l^2_0(\mathbf{D})$ indicates the linear span of the set $\{e_j \mid j \in \mathbf{D}\}$.

Definition 4.1.1 A collection $\Omega = \{v_j | j \in \mathbf{D}\}$ in H is called a frame with frame bounds m_F and M_F , $0 < m_F \leq M_F$, if for all $x \in H$ the sequence $(x, v_j)_{j \in \mathbf{D}}$ belongs to $l^2(\mathbf{D})$ and

$$m_F \|x\|^2 \leq \sum_{j \in \mathbf{D}} |(x, v_j)|^2 \leq M_F \|x\|^2. \quad (4.1)$$

Note that the wavelet frame (ψ, a_0, b_0) as introduced in Section 3.2 also satisfies this condition. Obviously, condition (3.16) equals (4.1) by taking

$$\mathbf{D} = \mathbf{Z}^2 \text{ and } v_{n,m} = T_{nb_0 a_0^n} \mathcal{D}_{a_0^n} \psi,$$

for some wavelet $\psi \in L^2(\mathbb{R})$.

For $\{v_j | j \in \mathbf{D}\}$ a frame in H , define $\mathcal{S}_F : H \rightarrow l^2(\mathbf{D})$ by

$$\mathcal{S}_F x = \sum_{j \in \mathbf{D}} (x, v_j) e_j, \quad \forall x \in H. \quad (4.2)$$

According to Definition 4.1.1,

$$m_F \|x\|^2 \leq \|\mathcal{S}_F x\|^2 \leq M_F \|x\|^2. \quad (4.3)$$

So \mathcal{S}_F is a bounded linear operator from H into $l^2(\mathbf{D})$, such that $\mathcal{S}_F^* \mathcal{S}_F$ has a bounded inverse. The optimal constants m_F and M_F are given by

$$m_F = \|(\mathcal{S}_F^* \mathcal{S}_F)^{-1}\|^{-1} \text{ and } M_F = \|\mathcal{S}_F^* \mathcal{S}_F\|.$$

The operator \mathcal{S}_F is called the frame generator associated with the frame Ω .

A straightforward computation shows that

$$\mathcal{S}_F^* \alpha = \sum_{j \in \mathbf{D}} (\alpha, e_j) v_j, \quad \forall \alpha \in l^2(\mathbf{D}).$$

Hence, the adjoint frame generator \mathcal{S}_F^* is given by

$$\mathcal{S}_F^* e_j = v_j.$$

The following lemma presents some auxiliary results on bounded operators. Using this lemma we are able to derive relations between the frame generator and its adjoint.

Lemma 4.1.2 Let $A \in B(H_1, H_2)$, with $B(H_1, H_2)$ the space of all bounded linear operators from the Hilbert space H_1 to the Hilbert space H_2 . Then, the following are equivalent

- (i) There exists an operator $B \in B(H_2, H_1)$ such that $BA = I$,

- (ii) There exists an $m > 0$ such that $\|Ax\| \geq m\|x\|$, for all $x \in H_1$,
- (iii) The null space $\text{Ker}(\mathcal{A}) = \{0\}$ and the range $\text{Ran}(\mathcal{A})$ is closed,
- (iv) There exists an operator $\mathcal{C} \in B(H_1, H_2)$ such that $\mathcal{A}^*\mathcal{C} = \mathcal{I}$,
- (v) The range $\text{Ran}(\mathcal{A}^*) = H_1$.

Proof

Assume (i) holds. Then

$$\|x\| = \|\mathcal{B}Ax\| \leq \|\mathcal{B}\| \|Ax\|.$$

So, property (ii) holds for $m = \|\mathcal{B}\|$. If property (ii) holds, then $\text{Ker}(\mathcal{A}) = \{0\}$, since $\|Ax\| = 0$ implies $\|x\| = 0$ using property (i). For proving that $\text{Ran}(\mathcal{A})$ is closed, we take a sequence $(x_n)_{n \in \mathbb{N}}$ in H_1 and a vector $y \in H_2$, such that

$$Ax_n \rightarrow y \quad (n \rightarrow \infty).$$

Then

$$\|x_k - x_n\| \leq \frac{1}{m} \|Ax_k - Ax_n\| \rightarrow 0 \quad (n \rightarrow \infty).$$

So, $(x_n)_{n \in \mathbb{N}}$ is a Cauchy sequence. For its limit $x \in H_1$ we have

$$Ax = \lim_{n \rightarrow \infty} Ax_n = y,$$

which yields that $\text{Ran}(\mathcal{A})$ is closed.

Obviously, (i) and (iv) are equivalent. If (i) holds, then also (iv) holds for $\mathcal{C} = \mathcal{B}^*$ and conversely. Therefore, instead of proving property (iv), if property (iii) holds, we prove property (i). To do this, we consider the mapping \mathcal{A}_1 from H_1 into $\text{Ran}(\mathcal{A})$, given by $\mathcal{A}_1 x = Ax$, for all $x \in H_1$. Since $\text{Ran}(\mathcal{A})$ is closed, we have that \mathcal{A}_1 is a continuous linear bijection from the Hilbert space H_1 onto the Hilbert space $\text{Ran}(\mathcal{A})$. By the inverse mapping theorem, \mathcal{A}_1 is invertible with bounded inverse $\mathcal{A}_1^{-1} : \text{Ran}(\mathcal{A}) \rightarrow H_1$. Now, define \mathcal{B} by $\mathcal{B}y = \mathcal{A}_1^{-1}y_1$, for $y = y_1 + y_2$, with $y_1 \in \text{Ran}(\mathcal{A})$ and $y_2 \in \text{Ran}(\mathcal{A})^\perp$. Then $\mathcal{B} \in B(H_2, H_1)$ and $\mathcal{B}\mathcal{A} = \mathcal{I}$.

Next, we assume that (iv) holds. Then

$$H_1 = \text{Ran}(\mathcal{A}^*\mathcal{C}) \subset \text{Ran}(\mathcal{A}^*) \subset H_1,$$

and therefore $\text{Ran}(\mathcal{A}^*) = H_1$. The last thing we have to prove is that property (iv) holds, given property (v). This is shown as follows. Take \mathcal{E} as \mathcal{A}^* restricted to $\text{Ker}(\mathcal{A}^*)^\perp$. Then \mathcal{E} is a continuous linear bijection from the Hilbert space $\text{Ker}(\mathcal{A}^*)^\perp$ onto H_1 . As before, the inverse mapping theorem yields that \mathcal{E} has a bounded inverse \mathcal{E}^{-1} from H_1 onto $\text{Ker}(\mathcal{A}^*)^\perp$. By defining an operator \mathcal{C} by $\mathcal{C}x = \mathcal{E}^{-1}x$, for all $x \in H_1$, property (iv) is established. \square

It follows from the previous lemma, that \mathcal{S}_F is an injective bounded linear operator from $l^2(\mathcal{D})$ into H with range, $\text{Ran}(\mathcal{S}_F)$, closed in H . This is equivalent with the fact that $\mathcal{S}_F^* \mathcal{S}_F$ is a boundedly invertible operator. Also $\mathcal{S}_F(\mathcal{S}_F^* \mathcal{S}_F)^{-1}$ is the right inverse of \mathcal{S}_F^* with minimal norm, which can be shown as follows. Take $\mathcal{P} = \mathcal{S}_F(\mathcal{S}_F^* \mathcal{S}_F)^{-1} \mathcal{S}_F^*$, the orthoprojector from $l^2(\mathcal{D})$ onto $\text{Ran}(\mathcal{S}_F)$. Let further \mathcal{A} be a right inverse of \mathcal{S}_F^* . Then

$$\|\mathcal{S}_F(\mathcal{S}_F^* \mathcal{S}_F)^{-1}\| = \|\mathcal{P} \mathcal{S}_F(\mathcal{S}_F^* \mathcal{S}_F)^{-1}\| = \|\mathcal{P} \mathcal{A}\| \leq \|\mathcal{A}\|.$$

In addition to the frame elements v_j we define \tilde{v}_j , $j \in \mathcal{D}$, in $\text{Ran}(\mathcal{S}_F^*)$ by

$$\tilde{v}_j = (\mathcal{S}_F^* \mathcal{S}_F)^{-1} \mathcal{S}_F^* v_j = (\mathcal{S}_F^* \mathcal{S}_F)^{-1} \mathcal{S}_F^* e_j.$$

Then for all $x \in \text{Ran}(\mathcal{S}_F^*)$

$$\begin{aligned} x &= \mathcal{S}_F^* (\mathcal{S}_F(\mathcal{S}_F^* \mathcal{S}_F)^{-1}) x = \mathcal{S}_F^* \sum_{j \in \mathcal{D}} (\mathcal{S}_F(\mathcal{S}_F^* \mathcal{S}_F)^{-1} x, e_j) e_j \\ &= \sum_{j \in \mathcal{D}} (x, (\mathcal{S}_F^* \mathcal{S}_F)^{-1} \mathcal{S}_F^* e_j) \mathcal{S}_F^* e_j = \sum_{j \in \mathcal{D}} (x, \tilde{v}_j) v_j \end{aligned}$$

and

$$\begin{aligned} x &= ((\mathcal{S}_F^* \mathcal{S}_F)^{-1} \mathcal{S}_F^*) \mathcal{S}_F x = (\mathcal{S}_F^* \mathcal{S}_F)^{-1} \mathcal{S}_F^* \sum_{j \in \mathcal{D}} (\mathcal{S}_F x, e_j) e_j \\ &= \sum_{j \in \mathcal{D}} (x, \mathcal{S}_F^* e_j) (\mathcal{S}_F^* \mathcal{S}_F)^{-1} \mathcal{S}_F^* e_j = \sum_{j \in \mathcal{D}} (x, v_j) \tilde{v}_j. \end{aligned}$$

Following these derivations and the equivalent properties as described in Lemma 4.1.2 we arrive at the following theorem.

Theorem 4.1.3 *Let $\Omega = \{v_j \mid j \in \mathcal{D}\}$ be a collection in H satisfying the right inequality in (4.1) and let \mathcal{S}_F be defined by (4.2). Then Ω is a frame if and only if the adjoint of the frame generator \mathcal{S}_F associated with Ω is surjective. If Ω is a frame, the collection $\{\tilde{v}_j \mid j \in \mathcal{D}\}$, defined by $\tilde{v}_j = (\mathcal{S}_F^* \mathcal{S}_F)^{-1} \mathcal{S}_F^* e_j$, is also a frame, which we call the frame dual to Ω .*

We observe that since $\{\tilde{v}_j \mid j \in \mathcal{D}\}$ is a frame, there exists also an associated frame generator given by $\mathcal{S}_F(\mathcal{S}_F^* \mathcal{S}_F)^{-1}$. Special cases of frames are exact frames and tight frames, which are defined as follows.

Definition 4.1.4 *An exact frame is a frame that is no longer a frame whenever any one of its elements is removed. A tight frame is a frame for which the frame bounds satisfy $m_F = M_F$.*

It can be shown in a rather straightforward way that if $\Omega = \{v_j \mid j \in \mathcal{D}\}$ is a tight frame for which $m_F = 1$ and $\|v_j\| = 1$ for all $j \in \mathcal{D}$, then Ω is an orthonormal system in H and conversely. A similar property also holds for Riesz systems, which are defined as follows.

Definition 4.1.5 The collection $\Omega = \{v_j \mid j \in \mathcal{D}\}$ in H is called a Riesz system with Riesz bounds $0 < m_R \leq M_R$, if

$$m_R \|\alpha\|_{\mathcal{D}}^2 \leq \left\| \sum_{j \in \mathcal{D}} (\alpha, e_j) v_j \right\|^2 \leq M_R \|\alpha\|_{\mathcal{D}}^2, \quad \forall \alpha \in l_0^2(\mathcal{D}). \quad (4.4)$$

Obviously, a Riesz system with Riesz bounds $m_R = M_R = 1$ is an orthonormal system, which means that orthonormal systems constitute a special set of Riesz systems. Also to a Riesz system we associate an operator in a similar way as we have done for frames.

For $\{v_j \mid j \in \mathcal{D}\}$ a Riesz system, define $S_R^{(0)} : l_0^2(\mathcal{D}) \rightarrow H$ by

$$S_R^{(0)} \alpha = \sum_{j \in \mathcal{D}} \alpha(j) v_j, \quad \forall \alpha \in l_0^2(\mathcal{D}).$$

Hence, $S_R^{(0)} e_j = v_j$. The operator $S_R^{(0)}$ extends to a bounded linear operator S_R from $l^2(\mathcal{D})$ into H , called the Riesz system!generator of Ω . This Riesz generator satisfies

$$m_R \|\alpha\|^2 \leq \|S_R \alpha\|^2 \leq M_R \|\alpha\|^2. \quad (4.5)$$

In the same manner as we concluded for the frame generator, we conclude that $S_R^* S_R$ is a boundedly invertible operator on $l^2(\mathcal{D})$. Moreover, the left inverse of S_R with minimal norm is given by $(S_R^* S_R)^{-1} S_R^*$. This can be verified in the same manner as for the right inverse of S_F^* .

Beside the Riesz system $\{v_j \mid j \in \mathcal{D}\}$ we also define the elements \tilde{v}_j , $j \in \mathcal{D}$, in $\text{Ran}(S_R)$ by $\tilde{v}_j = S_R (S_R^* S_R)^{-1} e_j$. Hence, $S_R^* \tilde{v}_j = e_j$ and consequently

$$(\tilde{v}_j, v_k) = (\tilde{v}_j, S_R e_k) = (e_j, e_k) = \delta_{j,k}.$$

Furthermore, for all $x \in \text{Ran}(S_R)$ and for $\alpha \in l^2(\mathcal{D})$ with $x = S_R \alpha$, we have

$$\begin{aligned} x &= S_R \alpha = S_R (S_R^* S_R)^{-1} (S_R^* S_R) \alpha \\ &= S_R (S_R^* S_R)^{-1} S_R^* x = S_R \sum_{j \in \mathcal{D}} ((S_R^* S_R)^{-1} S_R^* x, e_j) e_j \\ &= \sum_{j \in \mathcal{D}} (x, S_R (S_R^* S_R)^{-1} e_j) S_R e_j = \sum_{j \in \mathcal{D}} (x, \tilde{v}_j) v_j, \end{aligned}$$

and

$$\begin{aligned} x &= S_R (S_R^* S_R)^{-1} S_R^* x = S_R (S_R^* S_R)^{-1} \sum_{j \in \mathcal{D}} (S_R^* x, e_j) e_j \\ &= \sum_{j \in \mathcal{D}} (x, S_R e_j) S_R (S_R^* S_R)^{-1} e_j = \sum_{j \in \mathcal{D}} (x, v_j) \tilde{v}_j. \end{aligned}$$

These results are summarized in the following theorem.

Theorem 4.1.6 *Let $\Omega = \{v_j \mid j \in \mathcal{D}\}$ be a collection in H . Then Ω is a Riesz system if and only if there is a bounded linear injection $S_R : l^2(\mathcal{D}) \rightarrow H$ with closed range such that $S_R e_j = v_j$, $j \in \mathcal{D}$. If so, the collection $\{\tilde{v}_j \mid j \in \mathcal{D}\}$, defined by $\tilde{v}_j = S_R(S_R^* S_R)^{-1} e_j$, is the Riesz system dual to Ω .*

Note that by definition a Riesz system is a linearly independent set. A special set of Riesz systems are Riesz bases, which are defined as follows.

Definition 4.1.7 *A Riesz system which is total is a Riesz basis.*

Obviously, every Riesz system is a Riesz basis for the closure of its linear span. For a Riesz basis, the corresponding Riesz generator S_R is invertible. This yields immediately that the frame $\{v_j \mid j \in \mathcal{D}\}$ is a Riesz basis if and only if S_F^* is invertible. It can be proved, see [7], that the concepts of exact frame and of Riesz basis are equivalent. So, an exact frame can also be seen as a frame for which S_F^* is invertible. Connections between frames and Riesz systems are given in the following theorem, which results from the previous considerations.

Theorem 4.1.8 *Let $\Omega = \{v_j \mid j \in \mathcal{D}\}$ be any collection in H , and define the operator $S : l^2(\mathcal{D}) \rightarrow H$ by $S e_j = v_j$, $j \in \mathcal{D}$. Then*

- (i) *Ω is a frame if and only if SS^* is a boundedly invertible operator on H ,*
- (ii) *Ω is a Riesz system if and only if S^*S is a boundedly invertible operator on $l^2(\mathcal{D})$,*
- (iii) *Ω is a Riesz basis if and only if S is a boundedly invertible operator on $l^2(\mathcal{D})$, i.e., if both SS^* and S^*S are boundedly invertible operators.*

The relations between frames and Riesz systems as considered in this theorem are also depicted in Figure 4.1.

A characterization of Riesz systems which is used frequently in the sequel of this chapter is given in terms of a Gram matrix. For $\Omega = \{v_j \mid j \in \mathcal{D}\}$ in H , we define its Gram matrix Γ_Ω by $\Gamma_\Omega(i, j) = (v_j, v_i)_H$, $i, j \in \mathcal{D}$. Since $\Gamma_\Omega(i, j) = (S_R^* S_R e_j, e_i)$ we conclude Γ_Ω is the matrix of $S_R^* S_R$, yielding with (4.5) that Ω is a Riesz system if and only if

$$m_R I \leq \Gamma_\Omega \leq M_R I. \quad (4.6)$$

4.2 Multiresolution Analysis in Hilbert Spaces

In Section 3.2.1 we introduced the concept of an MRA for $L^2(\mathbb{R})$ following the ideas of Mallat [60] and Meyer [62]. This definition can be extended in a canonical way to an MRA for $L^2(\mathbb{R}^n)$, see e.g. [23, 62]. In this section we define an MRA for a separable Hilbert space H using mutually dependent unitary operators on H .

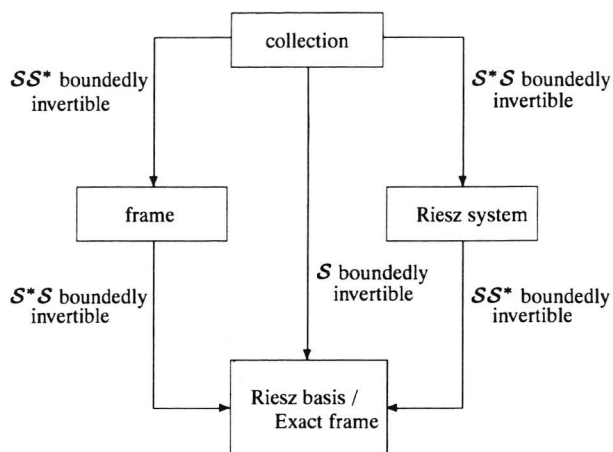


Figure 4.1: Relations between frames and Riesz systems, with \mathcal{S} the Riesz generator.

Definition 4.2.1 Let A be an $(n \times n)$ matrix with integer entries and eigenvalues λ_i , $i = 1, \dots, n$, such that $|\lambda_i| > 1$. Furthermore, let H be a separable Hilbert space, $\phi \in H$ and $\mathcal{U}_1, \mathcal{U}_{2,1}, \dots, \mathcal{U}_{2,n}$ unitary operators on H , such that $\mathcal{U}_{2,1}, \dots, \mathcal{U}_{2,n}$ mutually commute. Then $[\phi, \mathcal{U}_1, \mathcal{U}_{2,1}, \dots, \mathcal{U}_{2,n}]$ is an MRA in H if

- (i) $\{\mathcal{U}_2^k \phi \mid k \in \mathbb{Z}^n\}$ is a Riesz system in H ,
- (ii) $\phi \in \text{clos span}\{\mathcal{U}_1 \mathcal{U}_2^k \phi \mid k \in \mathbb{Z}^n\}$,
- (iii) $\mathcal{U}_2^k \mathcal{U}_1 = \mathcal{U}_1 \mathcal{U}_2^{A^k}$, for all $k \in \mathbb{Z}^n$,

with

$$\mathcal{U}_2^k = \mathcal{U}_{2,1}^{k_1} \cdots \mathcal{U}_{2,n}^{k_n}.$$

In the sequel ϕ is called the MRA generator.

To compare this definition with Definition 3.2.2 we construct a nested sequence of closed subspaces for H by

$$V_j = \text{clos span}\{\mathcal{U}_1^{-j} \mathcal{U}_2^k \phi \mid k \in \mathbb{Z}^n\}.$$

Then we have

$$\mathcal{U}_1(V_j) = V_{j-1}, \mathcal{U}_2^k(V_j) = V_j, k \in \mathbb{Z}^n, \text{ and } V_j \subset V_{j-1}.$$

For $n = 1$, $A = 2$, $\mathcal{U}_1 = \mathcal{D}^{-1}$, $\mathcal{U}_{2,1} = \mathcal{T}$ and $H = L^2(\mathbb{R})$, Definition 3.2.2 and Definition 4.2.1 are nearly the same. However in Definition 3.2.2 an orthonormal basis for V_0 was

constructed, which is a special case of a Riesz basis, and in the definition above we also did not introduce the conditions

$$\text{clos} \bigcup_{j \in \mathbb{Z}} V_j = H \quad \text{and} \quad \bigcap_{j \in \mathbb{Z}} V_j = \{0\}, \quad (4.7)$$

which occur in Definition 3.2.2. Whether or not inserting these conditions will not change any of the further derivations. Therefore these conditions have been omitted.

The concept of MRA in H is now used to construct Riesz systems in H of a special kind. We start this construction by defining a unique countable collection of closed subspaces W_j , $j \in \mathbb{Z}$, like we have done as well in Section 3.2.1, namely by writing

$$W_j = V_{j-1} \cap V_j^\perp,$$

for all $j \in \mathbb{Z}$. Since \mathcal{U}_1 and \mathcal{U}_2 are unitary operators on H , we have

$$\begin{aligned} \mathcal{U}_1(W_j) &= \mathcal{U}_1(V_{j-1} \cap V_j^\perp) = \mathcal{U}_1(V_{j-1}) \cap \mathcal{U}_1(V_j^\perp) \\ &= \mathcal{U}_1(V_{j-1}) \cap (\mathcal{U}_1(V_j))^\perp = W_{j-1}, \end{aligned}$$

and similarly $\mathcal{U}_{2,l}(W_j) = W_j$, $l = 1, \dots, n$. Obviously, since $W_j = \mathcal{U}_1^{-j}(W_0)$, each Riesz basis Ω for W_0 yields the Riesz basis $\mathcal{U}_1^{-j}(\Omega)$ for W_j , $j \in \mathbb{Z}$, with the same Riesz bounds.

Following the same orthogonal decomposition as in Section 3.2.1 and adding Condition 4.7 we arrive at

$$H = \bigoplus_{j=-\infty}^{\infty} W_j.$$

From this direct sum decomposition of H it follows immediately that $\bigcup_{j \in \mathbb{Z}} \mathcal{U}_1^j(\Omega)$ is a Riesz basis for H , if Ω is a Riesz basis for W_0 .

Now the idea is to construct Riesz systems in W_0 of the form $\{\mathcal{U}_2^k \psi \mid k \in \mathbb{Z}^n\}$. It will turn out that, for constructing a Riesz basis of this form in general more than one element ψ will be needed. Our aim here is to prove existence of an $N \in \mathbb{N}$ and of a collection

$$\{\psi_1, \dots, \psi_{N-1}\} \subset V_1,$$

such that

- (a) $(\psi_l, \mathcal{U}_2^k \phi) = 0$, $l = 1, \dots, N-1$, for all $k \in \mathbb{Z}^n$, i.e., $\psi_l \in V_0^\perp$,
- (b) $\{\mathcal{U}_2^k \psi_l \mid l = 1, \dots, N-1, k \in \mathbb{Z}^n\}$ is a Riesz basis for W_0 .

Since $V_{-1} = V_0 \oplus W_0$ Condition (b) can be also written as

(b') $\{\mathcal{U}_2^k \psi_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$ is a Riesz basis for V_{-1} , with $\psi_0 = \phi$.

Since $V_0 \subset V_{-1}$ and $W_0 \subset V_{-1}$ and since $\{\mathcal{U}_1 \mathcal{U}_2^k \phi \mid k \in \mathbb{Z}^n\}$ is a Riesz basis in V_{-1} , we come to

$$\phi = \sum_{k \in \mathbb{Z}^n} p(k) \mathcal{U}_1 \mathcal{U}_2^k \phi, \quad (4.8)$$

$$\psi_l = \sum_{k \in \mathbb{Z}^n} q_l(k) \mathcal{U}_1 \mathcal{U}_2^k \phi, \quad l = 1, \dots, N-1, \quad (4.9)$$

where $p \in l^2(\mathbb{Z}^n)$, known, and the $q_l \in l^2(\mathbb{Z}^n)$, to be determined, are the generating sequences. For $\mathcal{U}_1 = \mathcal{D}^{-1}$ and $\mathcal{U}_2 = \mathcal{T}$, these sequences are called scaling sequences as we have seen already in Section 3.2.1. In the sequel the term scaling sequence will only be used for cases in which $\mathcal{U}_1 = \mathcal{D}^{-1}$ has been chosen.

So the idea is to formulate constraints on the sequences q_l , given the sequence p , such that the Conditions (a) and (b') are satisfied. Therefore we reformulate these conditions in terms of the generating sequences.

Condition (a) can be put in a rather straightforward way in terms of the generating sequences by substituting (4.8) and (4.9) into this condition and using $\mathcal{U}_2^k \mathcal{U}_1 = \mathcal{U}_1 \mathcal{U}_2^{A^k}$. We get

$$(\psi_l, \mathcal{U}_2^k \phi) = (\tau_\phi * q_l, \mathcal{R}^{A^k} p)_{\mathbb{Z}^n},$$

with $\tau_\phi(k) = (\phi, \mathcal{U}_2^k \phi)$, $k \in \mathbb{Z}^n$, and $\mathcal{R}^m = \mathcal{R}_1^{m_1} \dots \mathcal{R}_n^{m_n}$ for $m \in \mathbb{Z}^n$, a composition of bilateral shift operators on $l^2(\mathbb{Z}^n)$, each one acting along a standard basis vector of \mathbb{Z}^n . So Condition (a) is equivalent with

$$(\tau_\phi * q_l, \mathcal{R}^{A^k} p)_{\mathbb{Z}^n} = 0, \quad \forall l \in \{1, \dots, N-1\} \quad \forall k \in \mathbb{Z}^n. \quad (4.10)$$

In order to put Condition (b') in terms of the generating sequences we present the following lemma.

Lemma 4.2.2 *Let $[\phi, \mathcal{U}_1, \mathcal{U}_{2,1}, \dots, \mathcal{U}_{2,n}]$ be an MRA and let p be the generating sequence of ϕ . Then*

$$\{\mathcal{R}^{A^k} p \mid k \in \mathbb{Z}^n\}$$

is a Riesz system in $l^2(\mathbb{Z}^n)$. Furthermore, let $\psi_0 = \phi$ and ψ_l , $l = 1, \dots, N-1$, be in W_0 with generating sequences q_l . Then

$$\{\mathcal{U}_2^k \psi_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$$

is a Riesz basis for V_{-1} if and only if

$$\{\mathcal{R}^{A^k} q_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\},$$

with $q_0 = p$, is a Riesz basis for $l^2(\mathbb{Z}^n)$.

Proof

We introduce the boundedly invertible operator $B : V_{-1} \rightarrow l^2(\mathbb{Z}^n)$ by

$$Bf = \alpha \text{ if and only if } f = \sum_{k \in \mathbb{Z}^n} \alpha(k) \mathcal{U}_1 \mathcal{U}_2^k \phi.$$

Since $B\mathcal{U}_2^k = \mathcal{R}^{A^k} B$, $k \in \mathbb{Z}^n$ and $B\phi = p$, applying B on the Riesz system $\{\mathcal{U}_2^k \phi \mid k \in \mathbb{Z}^n\}$ yields $\{\mathcal{R}^{A^k} p \mid k \in \mathbb{Z}^n\}$. This is also a Riesz system, since B is a boundedly invertible operator. The second result follows immediately by observing that

$$\{\mathcal{R}^{A^k} q_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\} = B(\{\mathcal{U}_2^k \psi_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}).$$

□

From this lemma it follows that if we can construct sequences $q_l \in l^2(\mathbb{Z}^n)$, such that

- $(\tau_\phi * q_l, \mathcal{R}^{A^k} p)_{\mathbb{Z}^n} = 0, \forall l \in \{1, \dots, N-1\} \quad \forall k \in \mathbb{Z}^n,$
- $\{\mathcal{R}^{A^k} q_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$ is a Riesz basis for $l^2(\mathbb{Z}^n)$, with $q_0 = p$,

then elements ψ_l , $l = 1, \dots, N-1$, in V_{-1} can be constructed for which Conditions (a) and (b) are satisfied. This naturally leads us to the next item.

4.3 Riesz Systems Generated by Unitary Operators

In Section 4.1 we already considered necessary and sufficient conditions on a set Ω such that it is a Riesz system. Now, we will deal with sets Ω of a special kind, namely those sets that are generated by unitary operators acting on one or more elements in H . This is done, since such Riesz systems occurred in Lemma 4.2.2.

First we consider the case of a Riesz system generated by several unitary operators all acting on one element in H . So, our aim is to derive necessary and sufficient conditions on the tuple $[\mathcal{U}_1, \dots, \mathcal{U}_n, \phi]$, such that $\{\mathcal{U}^j \phi \mid j \in \mathbb{Z}^n\}$ is a Riesz system, with $\mathcal{U}^j = \mathcal{U}_1^{j_1} \dots \mathcal{U}_n^{j_n}$. Besides, we compute its dual Riesz system. These derivations follow Section 4.1 with $\mathcal{D} = \mathbb{Z}^n$ and with the Riesz generator \mathcal{S} given by $\mathcal{S}e_j = \mathcal{U}^j \phi$.

By Section 4.1, $\{\mathcal{U}^j \phi \mid j \in \mathbb{Z}^n\}$ is a Riesz system if and only if its Gram matrix Γ , given by

$$\Gamma(i, j) = (\mathcal{U}^j \phi, \mathcal{U}^i \phi) = (\phi, \mathcal{U}^{i-j} \phi) = \tau_\phi(i - j),$$

satisfies (4.6). Observing that $\mathcal{S}^* \mathcal{S}$ with matrix Γ acts by convolution on $l^2(\mathbb{Z}^n)$,

$$\mathcal{S}^* \mathcal{S} \alpha = \tau_\phi * \alpha,$$

we arrive at the following theorem.

Theorem 4.3.1 For commuting unitary operators $\mathcal{U}_1, \dots, \mathcal{U}_n$ on H and $\phi \in H$, the collection $\{\mathcal{U}^j \phi \mid j \in \mathbb{Z}^n\}$ is a Riesz system if and only if the sequence τ_ϕ defined by

$$\tau_\phi(j) = (\phi, \mathcal{U}^j \phi), \quad j \in \mathbb{Z}^n,$$

yields a boundedly invertible convolution operator on $l^2(\mathbb{Z}^n)$, i.e., if and only if

$$0 < \operatorname{ess} \inf_{z \in \mathbb{T}^n} \hat{\tau}_\phi(z) \leq \operatorname{ess} \sup_{z \in \mathbb{T}^n} \hat{\tau}_\phi(z) < \infty, \quad (4.11)$$

where $\hat{\tau}_\phi$ denotes the discrete Fourier transform of τ_ϕ

$$\hat{\tau}_\phi(z) = \sum_{j \in \mathbb{Z}^n} \tau_\phi(j) z^{-j}, \quad z \in \mathbb{T}^n,$$

with $z^j = z_1^{j_1} \dots z_n^{j_n}$, $j \in \mathbb{Z}^n$, and \mathbb{T}^n the n -fold product of the unit circle with normalized Lebesgue measure μ .

Note that from this theorem it follows that for $\tau_\phi \in l^1(\mathbb{Z}^n)$ the collection

$$\Omega = \{\mathcal{U}^j \phi \mid j \in \mathbb{Z}^n\}$$

is a Riesz system if and only if $\hat{\tau}_\phi$ has no zeroes on \mathbb{T}^n .

By definition, the dual Riesz system is given by $\mathcal{S}(\mathcal{S}^* \mathcal{S})^{-1} e_j$, where $\mathcal{S} e_j = \mathcal{U}^j \phi$ and $\mathcal{S}^* \mathcal{S} \alpha = \tau_\phi * \alpha$, for all $\alpha \in l^2(\mathbb{Z})$. This yields immediately that the dual Riesz system $\tilde{\Omega}$ of Ω is of the form

$$\tilde{\Omega} = \{\mathcal{U}^j \tilde{\phi} \mid j \in \mathbb{Z}^n\}$$

with $\tilde{\phi}$ given by

$$\tilde{\phi} = \sum_{j \in \mathbb{Z}^n} \tilde{\tau}_\phi(j) \mathcal{U}^j \phi,$$

where $\tilde{\tau}_\phi * \tau_\phi = e_0$.

Next we replace the vector $\phi \in H$ by a finite collection $\{\phi_1, \dots, \phi_N\}$ and pose the same problem, namely under which conditions

$$\Omega_N = \{\mathcal{U}^j \phi_l \mid l = 1, \dots, N, j \in \mathbb{Z}^n\}$$

is a Riesz system. For this we take as index set $\mathcal{D} = \{1, \dots, N\} \times \mathbb{Z}^n$. Furthermore, we define the unitary operator \mathcal{E}_N from $l^2(\mathcal{D})$ into $L^2(\mathbb{T}^n, \mathbb{C}^N) = L^2(\mathbb{T}^n) \otimes \mathbb{C}^N$ by $(\mathcal{E}_N e_{l,j})(z) = z^j \varepsilon_l$, with $\{\varepsilon_1, \dots, \varepsilon_N\}$ the standard orthonormal basis in \mathbb{C}^N . We see that \mathcal{E}_N is a Riesz system if and only if the Gram matrix Γ with entries $(\mathcal{U}^j \phi_l, \mathcal{U}^i \phi_m)_{(m,j),(k,i)}$ represents a bounded invertible operator on $l^2(\mathbb{Z}^n)$. A straightforward computation shows

$$(\mathcal{E}_N \Gamma e_{m,j})(z) = z^j \hat{\Gamma}(z) \varepsilon_m,$$

where $\hat{\Gamma}(z) \in \mathcal{C}^{N \times N}$ is defined a.e. by $(\hat{\Gamma}(z))_{k,m} = \sum_{j \in \mathbb{Z}^n} (\phi_m, U^j \phi_k) z^{-j}$. Hence the Gram matrix Γ represents a boundedly invertible operator on $l^2(\mathcal{D})$ if and only if the matrix valued function $\hat{\Gamma}$ from \mathbb{T}^n into $\mathcal{C}^{N \times N}$ satisfies

$$\exists m, M > 0 \quad \forall z \in \mathbb{T}^n \quad mI_N \leq \hat{\Gamma}(z) \leq MI_N \quad \text{a.e.} \quad (4.12)$$

4.4 Riesz Bases in $l^2(\mathbb{Z}^n)$

In the previous section we derived necessary and sufficient conditions such that

$$\Omega_N = \{U^j \phi_l \mid l = 1, \dots, N, j \in \mathbb{Z}^n\}$$

is a Riesz system. According to Lemma 4.2.2 such conditions can also be derived in terms of the generating sequences. Therefore in this section we deal with the following problem. Let the sequence γ yield a boundedly invertible convolution operator on $l^2(\mathbb{Z}^n)$ and let $\beta_0 \in l^2(\mathbb{Z}^n)$. Find necessary and sufficient conditions on sequences β_l , $l = 1, \dots, N-1$, in $l^2(\mathbb{Z}^n)$, and determine N such that

- (i) $(\gamma * \beta_l, \mathcal{R}^{A^k} \beta_0)_Z = 0$, $\forall l \in \{1, \dots, N-1\} \quad \forall k \in \mathbb{Z}^n$,
- (ii) $\{\mathcal{R}^{A^k} \beta_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$ is a Riesz basis for $l^2(\mathbb{Z}^n)$.

We reformulate this into terms of the Hilbert space $L^2(\mathbb{T}^n)$.

Since we deal with a rather arbitrary matrix $A \in \mathbb{Z}^{n \times n}$ we introduce the so-called Smith normal form of a matrix with integer entries, which is given in the following theorem. In [56] one can find a proof of this theorem for matrices over a ring of polynomials in one variable. This result generalizes immediately to the case of matrices over the ring of integers.

Theorem 4.4.1 (Smith normal form) *Let $A \in \mathbb{Z}^{n \times n}$. Then there are unimodular matrices $U, V \in \mathbb{Z}^{n \times n}$, i.e., $\det(U) = \det(V) = 1$, and a diagonal matrix $D \in \mathbb{Z}^{n \times n}$, such that*

$$A = UDV. \quad (4.13)$$

This factorization is not unique.

In the sequel we use the notation $L = |\det(D)|$.

It can be proved by some straightforward computations that the problem posed in the beginning of this section is equivalent with the following one. Give necessary and sufficient conditions on sequences β_l , $l = 1, \dots, N-1$, in $l^2(\mathbb{Z}^n)$, and determine N such that

- (i) $(\gamma * \beta_l, \mathcal{R}^{D^k} \beta_0)_Z = 0$, $\forall l \in \{1, \dots, N-1\} \quad \forall k \in \mathbb{Z}^n$,

(ii) $\{\mathcal{R}^{Dk}\beta_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$ is a Riesz basis for $l^2(\mathbb{Z}^n)$, with $D \in \mathbb{Z}^{n \times n}$ a diagonal matrix involved in the Smith normal form of D .

Let now $d_i = D(i, i)$, $i = 1, \dots, n$. Define $\omega_{d_i} = e^{2\pi i/d_i}$, $i = 1, \dots, n$, and K_n the n -fold segment of all $z \in \mathbb{T}^n$ such that

$$\arg(z_i) \in [0, 2\pi/d_i), \quad i = 1, \dots, n.$$

We observe, that $\{z^k \mid k \in \mathbb{Z}^n\}$ is an orthonormal basis for $L^2(\mathbb{T}^n)$ and

$$\{\sqrt{L} z_1^{k_1 d_1} \dots z_n^{k_n d_n} \mid k \in \mathbb{Z}^n\}$$

is an orthonormal basis for $L^2(K_n)$. So

$$\{\sqrt{L} z_1^{k_1 d_1} \dots z_n^{k_n d_n} e_i \mid i = 1, \dots, N, k \in \mathbb{Z}^n\}$$

is an orthonormal basis for $L^2(K_n, \mathbb{C}^N)$, the Hilbert space consisting of all \mathbb{C}^N -valued Euclidean square integrable functions on K_n . In dealing with the above stated problem, we present some auxiliary results.

The proof of the following lemma is based on the fact, that the $(n \times n)$ Fourier matrix F_n with entries

$$F_n(i, j) = 1/\sqrt{n} \omega_n^{ij}, \quad i, j = 0, \dots, n-1,$$

is unitary. Furthermore, we use the notation

$$P_{\hat{g}, \hat{h}}(z) = \sum_{j_1=0}^{|d_1|-1} \dots \sum_{j_n=0}^{|d_n|-1} \hat{g}(\omega_{d_1}^{j_1} z_1, \dots, \omega_{d_n}^{j_n} z_n) \overline{\hat{h}(\omega_{d_1}^{j_1} z_1, \dots, \omega_{d_n}^{j_n} z_n)}. \quad (4.14)$$

Lemma 4.4.2 Let $g, h \in l^2(\mathbb{Z}^n)$. Then

$$\frac{1}{L} \int_{\mathbb{T}^n} P_{\hat{g}, \hat{h}}(z) z^m d\mu_n(z) = \begin{cases} (g, \mathcal{R}^{Dk} h)_{\mathbb{Z}^n} & \text{if } m = Dk, k \in \mathbb{Z}^n, \\ 0 & \text{if } m \neq Dk, k \in \mathbb{Z}^n. \end{cases}$$

Proof

We consider the following computation

$$\begin{aligned} & \frac{1}{L} \int_{\mathbb{T}^n} P_{\hat{g}, \hat{h}}(z) z^m d\mu_n(z) \\ &= \frac{1}{L} \sum_{j_1=0}^{|d_1|-1} \dots \sum_{j_n=0}^{|d_n|-1} \omega_{d_1}^{-j_1 m_1} \dots \omega_{d_n}^{-j_n m_n} \int_{\mathbb{T}^n} \hat{g}(z) \overline{\hat{h}(z) z^{-m}} d\mu_n(z) \\ &= \begin{cases} \int_{\mathbb{T}^n} \hat{g}(z) \overline{\hat{h}(z) z^{-m}} d\mu_n(z) & \text{if } m = Dk, k \in \mathbb{Z}^n, \\ 0 & \text{if } m \neq Dk, k \in \mathbb{Z}^n. \end{cases} \end{aligned}$$

The proof is completed by observing that the n -dimensional discrete Fourier transform of $\mathcal{R}^l h$ is given by $z^{-l} \hat{h}$. \square

From this lemma we deduce the following result.

Lemma 4.4.3 *Let $g, h \in l^2(\mathbb{Z}^n)$. Then for all $k \in \mathbb{Z}^n$*

$$\int_{K_n} P_{\hat{g}, \hat{h}}(z) z_1^{k_1 d_1} \dots z_n^{k_n d_n} d\mu_n(z) = (g, \mathcal{R}^{Dk} h)_{\mathbb{Z}^n}.$$

Proof

Obviously, $P_{\hat{g}, \hat{h}}(z) z_1^{k_1 d_1} \dots z_n^{k_n d_n}$ remains unchanged if z_i is replaced by $\omega_{\lambda_i}^{m_i} z_i$, $i = 1, \dots, n$, for $m \in \mathbb{Z}^n$ and so

$$\int_{K_n} P_{\hat{g}, \hat{h}}(z) z_1^{k_1 d_1} \dots z_n^{k_n d_n} d\mu_n(z) = \int_{K_n^m} P_{\hat{g}, \hat{h}}(z) z_1^{k_1 d_1} \dots z_n^{k_n d_n} d\mu_n(z),$$

with

$$K_n^m = [m_1 \omega_{d_1}, (m_1 + 1) \omega_{d_1}) \times \dots \times [m_n \omega_{d_n}, (m_n + 1) \omega_{d_n}).$$

Consequently the result follows from Lemma 4.4.2. \square

By Lemma 4.4.3, Condition (i) can be written as

$$\int_{K_n} P_{\gamma \beta_l, \beta_0}(z) z_1^{k_1 d_1} \dots z_n^{k_n d_n} d\mu_n(z) = 0.$$

Since this relation must hold for every $k \in \mathbb{Z}^n$ and since

$$\{\sqrt{L} z^{Dk} \mid k \in \mathbb{Z}^n\}$$

is an orthonormal basis for $L^2(K_n)$, we get that

$$P_{\gamma \beta_l, \beta_0} = 0 \text{ a.e. on } K_n, l = 1, \dots, N-1. \quad (4.15)$$

So, sequences β_l that should satisfy Condition (i) are given in terms of their Fourier transforms that satisfy (4.15). Note, that for finite sequences β_0 and γ Condition (4.15) only deals with polynomial function on the n -dimensional unit sphere.

Condition (ii) can also be reformulated in terms of function on \mathbb{T}^n , using (4.12). Therefore we introduce

$$B_N = \{\mathcal{R}^{Dk} \beta_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}.$$

Since B_N is generated by N vectors $\beta_0, \dots, \beta_{N-1}$ and the unitary operators $\mathcal{R}_1, \dots, \mathcal{R}_n$, we may use result (4.12). This yields that B_N is a Riesz system if and only if

$$mI_N \leq \widehat{\Gamma}(z) \leq MI_N \text{ a.e. } z \in \mathbb{T}^n,$$

with

$$\widehat{\Gamma}(z)_{k,m} = \sum_{l \in \mathbb{Z}^n} (\beta_m, \mathcal{R}^{Dl} \beta_k)_{\mathbb{Z}^n} z^{-l}, \quad k, m = 0, \dots, N-1. \quad (4.16)$$

This result can also be put in terms of the Fourier transforms of β_l . Therefore we derive the relation

$$\widehat{\Gamma}(z_1^{d_1}, \dots, z_n^{d_n})_{k,m} = 1/L P_{\hat{\beta}_m, \hat{\beta}_k}(z), \quad k, m = 0, \dots, N-1, \quad (4.17)$$

using Lemma 4.4.3, and the fact that

$$\{\sqrt{L} z^{k_1 d_1} \dots z^{k_n d_n} \mid k \in \mathbb{Z}^n\}$$

is an orthonormal basis for $L^2(K_n)$.

Define $\widehat{\Gamma}_d(z) = \widehat{\Gamma}_d(z_1, \dots, z_n) = \widehat{\Gamma}(z_1^{d_1}, \dots, z_n^{d_n})$. Then we arrive at the following theorem by combining (4.12) and (4.17).

Theorem 4.4.4 *Let $N \in \mathbb{N}$ be fixed and $\{\beta_0, \dots, \beta_{N-1}\}$ be a subset of $l^2(\mathbb{Z}^n)$. Let further D be an $(n \times n)$ diagonal matrix with integer entries. Then the collection*

$$B_N = \{\mathcal{R}^{Dk} \beta_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$$

is a Riesz system if and only if the $(N \times N)$ matrix valued function $z \mapsto \widehat{\Gamma}_d(z)$, $z \in K_n$, with entries

$$\widehat{\Gamma}_d(z)_{k,m} = 1/L P_{\hat{\beta}_m, \hat{\beta}_k}(z), \quad k, m = 0, \dots, N-1,$$

admits real positive constants m and M , such that

$$mI_N \leq \widehat{\Gamma}_d(z) \leq MI_N \text{ a.e. } z \in K_n. \quad (4.18)$$

So, Theorem 4.4.4 presents necessary and sufficient conditions on $\hat{\beta}_0, \dots, \hat{\beta}_{N-1}$, so that B_N is a Riesz system. We proceed by searching for similar conditions on the Fourier transforms of $\hat{\beta}_0, \dots, \hat{\beta}_{N-1}$, such that B_N is a Riesz basis for $l^2(\mathbb{Z}^n)$. As a starting point for deriving such conditions we present a corollary of the preceding theorem.

Corollary 4.4.5 *If B_N is a Riesz system, then $N \leq L$.*

Proof

Define the $(L \times N)$ matrix valued function $z \mapsto \widehat{M}(z)$, $z \in K_n$, with entries

$$\widehat{M}(z)_{r,l} = L^{-1/2} \hat{\beta}_l(\omega_{d_1}^{(\pi(r))(1)} z_1, \dots, \omega_{d_n}^{(\pi(r))(n)} z_n), \quad (4.19)$$

$l = 0, \dots, N-1$, $r = 0, \dots, L-1$, where π is an arbitrary bijection from the collection $\{0, 1, \dots, L-1\}$ onto

$$\{r \in \mathbb{Z}^n \mid 0 \leq r_i \leq |d_i| - 1, i = 1, \dots, n\}.$$

Since

$$\widehat{\Gamma}_d(z) = \widehat{M}^*(z)\widehat{M}(z), \quad z \in K_n,$$

$\widehat{\Gamma}_d$ is invertible a.e. if and only if \widehat{M} is injective a.e. If therefore $\widehat{\Gamma}_d$ satisfies (4.18), i.e., $\widehat{\Gamma}_d(z)$ is invertible for almost all $z \in K_d$, then $N \leq L$. \square

For deriving conditions on $\widehat{\beta}_0, \dots, \widehat{\beta}_{N-1}$, such that B_N is a Riesz basis for $l^2(\mathbb{Z}^n)$ we consider the special case $N = L$ and we assume that $\widehat{\Gamma}_d$ satisfies (4.18). Now the proof of Corollary 4.4.5 yields that $\widehat{\Gamma}_d$ is invertible a.e. if and only if \widehat{M} , as introduced in the above proof, is invertible a.e. So (4.18) is equivalent with \widehat{M} being invertible a.e. on K_n . Furthermore, let \mathcal{G} and \mathcal{M} denote bounded linear operators on $L^2(K_n; \mathcal{C}^L)$ corresponding to $\widehat{\Gamma}_d$ and \widehat{M} , respectively, i.e.,

$$(\mathcal{G}\eta)(z) = \widehat{\Gamma}_d(z)\eta(z) \quad \text{and} \quad (\mathcal{M}\eta)(z) = \widehat{M}(z)\eta(z) \quad \text{a.e. } z \in K_n, \quad (4.20)$$

for all $\eta \in L^2(K_n; \mathcal{C}^L)$. Then $\mathcal{G} = \mathcal{M}^*\mathcal{M}$ and $\mathcal{M}^{-1} = \mathcal{G}^{-1}\mathcal{M}^*$. Thus \mathcal{M} is a boundedly invertible operator, since \mathcal{G} is a boundedly invertible operator and

$$(\mathcal{M}^{-1}\eta)(z) = \widehat{M}(z)^{-1}\eta(z).$$

These considerations are used to give conditions on $\widehat{\beta}_0, \dots, \widehat{\beta}_{N-1}$, such that B_N is a Riesz basis for $l^2(\mathbb{Z}^n)$.

Theorem 4.4.6 *Let $N \in \mathbb{N}$ be fixed and $D = \text{diag}(d_1, \dots, d_n)$ with $d_i \in \mathbb{Z}$. Let further $\{\beta_0, \dots, \beta_{N-1}\}$ be a subset of $l^2(\mathbb{Z}^n)$, such that the collection*

$$B_N = \{\mathcal{R}^{Dk}\beta_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$$

is a Riesz system. Then this collection is a Riesz basis if and only if $N = L$.

Proof

Take $N = L$ and let B_L be a Riesz system in $l^2(\mathbb{Z}^n)$. Besides, let \widehat{M} be defined as in (4.19) and let \mathcal{M} be associated with \widehat{M} by (4.20). Then $\widehat{M}(z)$ is invertible for almost all $z \in K_n$ and \mathcal{M} is invertible, since $\widehat{\Gamma}_d$ satisfies (4.18). Define $\tilde{\epsilon}_{l,k} \in L^2(K_n; \mathcal{C}^L)$ for $l = 0, \dots, L-1$, and $k \in \mathbb{Z}^n$ by

$$\tilde{\epsilon}_{l,k}(z) = \sqrt{L} z^{Dk} \epsilon_l, \quad z \in K_n.$$

Furthermore, introduce $\mathcal{V}_D : l^2(\mathbb{Z}^n) \rightarrow L^2(K_n; \mathcal{C}^L)$ by

$$(\mathcal{V}_D h)(z) = \begin{pmatrix} \hat{h}(\omega_{d_1}^{(\pi(0))(1)} z_1, \dots, \omega_{d_n}^{(\pi(0))(n)} z_n) \\ \vdots \\ \hat{h}(\omega_{d_1}^{(\pi(L-1))(1)} z_1, \dots, \omega_{d_n}^{(\pi(L-1))(n)} z_n) \end{pmatrix}$$

where π is an arbitrary bijection from the set $\{0, 1, \dots, L-1\}$ onto the collection

$$E = \{r \in \mathbb{Z}^n \mid 0 \leq r_i \leq |d_i| - 1, i = 1, \dots, n\}.$$

With this definition

$$\begin{aligned} \|\mathcal{V}_D h\|_{L^2(K_n; \mathcal{C}^L)}^2 &= \int_{K_n} \|(\mathcal{V}_D h)(z)\|_{\mathcal{C}^L}^2 d\mu_n(z) \\ &= \int_{K_n} P_{\hat{h}, \hat{h}}(z) d\mu_n(z) = \|h\|_{\mathbb{Z}^n}^2, \end{aligned}$$

for all $h \in l^2(\mathbb{Z}^n)$, so that \mathcal{V}_D is an isometry. Define $h_{l,k} \in l^2$ by

$$\hat{h}_{l,k}(z) = \begin{cases} \sqrt{L} z_1^{d_1 k_1} \dots z_n^{d_n k_n} & \text{if } (\omega_{d_1}^{-l_1} z_1, \dots, \omega_{d_n}^{-l_n} z_n) \in K_n, \\ 0 & \text{otherwise,} \end{cases}$$

with $l \in E, k \in \mathbb{Z}^n$. Then $\{h_{l,k} \mid l \in E, k \in \mathbb{Z}^n\}$ is an orthonormal basis for $l^2(\mathbb{Z}^n)$. This yields $\mathcal{V}_D h_{\pi(l),k} = \tilde{\varepsilon}_{l,k}$ and so the operator \mathcal{V}_D is unitary. Applying \mathcal{V}_D on $\mathcal{R}^{Dk} \beta_l$ now yields

$$\begin{aligned} \mathcal{V}_D(\mathcal{R}^{Dk} \beta_l)(z) &= z^{-Dk} \begin{pmatrix} \hat{\beta}_l(\omega_{d_1}^{(\pi(0))(1)} z_1, \dots, \omega_{d_n}^{(\pi(0))(n)} z_n) \\ \vdots \\ \hat{\beta}_l(\omega_{d_1}^{(\pi(L-1))(1)} z_1, \dots, \omega_{d_n}^{(\pi(L-1))(n)} z_n) \end{pmatrix} \\ &= \sqrt{L} z^{-Dk} \widehat{M}(z) \varepsilon_l = \widehat{M}(z) \tilde{\varepsilon}_{l,-k}(z) = (\mathcal{M} \tilde{\varepsilon}_{l,-k})(z), \end{aligned}$$

for all $l = 0, \dots, L-1$. So $\mathcal{R}^{Dk} \beta_l = (\mathcal{V}_D)^* \mathcal{M} \tilde{\varepsilon}_{l,-k}$.

Since $\{\tilde{\varepsilon}_{l,k} \mid l = 0, \dots, L-1, k \in \mathbb{Z}^n\}$ is an orthonormal basis for $L^2(K_n; \mathcal{C}^L)$ and $(\mathcal{V}_D)^* \mathcal{M}$ is boundedly invertible, it follows that B_N is a Riesz basis for $l^2(\mathbb{Z}^n)$.

For proving the converse, we assume B_L to be a Riesz basis for $l^2(\mathbb{Z}^n)$. Then \mathcal{M} has to be invertible, since $\mathcal{R}^{Dk} \beta_l = \mathcal{V}_D^* \mathcal{M} \tilde{\varepsilon}_{l,k}$. It follows that the matrix valued function \widehat{M} has to be invertible a.e. on K_n , and thus $N = L$. \square

From the preceding theorem we conclude that for satisfying Condition (ii) we have to search for sequences $\beta_1, \dots, \beta_{|\det A|-1}$, given β_0 , such that (4.18) holds.

Although we are not dealing with the concept of frames in $l^2(\mathbb{Z}^n)$ in this section, similar results can now be given in rather straightforward way such that

$$B_N = \{\mathcal{R}^{Dk}\beta_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$$

is a frame. We can write

$$\mathcal{R}^{Dk}\beta_l = \mathcal{V}_D^* \mathcal{M} \tilde{e}_{l,k} = \mathcal{V}_D^* \mathcal{M} \mathcal{U} e_{l,k},$$

with $\mathcal{U} : l^2(\{0, \dots, N-1\} \times \mathbb{Z}^n) \rightarrow L^2(K_n; \mathbb{C}^N)$ the unitary operator given by $\mathcal{U} e_{l,k} = \tilde{e}_{l,k}$. So $\mathcal{S}_F = \mathcal{U}^* \mathcal{M}^* \mathcal{V}_D$ is the frame generator of B_N if B_N is a frame. Now Theorem 4.1.8 immediately yields the following theorem.

Theorem 4.4.7 *Let $N \in \mathbb{N}$ be fixed and $\{\beta_0, \dots, \beta_{N-1}\}$ be a subset of $l^2(\mathbb{Z}^n)$. Then the collection*

$$\{\mathcal{R}^{Dk}\beta_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$$

is a frame if and only if for the $(L \times L)$ matrix valued function $z \mapsto \widehat{M}(z)\widehat{M}^(z)$, $z \in K_n$, with \widehat{M} defined as in (4.19) there exists real positive constants m_F and M_F , such that*

$$m_F I_L \leq \widehat{M}(z)\widehat{M}^*(z) \leq M_F I_L \text{ a.e. } z \in K_n. \quad (4.21)$$

So, we presented necessary and sufficient conditions on $\hat{\beta}_0, \dots, \hat{\beta}_{N-1}$, so that B_N is a frame in a similar way as in Theorem 4.4.4. Finally, we also present a corollary of Theorem 4.4.7, analogous to Corollary 4.4.5.

Corollary 4.4.8 *If B_N is a frame, then $N \geq L$.*

4.5 MRA and Riesz Bases in Hilbert Spaces

In Section 4.2 we used the concept of MRA to construct Riesz bases of the form

$$\{\mathcal{U}_1^j \mathcal{U}_2^k \psi_l \mid l = 1, \dots, N-1, j \in \mathbb{Z}, k \in \mathbb{Z}^n\}$$

for the separable Hilbert space H . We showed that the vectors ψ_l were uniquely determined by (4.9) and that their generating sequences q_l had to be determined such that

- $(\tau_\phi * q_l, \mathcal{R}^{Ak} q_0)_{\mathbb{Z}^n} = 0, \quad \forall l \in \{1, \dots, N-1\} \quad \forall k \in \mathbb{Z}^n,$
- $\{\mathcal{R}^{Ak} q_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$ is a Riesz basis for $l^2(\mathbb{Z}^n)$, with $q_0 = p$, the generating sequence of the MRA generator ϕ .

By taking $\gamma = \tau_\phi$ and $\beta_l = q_l$, $l = 0, \dots, N-1$, in (4.15), Theorem 4.4.4 and Theorem 4.4.6, we arrive at the following theorem on the construction of Riesz bases in Hilbert spaces using MRA.

Theorem 4.5.1 *Given a sequence $q_0 \in l^2(\mathbb{Z}^n)$. Then the following two problems are equivalent.*

Problem 1: Construct sequences q_l , $l = 1, \dots, |\det A| - 1$, in $l^2(\mathbb{Z}^n)$ such that

$$1.1 \quad (\tau_\phi * q_l, \mathcal{R}^{A^k} q_0)_{\mathbb{Z}^n} = 0, \quad \forall l \in \{1, \dots, |\det A| - 1\} \quad \forall k \in \mathbb{Z}^n,$$

$$1.2 \quad \{\mathcal{R}^{A^k} q_l \mid l = 1, \dots, |\det A| - 1, k \in \mathbb{Z}^n\} \text{ is a Riesz basis for } l^2(\mathbb{Z}^n).$$

Problem 2: Construct \hat{q}_l , $l = 1, \dots, |\det A| - 1$, in $L^2(\mathbb{T}^n)$, the n -dimensional discrete Fourier transforms of $q_l \in l^2(\mathbb{Z}^n)$, such that

$$2.1 \quad P_{\hat{\tau}_\phi \hat{q}_l, \hat{q}_0}(z) = 0 \quad \text{a.e. on } K_n, \quad l = 1, \dots, |\det A| - 1, \text{ with } P \text{ as defined in (4.14).}$$

$$2.2 \quad \text{The matrix-valued function } z \mapsto \widehat{M}(z), \quad z \in K_n, \text{ with entries}$$

$$(\widehat{M}(z))_{r,l} = |\det A|^{-1/2} \hat{q}_l(\omega_{d_1}^{(\pi(r))(1)} z_1, \dots, \omega_{d_n}^{(\pi(r))(n)} z_n),$$

$$l, r = 0, \dots, |\det A| - 1, \text{ where } \pi \text{ is an arbitrary enumeration from}$$

$$\{0, 1, \dots, |\det A| - 1\}$$

onto

$$\{m \in \mathbb{Z}^n \mid 0 \leq m_i \leq |d_i| - 1, i = 1, \dots, n\}$$

and with $d_i = D(i, i)$ as in Theorem 4.4.1, is invertible for almost all $z \in K_n$.

Possible solutions to these problems are given in [37, 62] and [74].

To illustrate how to deal with Theorem 4.5.1 we consider an example of an MRA, which we already mentioned in Section 3.2.1, namely an MRA for $L^2(\mathbb{R})$ using Riesz systems. It will turn out that in this example, Problem 2 is not hard to solve. Moreover, for this example Conditions 2.1 and 2.2 will reduce to conditions, which are described thoroughly in the literature [12, 23, 62].

Example 4.5.2 This example deals with an MRA for $L^2(\mathbb{R})$ as introduced in Section 3.2.1. However, here we take an MRA generator $\phi \in L^2(\mathbb{R})$, such that

$$\{\mathcal{T}^k \phi \mid k \in \mathbb{Z}\}$$

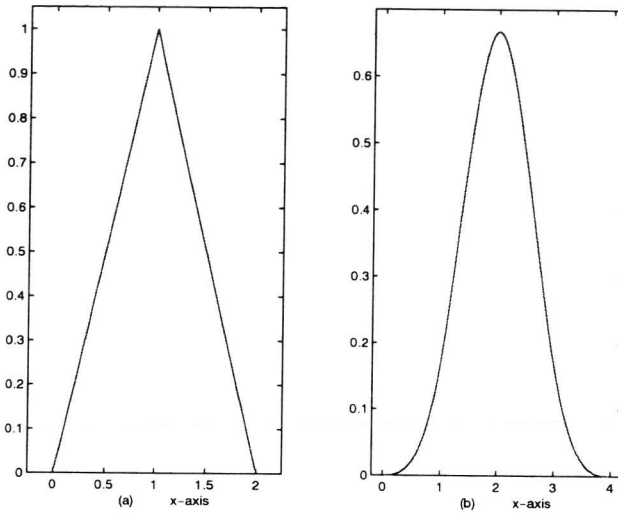


Figure 4.2: Two cardinal B-splines: a) order 2, b) order 4.

is a Riesz system. So, according to Definition 4.2.1 we take $H = L^2(\mathbb{R})$, $\mathcal{U}_1 = \mathcal{D}$ and $\mathcal{U}_{2,1} = \mathcal{T}$. Obviously, Condition (iii) in Definition 4.2.1 holds for $A = 2$. As MRA generator we take $\phi = \phi_m$, the cardinal B-spline of order $m \geq 1$, which is defined by

$$\phi_m = \begin{cases} \chi_{[0,1]}, & m = 1, \\ \phi_1 * \phi_{m-1}, & m \geq 2. \end{cases} \quad (4.22)$$

In Figure 4.2, ϕ_2 and ϕ_4 have been depicted.

For cardinal B-splines we have the following properties

1. $\text{supp } \phi_m = [0, m]$,
2. $\sum_{k \in \mathbb{Z}} \phi_m(x - k) = 1 \quad \forall x \in \mathbb{R}$,
3. $\phi_m(m/2 - x) = \phi_m(m/2 + x) \quad \forall x \in \mathbb{R}$
4. $\{\mathcal{T}^k \phi_m \mid k \in \mathbb{Z}\}$ is a Riesz system,
5. $\phi_m \in \text{clos span } \{\mathcal{D}_a \mathcal{T}^k \phi_m \mid k \in \mathbb{Z}\} \quad \forall a \in \mathbb{N} \setminus \{0\}$.

For a proof of these and other properties of cardinal B-splines we refer to [12, 88].

For solving Problem 2, i.e., to search for a sequence q_1 whose Fourier transform satisfies Conditions 2.1 and 2.2, we have to determine \hat{p} and $\hat{\tau}_\phi$.

By p_m we denote the generating sequence of ϕ_m . To give an expression for p_m , we derive

$$\begin{aligned}\phi_m &= \phi_1 * \phi_{m-1} = \left(\sum_{k \in \mathbb{Z}} p_1(k) \mathcal{DT}^k \phi_1 \right) * \left(\sum_{l \in \mathbb{Z}} p_{m-1}(l) \mathcal{DT}^l \phi_{m-1} \right) \\ &= \sum_{k, l \in \mathbb{Z}} p_1(k) p_{m-1}(l) (\mathcal{DT}^k \phi_1 * \mathcal{DT}^l \phi_{m-1}) \\ &= 1/\sqrt{2} \sum_{k, l \in \mathbb{Z}} p_1(k) p_{m-1}(l) \mathcal{DT}^{k+l} (\phi_1 * \phi_{m-1}) \\ &= 1/\sqrt{2} \sum_{k \in \mathbb{Z}} (p_1 * p_{m-1})(k) \mathcal{DT}^k \phi_m.\end{aligned}$$

Recursively we get $\hat{p}_m(z) = 2^{1/2-m} (\hat{p}_1(z))^m$. For ϕ_1 we find in a straightforward way

$$p_1(z) = (1 + z^{-1})/\sqrt{2},$$

which yields

$$\hat{p}_m(z) = 2^{1/2-m} (1 + z^{-1})^m = 2^{1/2-m} \sum_{k=0}^m \binom{m}{k} z^{-k}.$$

So, the generating sequence p_m of ϕ_m is given by

$$p_m(k) = \begin{cases} 2^{1/2-m} \binom{m}{k}, & k = 0, \dots, m, \\ 0, & \text{otherwise.} \end{cases}$$

Using Property 3 of cardinal B-splines we derive

$$\begin{aligned}\tau_{\phi_m}(k) &= \int_{\mathbb{R}} \phi_m(x) \phi_m(x-k) dx = \int_{\mathbb{R}} \phi_m(x) \phi_m(M+k-x) dx \\ &= (\phi_m * \phi_m)(m+k) = \phi_{2m}(m+k).\end{aligned}$$

Its Fourier transform is given by

$$\hat{\tau}_{\phi_m}(z) = \sum_{k=-m}^m \phi_{2m}(m+k) z^{-k} = \frac{z^{1-m}}{(2m-1)!} E_{2m-1}(z),$$

with E_{2m-1} the Euler-Frobenius polynomial of order $2m-1$, see [12]. We observe that $\hat{\tau}_{\phi_m}(1) = 1$, which follows from Property 2 of cardinal B-splines. Besides, as a property of Euler-Frobenius polynomials we have $E_{2m-1}(z) \neq 0$, $z \in \mathbb{T}$, for all $m \in \mathbb{N} \setminus \{0\}$. These

two considerations yield that $\hat{\tau}_{\phi_m}$ satisfies (4.11). So, indeed $\{\mathcal{T}^k \phi_m \mid k \in \mathbb{Z}\}$ is a Riesz system for all $m \in \mathbb{N} \setminus \{0\}$.

According to Problem 2, we have to search for a $q_m \in l^2(\mathbb{Z})$, such that

- $\hat{\tau}_{\phi_m}(z) \hat{q}_m(z) \overline{\hat{p}_m(z)} + \hat{\tau}_{\phi_m}(-z) \hat{q}_m(-z) \overline{\hat{p}_m(-z)} = 0$ a.e. $z \in \mathbb{T}$,
- $\widehat{M}(z) = \begin{pmatrix} \hat{p}_m(z) & \hat{q}_m(z) \\ \hat{p}_m(-z) & \hat{q}_m(-z) \end{pmatrix}$ is invertible for almost all $z \in \mathbb{T}$.

It can be verified that the first condition holds for

$$\hat{q}_m(z) = z^{2k+1} \hat{\tau}_{\phi_m}(-z) \overline{\hat{p}_m(-z)},$$

$k \in \mathbb{Z}$. With this choice for \hat{q}_m we compute

$$\begin{aligned} |\det(\widehat{M}(z))| &= |\hat{p}_m(z) \hat{q}_m(-z) - \hat{p}_m(-z) \hat{q}_m(z)| \\ &= (\hat{\tau}_{\phi_m}(z) |\hat{p}_m(z)|^2 + \hat{\tau}_{\phi_m}(-z) |\hat{p}_m(-z)|^2) \\ &= \hat{\tau}_{\phi_m}(z) |\hat{p}_m(z)|^2 + \hat{\tau}_{\phi_m}(-z) |\hat{p}_m(-z)|^2 \text{ a.e. } z \in \mathbb{T}. \end{aligned}$$

We already observed that

$$\operatorname{ess\,inf}_{z \in \mathbb{T}} \hat{\tau}_{\phi_m}(z) \geq m_\tau,$$

for a certain positive constant m_τ . Furthermore, we derive

$$\begin{aligned} |\hat{p}_m(z)|^2 + |\hat{p}_m(-z)|^2 &= \sum_{k,l} (1 + (-1)^l) p_m(k) p_m(k-l) z^{-l} \\ &= 2 \sum_{k,l} p_m(k) (\mathcal{R}^{2l} p_m)(k) z^{-2l} \\ &= 2 \widehat{\Gamma}(z^2), \end{aligned}$$

with $\widehat{\Gamma}$ as in (4.16). So,

$$|\hat{p}_m(z)|^2 + |\hat{p}_m(-z)|^2 \geq m_p \text{ a.e. } z \in \mathbb{T},$$

for a certain positive constant m_p . Together these results yield

$$|\det(\widehat{M}(z))| \geq m_\tau m_p > 0.$$

Concluding, q_m and its corresponding wavelet function ψ_m can be obtained from the coefficients of the polynomial

$$\hat{q}_m(z) = \frac{(-2z)^{1-m}}{\sqrt{2}(2m-1)!} (1+z)^m E_{2m-1}(-z).$$

Chapter 5

The FRFT and Affine Transformations in the Wigner Plane

This chapter provides a classification of all unitary operators that act as affine transformations in the multi-dimensional Wigner plane. Moreover, a representation formula is given that encloses all these operators.

The problem of finding these operators is inspired by studying the fractional Fourier transform. This operator, which is introduced in the first section of this chapter, turns out to be acting as a rotation in the Wigner plane. Using a group theoretical approach we arrive at a classification of all linear transformations in the Wigner plane that correspond to unitary operators. This classification is used to come to a representation formula for the corresponding operators on $L^2(\mathbb{R}^n)$. This is done in the second section of this chapter.

This chapter is mainly based on [63].

5.1 The Fractional Fourier Transform

The fractional Fourier transform (FRFT) was introduced by Namias in [68] as a Fourier transform of fractional order. This was done starting from fractional powers of the eigenvalues of the Fourier transform and their corresponding eigenvalues. With this formalism he derived in a heuristic manner an integral representation of this operator. In [53, 61], McBride and Kerr provided a rigorous mathematical framework in which the formal work of Namias could be situated. We discuss this mathematical framework and Namias formal work in the first part

of this section.

Recently, the FRFT turned out to be an interesting transformation for time-frequency signal processing and optical engineering. The growing interest for the FRFT is the consequence of a series of papers that deal with the relation of the FRFT to time-frequency representations of a signal, like the Wigner distribution, see e.g. [4, 67, 77, 78]. This relation is discussed in the second part of this section.

5.1.1 Definition and Properties

We start with the definition of the FRFT for functions in $L^2(\mathbb{R})$.

Definition 5.1.1 Take $f \in L^2(\mathbb{R})$. Its fractional Fourier transform of order $\alpha \in (-\pi, \pi]$ is given by

$$\mathcal{F}_\alpha[f](x) = \frac{C_\alpha}{\sqrt{2\pi} |\sin \alpha|} \int_{\mathbb{R}} f(u) e^{i((u^2+x^2) \cdot (\cot \alpha)/2 - ux \csc \alpha)} du, \quad (5.1)$$

for $0 < |\alpha| < \pi$, with

$$C_\alpha = e^{i(\frac{\pi}{4} \operatorname{sgn} \alpha - \alpha/2)}. \quad (5.2)$$

Furthermore, for $\alpha = 0$ and $\alpha = \pi$ the FRFT is defined by

$$\mathcal{F}_0[f](x) = f(x) \text{ and } \mathcal{F}_\pi[f](x) = f(-x).$$

For $\alpha \notin (\pi, \pi]$ the FRFT is defined by periodicity $\mathcal{F}_{\alpha+2\pi} = \mathcal{F}_\alpha$.

Particularly, we have from this definition

$$\mathcal{F}_{\pi/2} = \mathcal{F} \text{ and } \mathcal{F}_{n\pi/2} = \mathcal{F}^n \quad \forall n \in \mathbb{Z},$$

with \mathcal{F} the Fourier transform on $L^2(\mathbb{R})$.

The factor C_α in (5.2) is chosen to guarantee that \mathcal{F}_α is continuous in α and that \mathcal{F}_α is properly normalized. Indeed, it can be shown that

$$\lim_{\beta \rightarrow \alpha} \|\mathcal{F}_\beta f - \mathcal{F}_\alpha f\|_2 = 0, \quad (5.3)$$

for all $f \in L^2(\mathbb{R})$ and for this particular choice of C_α .

This result is obtained by combining two properties of the FRFT. The first property of the FRFT is known as the index law, i.e.,

$$\mathcal{F}_\alpha \mathcal{F}_\beta f = \mathcal{F}_{\alpha+\beta} f, \quad (5.4)$$

for all $\alpha, \beta \in \mathbb{R}$ and $f \in L^2(\mathbb{R})$. A rigorous proof of this property for functions in the Schwartz space $S(\mathbb{R})$ is given in [61]. Consequently, this result can be extended to functions in $L^2(\mathbb{R})$ since $S(\mathbb{R})$ is dense in $L^2(\mathbb{R})$.

The second property we need for proving the continuity of \mathcal{F}_α is the continuity of the FRFT either in $\alpha = 0$ or $\alpha = \pi$. In [53], it is proven that

$$\lim_{\alpha \rightarrow 0} \|\mathcal{F}_\alpha f - f\|_2 = 0, \quad (5.5)$$

for all $f \in L^2(\mathbb{R})$. Result (5.3) can now be obtained in a straightforward way by combining (5.4) and (5.5). We observe, that (5.3) also holds for other choices of C_α , see e.g. [4].

Considering again (5.4) we have in particular

$$\mathcal{F}_\alpha \mathcal{F}_{-\alpha} = \mathcal{I} \text{ and } \mathcal{F}_{-\alpha} \mathcal{F}_\alpha = \mathcal{I}.$$

It follows that the inverse of \mathcal{F}_α is given by $\mathcal{F}_{-\alpha}$, for all $\alpha \in \mathbb{R}$.

For $t \in \mathbb{R}$, we introduce the unitary operator \mathcal{C}_t on $L^2(\mathbb{R})$ by

$$\mathcal{C}_t[f](x) = e^{itx^2/2} f(x). \quad (5.6)$$

Obviously, \mathcal{C}_t multiplies a given function $f \in L^2(\mathbb{R})$ with a quadratic chirp, i.e., a Fourier mode with a quadratic argument. Using this chirp multiplication and the dilation operator \mathcal{D} as defined in (3.2), we can write \mathcal{F}_α , $\alpha \in (-\pi, \pi)$, also as

$$\mathcal{F}_\alpha f = C_\alpha \mathcal{C}_{\cot \alpha} \mathcal{D}_{\sin \alpha} \mathcal{F} \mathcal{C}_{\cot \alpha}. \quad (5.7)$$

The fact that all operators in the right-hand side of (5.7) are unitary operators on $L^2(\mathbb{R})$ and that $|C_\alpha| = 1$ yields that \mathcal{F}_α is a unitary operator on $L^2(\mathbb{R})$, for all $\alpha \in \mathbb{R}$. Note, that \mathcal{F}_0 and \mathcal{F}_π are also unitary, which follows immediately from Definition 5.1.1. As a consequence we also have Parseval's formula for the FRFT

$$\int_{\mathbb{R}} f(x) \overline{g(x)} dx = \int_{\mathbb{R}} \mathcal{F}_\alpha[f](x) \overline{\mathcal{F}_\alpha[g](x)} dx, \quad (5.8)$$

for all $\alpha \in \mathbb{R}$ and $f, g \in L^2(\mathbb{R})$. Furthermore, as a result we have Plancherel's formula for the FRFT

$$\int_{\mathbb{R}} |f(x)|^2 dx = \int_{\mathbb{R}} |\mathcal{F}_\alpha[f](x)|^2 dx, \quad (5.9)$$

for all $\alpha \in \mathbb{R}$ and $f \in L^2(\mathbb{R})$.

From the preceding derivations and the definition of \mathcal{F}_0 it follows that

$$G_{fr} = \{\mathcal{F}_\alpha \mid \alpha \in \mathbf{R}\}$$

is a strongly continuous subgroup of unitary operators on $L^2(\mathbf{R})$. A cyclic subgroup of order 4 is given by the integer powers of the Fourier transform

$$\{\mathcal{F}^n \mid n = 0, 1, 2, 3\}.$$

Consequently, the discrete cyclic group with generating element \mathcal{F} is embedded in the continuous group G_{fr} .

A further relation with the classical Fourier transform on $L^2(\mathbf{R})$ can be observed by considering the formal derivation of the FRFT by Namias in [68]. His starting point was to consider the eigenvalues and eigenfunctions of the Fourier transform.

It is known, see e.g. [29], that the eigenfunctions of the Fourier transform are given by the Hermite functions

$$h_k(x) = (2^k k! \sqrt{\pi})^{-1/2} e^{-x^2/2} H_k(x), \quad (5.10)$$

where H_k are the Hermite polynomials given by

$$H_k(x) = (-1)^k e^{x^2} \left(\frac{d}{dx} \right)^k e^{-x^2}. \quad (5.11)$$

The Hermite functions form an orthonormal basis for $L^2(\mathbf{R})$ and they satisfy

$$\mathcal{F}h_k = e^{ik\pi/2} h_k.$$

The first idea for the construction of the FRFT was to define an operator \mathcal{F}_α , satisfying

$$\mathcal{F}_\alpha h_k = e^{ik\alpha} h_k, \quad (5.12)$$

for $\alpha \in \mathbf{R}$. For $\alpha = m\pi/2$, with $m \in \mathbf{Z}$, we have $\mathcal{F}_{m\pi/2} = \mathcal{F}^m$. Particularly, if $m \bmod 4 = 0$, then $\mathcal{F}^m = \mathcal{I}$. For a formal representation of \mathcal{F}_α , with $0 < \alpha < \pi/2$, we follow Namias in [68].

We write $f \in L^2(\mathbf{R})$ as $f = \sum_{k=0}^{\infty} (f, h_k)_2 h_k$. Consequently, we have

$$\begin{aligned} \mathcal{F}_\alpha[f](x) &= \sum_{k=0}^{\infty} (f, h_k)_2 \mathcal{F}_\alpha[h_k](x) = \sum_{k=0}^{\infty} (f, h_k)_2 e^{ik\alpha} h_k(x) \\ &= \int_{\mathbf{R}} f(u) \left(\sum_{k=0}^{\infty} e^{ik\alpha} h_k(u) h_k(x) \right) du \\ &= \int_{\mathbf{R}} f(u) \left(\sum_{k=0}^{\infty} \frac{e^{ik\alpha}}{2^k k! \sqrt{\pi}} H_k(u) H_k(x) e^{-u^2/2 - x^2/2} \right) du. \end{aligned}$$

The latter expression can be rewritten using Mehler's formula, see [64],

$$\sum_{k=0}^{\infty} \frac{z^k}{2^k k! \sqrt{\pi}} H_k(u) H_k(x) = \frac{1}{\sqrt{\pi(1-z^2)}} \exp\left(\frac{2xuz - z^2(x^2 + u^2)}{1-z^2}\right). \quad (5.13)$$

We observe that the series converges in L^2 with respect to u , for all x and z , see [29]. Using Mehler's formula in the previous result yields

$$\begin{aligned} \mathcal{F}_\alpha[f](x) &= \frac{1}{\sqrt{\pi e^{i\alpha}} \cdot \sqrt{e^{-i\alpha} - e^{i\alpha}}} \int_{\mathbb{R}} f(u) \exp\left(i \frac{2ixu - i(e^{i\alpha} + e^{-i\alpha})(x^2 + u^2)/2}{e^{i\alpha} - e^{-i\alpha}}\right) du \\ &= \frac{e^{i\pi/4 - i\alpha/2}}{\sqrt{2\pi \sin \alpha}} \int_{\mathbb{R}} f(u) e^{i((u^2 + x^2) \cdot (\cot \alpha)/2 - ux \csc \alpha)} du. \end{aligned}$$

For a rigorous framework in which this formal work of Namias can be studied we refer to [53, 61].

5.1.2 The FRFT and the Wigner Plane

For time-frequency analysis it is interesting to consider the relation of the FRFT with time-frequency operators like the Wigner distribution. Therefore, we compute the Wigner distribution of the FRFT. This will give us insight in how the FRFT acts in phase space.

For this computation we need the following lemma.

Lemma 5.1.2 *Let \mathcal{T}_b and \mathcal{M}_ω , $b, \omega \in \mathbb{R}$, denote the shift operator and frequency modulation on $L^2(\mathbb{R})$ as given in (2.12) and (2.13) respectively. Furthermore, let \mathcal{F}_α , $\alpha \in \mathbb{R}$, the fractional Fourier transform on $L^2(\mathbb{R})$ as given in Definition 5.1.1. Then*

$$\mathcal{F}_\alpha \mathcal{T}_b = e^{ib^2(\sin 2\alpha)/4} \mathcal{M}_{-b \sin \alpha} \mathcal{T}_{b \cos \alpha} \mathcal{F}_\alpha, \quad (5.14)$$

$$\mathcal{F}_\alpha \mathcal{M}_\omega = e^{-i\omega^2(\sin 2\alpha)/4} \mathcal{M}_{\omega \cos \alpha} \mathcal{T}_{\omega \sin \alpha} \mathcal{F}_\alpha. \quad (5.15)$$

Proof

For $\alpha = 0$ both results are trivial, since $\mathcal{F}_0 = \mathcal{I}$. For $\alpha = \pi$ both results follow directly from Definition 5.1.1. Furthermore, equation (5.15) follows from (5.14) by observing that $\mathcal{F} \mathcal{M}_\omega = \mathcal{T}_\omega \mathcal{F}$, with \mathcal{F} the Fourier transform. Indeed, if (5.14) holds, this observation yields

$$\begin{aligned} \mathcal{F}_\alpha \mathcal{M}_\omega &= \mathcal{F}_\alpha \mathcal{F}^* \mathcal{T}_\omega \mathcal{F} = \mathcal{F}_{\alpha-\pi/2} \mathcal{T}_\omega \mathcal{F}_{\pi/2} \\ &= e^{i\omega^2(\sin(2\alpha-\pi))/4} \mathcal{M}_{-\omega \sin(\alpha-\pi/2)} \mathcal{T}_{\omega \cos(\alpha-\pi/2)} \mathcal{F}_{\alpha-\pi/2} \mathcal{F}_{\pi/2} \\ &= e^{-i\omega^2(\sin 2\alpha)/4} \mathcal{M}_{\omega \cos \alpha} \mathcal{T}_{\omega \sin \alpha} \mathcal{F}_\alpha, \end{aligned}$$

using the index law for the FRFT. Consequently, the proof is established by showing that (5.14) holds for $0 < |\alpha| < \pi$. We derive for $f \in L^2(\mathbb{R})$, $b \in \mathbb{R}$ and $0 < |\alpha| < \pi$

$$\begin{aligned}
 \mathcal{F}_\alpha \mathcal{T}_b[f](x) &= \frac{C_\alpha}{\sqrt{2\pi} |\sin \alpha|} \int_{\mathbb{R}} f(u-b) e^{i((u^2+x^2) \cdot (\cot \alpha)/2 - ux \csc \alpha)} du \\
 &= \frac{C_\alpha}{\sqrt{2\pi} |\sin \alpha|} \int_{\mathbb{R}} f(u) e^{i((u^2+x^2+b^2+2ub) \cdot (\cot \alpha)/2 - (u+b)x \csc \alpha)} du \\
 &= \frac{C_\alpha}{\sqrt{2\pi} |\sin \alpha|} e^{i(b^2 \cdot (\cos \alpha)/2 - bx)(1 - \cos^2 \alpha) \csc \alpha} \times \\
 &\quad \int_{\mathbb{R}} f(u) e^{i((u^2+(x-b \cos \alpha)^2)(\cot \alpha)/2 - (u(x-b \cos \alpha)) \csc \alpha)} du \\
 &= e^{i(b^2(\sin 2\alpha)/4 - bx \sin \alpha)} \mathcal{F}_\alpha[f](x - b \cos \alpha) \\
 &= e^{ib^2(\sin 2\alpha)/4} \mathcal{M}_{-b \sin \alpha} \mathcal{T}_{b \cos \alpha} \mathcal{F}_\alpha[f](x).
 \end{aligned}$$

□

Using this lemma, we can compute the action of the FRFT in phase space by means of the Wigner distribution. For this we write

$$\begin{aligned}
 \mathcal{WV}[f](x, \omega) &= \frac{1}{2\pi} \int_{\mathbb{R}} f(x+t/2) \overline{f(x-t/2)} e^{-it\omega} dt \\
 &= \frac{1}{\pi} \int_{\mathbb{R}} f(t+x) \overline{f(x-t)} e^{-2it\omega} dt \\
 &= (\mathcal{M}_{-\omega} \mathcal{T}_{-x} f, \mathcal{M}_\omega \mathcal{T}_x \mathcal{F}_\pi f) / \pi.
 \end{aligned}$$

Using Lemma 5.1.2 we compute

$$\begin{aligned}
 \mathcal{F}_{-\alpha} \mathcal{M}_\omega \mathcal{T}_x &= e^{i(\omega^2 - x^2) \cdot (\sin 2\alpha)/4} \mathcal{M}_{\omega \cos \alpha} \mathcal{T}_{-\omega \sin \alpha} \mathcal{M}_{x \sin \alpha} \mathcal{T}_{x \cos \alpha} \mathcal{F}_{-\alpha} \\
 &= e^{i(\omega^2 - x^2) \cdot (\sin 2\alpha)/4} e^{ix\omega \sin^2 \alpha} \mathcal{M}_{x \sin \alpha + \omega \cos \alpha} \mathcal{T}_{x \cos \alpha - \omega \sin \alpha} \mathcal{F}_{-\alpha}.
 \end{aligned}$$

Combining these two results yields

$$\begin{aligned}
 &\mathcal{WV}[\mathcal{F}_\alpha f](x, \omega) \\
 &= (\mathcal{M}_{-\omega} \mathcal{T}_{-x} \mathcal{F}_\alpha f, \mathcal{M}_\omega \mathcal{T}_x \mathcal{F}_\pi \mathcal{F}_\alpha f) / \pi \\
 &= (\mathcal{F}_{-\alpha} \mathcal{M}_{-\omega} \mathcal{T}_{-x} \mathcal{F}_\alpha f, \mathcal{F}_{-\alpha} \mathcal{M}_\omega \mathcal{T}_x \mathcal{F}_\pi \mathcal{F}_\alpha f) / \pi \\
 &= (\mathcal{M}_{-x \sin \alpha - \omega \cos \alpha} \mathcal{T}_{-x \cos \alpha + \omega \sin \alpha} f, \mathcal{M}_{x \sin \alpha + \omega \cos \alpha} \mathcal{T}_{x \cos \alpha - \omega \sin \alpha} \mathcal{F}_\pi f) / \pi \\
 &= \mathcal{WV}[f](x \cos \alpha - \omega \sin \alpha, x \sin \alpha + \omega \cos \alpha) = \mathcal{WV}[f](R_\alpha(x, \omega)), \quad (5.16)
 \end{aligned}$$

where $R_\alpha(x, \omega)$ represents the matrix vector product with matrix

$$R_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}. \quad (5.17)$$

We conclude from this derivation that the FRFT of order α acts like a rotation by α in the Wigner plane. In particular, we have a rotation by $\pi/2$ in the Wigner plane for $\mathcal{F}_{\pi/2}$. We observe, that this result coincides with the action of the Fourier transform in the Wigner plane as given in (2.32).

The action of the FRFT in the Wigner plane leads us in a natural way to the question which operators on $L^2(\mathbb{R})$ act like a linear transformation in the Wigner plane. The sequel of this chapter is devoted to this question. However, instead of operators on $L^2(\mathbb{R})$ we consider operators acting on $L^2(\mathbb{R}^n)$. It will turn out that finding a solution for the multi-dimensional problem does not follow straightforwardly from the solution for the one-dimensional case.

Since we want to give an answer to our problem for operators on $L^2(\mathbb{R}^n)$, we introduce the fractional Fourier transform on $L^2(\mathbb{R}^n)$ by

$$\mathcal{F}_{\alpha_1, \dots, \alpha_n} = \mathcal{F}_{\alpha_1} \cdots \mathcal{F}_{\alpha_n}, \quad (5.18)$$

for $\alpha_1, \dots, \alpha_n \in \mathbb{R}$. Computing the multi-dimensional Wigner distribution of this FRFT yields

$$\mathcal{WV}[\mathcal{F}_{\alpha_1, \dots, \alpha_n} f](x, \omega) = \mathcal{WV}[f](R_{\alpha_1, \dots, \alpha_n}(x, \omega)), \quad (5.19)$$

with

$$R_{\alpha_1, \dots, \alpha_n} = \begin{pmatrix} \cos \alpha_1 & & 0 & -\sin \alpha_1 & & 0 \\ & \ddots & & & \ddots & \\ 0 & & \cos \alpha_n & & 0 & -\sin \alpha_n \\ \sin \alpha_1 & & 0 & \cos \alpha_1 & & 0 \\ & \ddots & & & \ddots & \\ 0 & & \sin \alpha_n & & 0 & \cos \alpha_n \end{pmatrix} \quad (5.20)$$

This result follows in a straightforward way from (5.16).

5.2 Affine Transformations in the Wigner Plane

Inspired by the fractional Fourier transform and its action in the Wigner plane, we search for linear operators \mathcal{V} on $L^2(\mathbb{R}^n)$ such that there exist a matrix $A \in \mathbb{R}^{n \times n}$ and a vector $b \in \mathbb{R}^n$ for which

$$\mathcal{WV}[\mathcal{V}f](x, \omega) = \mathcal{WV}[f](A(x, \omega) + b), \quad (5.21)$$

holds for all $f \in L^2(\mathbb{R}^n)$. We observe, that De Bruijn already considered this problem in [9] using a new class of generalized functions. Here we will follow an approach based on group theory, see [86, 87, 100]. These results will be placed within the concept of the FRFT in order to embed this transform in a larger class of unitary transformations. Also new results will be added.

We restrict ourselves to matrices A for which $\det A = \pm 1$. For these matrices we have

$$\int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathcal{WV}[f](A[x, \omega] + b) d\omega dx = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathcal{WV}[f](x, \omega) d\omega dx.$$

We shall refer to such affine transformations in the Wigner plane as energy preserving affine transformations. For these transformations the corresponding operators \mathcal{V} on $L^2(\mathbb{R}^n)$ satisfy

$$\begin{aligned} (\mathcal{V}f, \mathcal{V}f) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathcal{WV}[\mathcal{V}f](x, \omega) d\omega dx \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathcal{WV}[f](A(x, \omega) + b) d\omega dx \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \mathcal{WV}[f](x, \omega) d\omega dx = (f, f), \end{aligned}$$

for $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ or $\hat{f} \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ which follows from (2.41). We observe that $L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ is a dense subspace of $L^2(\mathbb{R}^n)$. Concluding, an operator on $L^2(\mathbb{R}^n)$ that yields an energy preserving affine transformation in the Wigner plane has to be an isometry on $L^2(\mathbb{R}^n)$. On the other hand, Equation (5.21) follows directly from applying (2.41) on both sides of the equation $(\mathcal{V}f, \mathcal{V}f) = (f, f)$, for $f \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$ or $\hat{f} \in L^1(\mathbb{R}^n) \cap L^2(\mathbb{R}^n)$.

Before dealing with a classification of all unitary operators that satisfy (5.21), we present some well-known operators for which (5.21) holds.

Multiplication

We start our set of unitary operators on $L^2(\mathbb{R}^n)$ with a trivial one, namely multiplication by a constant C with $|C| = 1$. Result (2.51) already showed that $\mathcal{WV}[f] = \mathcal{WV}[Cf]$, for all $|C| = 1$. Consequently, this multiplication operator satisfies (5.21) with $A = I_{2n}$, the $(2n \times 2n)$ identity matrix, and $b = 0$.

Complex conjugation

Besides linear operators there also exists a non-linear operator for which (5.21) holds, namely the operator $f \mapsto \bar{f}$. For the one-dimensional case we have already seen in (2.30) that

$$\mathcal{WV}[\bar{f}](x, \omega) = \mathcal{WV}[f](x, -\omega).$$

For $f \in L^2(\mathbb{R}^n)$ this result also holds. This follows from a straightforward generalization of (2.30). We conclude, that taking the complex conjugate also satisfies (5.21) with

$$A = \begin{pmatrix} I_n & 0 \\ 0 & -I_n \end{pmatrix} \quad \text{and} \quad b = 0.$$

We observe that we have $\det A = (-1)^n$ for the complex conjugation. Later in this section it will turn out that a necessary condition on a linear operator \mathcal{V} , such that (5.21) holds, is given by $\det A = 1$.

Space and frequency shift

For $x_0, \omega_0 \in \mathbb{R}^n$ we introduce the shift operator and the frequency shift operator on $L^2(\mathbb{R}^n)$ by

$$T_{x_0}[f](x) = f(x - x_0) \quad \text{and} \quad M_{\omega_0}[f](x) = e^{i(\omega_0, x)} f(x)$$

respectively, with $f \in L^2(\mathbb{R}^n)$. Remark, that these operators coincide with the shift and frequency shift operators (2.12) and (2.13) in the one-dimensional case.

We combine the introduced unitary operators T_{x_0} and M_{ω_0} into one unitary operator on $L^2(\mathbb{R}^n)$, given by

$$\mathcal{N}_{(x_0, \omega_0)}[f](x) = T_{x_0} M_{\omega_0}[f](x) = e^{i(\omega_0, x)} f(x - x_0). \quad (5.22)$$

Computing the Wigner transform of this operator yields

$$\mathcal{WV}[\mathcal{N}_{(x_0, \omega_0)} f](x, \omega) = \mathcal{WV}[f](x - x_0, \omega - \omega_0),$$

which is a result we have seen before in discussing the one-dimensional Wigner distribution. From this result we conclude, that (5.21) holds for $\mathcal{N}_{(x_0, \omega_0)}$, namely by taking $A = 0$ and $b = (x_0, \omega_0)$.

We observe that all possible translations $b \in \mathbb{R}^n$ in (5.21) can be obtained from \mathcal{N}_b . This means, that if we are looking for a unitary operator \mathcal{V} on $L^2(\mathbb{R}^n)$ such that (5.21) holds, then we only have to find a linear operator \mathcal{U} on $L^2(\mathbb{R}^n)$ such that

$$\mathcal{WV}[\mathcal{U}f](x, \omega) = \mathcal{WV}[f](A(x, \omega)), \quad (5.23)$$

for all $f \in \mathbb{R}^n$. The operator \mathcal{V} we are looking for is then given by $\mathcal{V} = \mathcal{N}_b \mathcal{U}$. Therefore, we will restrict ourselves from now on to operators \mathcal{U} that satisfy (5.23) with $\det A = \pm 1$.

The Fourier transform

In Section 2.3 we already derived for the Fourier transform \mathcal{F} on $L^2(\mathbb{R})$

$$\mathcal{WV}[\mathcal{F}f](x, \omega) = \mathcal{WV}[f](-\omega, x). \quad (5.24)$$

For $f \in L^2(\mathbb{R}^n)$ and the n -dimensional Fourier transform \mathcal{F} this relation remains the same, which follows straightforwardly from a generalization of Relation (2.31) for the multi-dimensional Wigner distribution. Consequently, the Fourier transform on $L^2(\mathbb{R}^n)$ satisfies (5.23) with $A = J_n^T$. Here J_n denotes the $(2n \times 2n)$ matrix given by

$$J_n = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix}. \quad (5.25)$$

In the sequel of this section this matrix will play an important role in classifying all unitary operators \mathcal{U} that satisfy (5.23).

The dilation operator

For $B \in \mathbb{R}^{n \times n}$, with $\det B \neq 0$, the dilation operator \mathcal{D}_B on $L^2(\mathbb{R}^n)$ is defined by

$$\mathcal{D}_B[f](x) = \frac{1}{\sqrt{|\det B|}} f(B^{-1}x), \quad (5.26)$$

with inverse

$$\mathcal{D}_B^{-1}[f](x) = \sqrt{|\det B|} f(Bx).$$

We use the definition of the Wigner distribution to derive the action of \mathcal{D}_B in the Wigner plane. We compute

$$\begin{aligned} & \mathcal{WV}[\mathcal{D}_B f](x, \omega) \\ &= \frac{1}{(2\pi)^n |\det B|} \int_{\mathbb{R}^n} f(B^{-1}(x + \tau/2)) \overline{f(B^{-1}(x - \tau/2))} e^{-i(\tau, \omega)} d\tau \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} f(B^{-1}x + \tau/2) \overline{f(B^{-1}x - \tau/2)} e^{-i(\tau, B^T \omega)} d\tau \\ &= \mathcal{WV}[f](B^{-1}x, B^T \omega). \end{aligned} \quad (5.27)$$

Concluding, also \mathcal{D}_B corresponds to a linear transformation in the Wigner plane. For \mathcal{D}_B Relation (5.23) holds with

$$A = \begin{pmatrix} B^{-1} & 0 \\ 0 & B^T \end{pmatrix}.$$

Multiplication with a chirp

The last example of a unitary operator that satisfies (5.23) is the operator that multiplies a function in $L^2(\mathbb{R}^n)$ with a quadratic chirp. This operator is given by

$$C_S[f](x) = e^{i(Sx, x)/2} f(x), \quad (5.28)$$

with $S \in \mathbb{R}^{n \times n}$ symmetric. Remark, that we have seen this operator already for the one-dimensional case in (5.6), which coincides with (5.28) for $n = 1$. Obviously its inverse is given by

$$C_S^{-1}[f](x) = e^{-i(Sx, x)/2} f(x).$$

We use (2.52) to derive the action of C_S in the Wigner plane

$$\mathcal{WV}[C_S f](x, \omega) = (2\pi)^{-2n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} ((C_S^* \mu(p, q, 0) C_S) f, f)_2 e^{-i(p, x)} e^{-i(q, \omega)} dp dq.$$

In a direct way we get

$$\begin{aligned} (C_S^* \mu(p, q, 0) C_S)[f](x) &= e^{-i(Sx, x)/2} e^{i(p, x)} e^{i(p, q)/2} e^{i(S(x+q), x+q)/2} f(x+q) \\ &= e^{i(p+Sq, x)} e^{i(p+S(q, q))/2} f(x+q) \\ &= \mu(p+Sq, q, 0)[f](x), \end{aligned}$$

which yields

$$\begin{aligned} \mathcal{WV}[C_S f](x, \omega) &= (2\pi)^{-2n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (\mu(p+Sq, q, 0) f, f)_2 e^{-i(p, x)} e^{-i(q, \omega)} dp dq \\ &= (2\pi)^{-2n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (\mu(p, q, 0) f, f)_2 e^{-i((p, q), A(x, \omega))} dp dq \\ &= \mathcal{WV}[f](A(x, \omega)), \end{aligned} \quad (5.29)$$

with $A = \begin{pmatrix} I_n & 0 \\ -S & I_n \end{pmatrix}$. Consequently, also C_S satisfies (5.23) with A as given before.

5.2.1 A Group Theoretical Approach

In the last example of the previous subsection we have already seen that the relation between a unitary operator on $L^2(\mathbb{R}^n)$ and its affine action in the Wigner plane can be given by using (2.52). This relation can also be used to translate our problem in terms of group theory. This can be done in the following way.

Given a unitary operator \mathcal{V} on $L^2(\mathbb{R}^n)$, we define a unitary representation ρ of the Heisenberg group H_n by $\rho(g) = \mathcal{V}^* \mu(g) \mathcal{V}$, for all $g \in H_n$ and μ the Schrödinger representation. Then by (2.52) we have for such ρ and \mathcal{V}

$$\begin{aligned} \mathcal{W}\mathcal{V}[\mathcal{V}f](x, \omega) &= (2\pi)^{-2n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} ((\mathcal{V}^* \mu(p, q, 0) \mathcal{V})f, f)_2 e^{-i(p, x)} e^{-i(q, \omega)} dp dq. \\ &= (2\pi)^{-2n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (\rho(p, q, 0)f, f)_2 e^{-i(p, x)} e^{-i(q, \omega)} dp dq. \end{aligned}$$

Consequently, if there exists a linear transformation A such that $\mu(g, 0) = \rho(A^T g, 0)$ for all $g \in H'_n$, with

$$H'_n = \{g \in \mathbb{R}^{2n} \mid \forall t \in \mathbb{R} (g, t) \in H_n\},$$

then

$$\begin{aligned} \mathcal{W}\mathcal{V}[\mathcal{V}f](x, \omega) &= (2\pi)^{-2n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} (\mu(A^{-T}(p, q), 0)f, f)_2 e^{-i(p, x)} e^{-i(q, \omega)} dp dq. \\ &= |\det A| \cdot \mathcal{W}\mathcal{V}[f](A(x, \omega)), \end{aligned} \quad (5.30)$$

using the notation $A^{-T} = (A^{-1})^T$.

This derivation shows that the problem we are considering is equivalent to problem of finding operators $\mathcal{V} \in U(L^2(\mathbb{R}^n))$ for which there exist matrices $A \in \mathbb{R}^{n \times n}$ such that

$$\mathcal{V}^* \mu(g, t) \mathcal{V} = \mu(A^{-T} g, t), \quad (5.31)$$

for all $g \in H'_n$ and $t \in \mathbb{R}$.

Besides the Lie groups that have been discussed in Example 2.4.2 we will use another Lie group for solving this problem, namely the symplectic group $Sp(n)$. This group is defined by

$$Sp(n) = \{M \in GL(2n) \mid J_n M^T J_n^T = M^{-1}\}, \quad (5.32)$$

with J_n as given in (5.25). Note that by definition $M^T \in Sp(n)$ and $\det M = \pm 1$ for any $M \in Sp(n)$. Moreover, it can be shown that $Sp(n)$ is connected, see [29]. This yields that $\det M = 1$ if $M \in Sp(n)$. Furthermore, we observe, that $Sp(n) \subset SL(2n)$, but $Sp(1) = SL(2)$. It will turn out later in this section, that this property of the symplectic group causes the fact that solutions for the multi-dimensional problem do not follow straightforwardly from the solution for the one-dimensional case.

To solve our problem we start with the introduction of G , the subgroup of $U(L^2(\mathbb{R}^n))$ given by

$$G = \{\mathcal{V} \in U(L^2(\mathbb{R}^n)) \mid \forall g \in \mathbb{R}^{2n} \forall t \in \mathbb{R} \exists g' \in \mathbb{R}^{2n} : \mathcal{V}^* \mu(g, t) \mathcal{V} = \mu(g', t)\}. \quad (5.33)$$

Obviously, G is a semi-group. Later we will show that every $g \in G$ has an inverse element in G , which yields that G is a group. This group can be equipped with the strong operator topology of $U(L^2(\mathbb{R}^n))$. Furthermore, it is clear from (2.45) that g' in (5.33) is uniquely determined. So a mapping $\nu(\mathcal{V}) : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ can be defined, which depends on $\mathcal{V} \in G$. This $\nu(\mathcal{V})$ is given by $\nu(\mathcal{V})(p, q) = (p', q')$, with p, p', q and q' as in (5.33). Also $\nu(\mathcal{V})$ is a homomorphism for all $\mathcal{V} \in G$. This is shown in the following way.

For $\alpha, \beta \in \mathbb{R}$ and $p_1, p_2, q_1, q_2 \in \mathbb{R}^n$ we have

$$\begin{aligned} & \mathcal{V}^* \mu(\alpha p_1, \alpha q_1, 0) \mu(\beta p_2, \beta q_2, 0) \mathcal{V} \\ &= \mathcal{V}^* \mu(\alpha p_1 + \beta p_2, \alpha q_1 + \beta q_2, (\alpha q_1, \beta p_2)/2 - (\alpha p_1, \beta q_2)/2) \mathcal{V} \\ &= \mu(\nu(\mathcal{V})(\alpha p_1 + \beta p_2, \alpha q_1 + \beta q_2), (\alpha J_n(p_1, q_1), \beta(p_2, q_2))/2). \end{aligned}$$

On the other hand we also have

$$\begin{aligned} & \mathcal{V}^* \mu(\alpha p_1, \alpha q_1, 0) \mu(\beta p_2, \beta q_2, 0) \mathcal{V} \\ &= \mu(\alpha \nu(\mathcal{V})(p_1, q_1), 0) \mu(\beta \nu(\mathcal{V})(p_2, q_2), 0) \\ &= \mu(\alpha \nu(\mathcal{V})(p_1, q_1) + \beta \nu(\mathcal{V})(p_2, q_2), (y_1, x_2)/2 - (x_1, y_2)/2), \end{aligned}$$

with $(x_1, y_1) = \alpha \nu(\mathcal{V})(p_1, q_1)$ and $(x_2, y_2) = \beta \nu(\mathcal{V})(p_2, q_2)$. Taking these results together yields

$$\begin{aligned} & \mu(\nu(\mathcal{V})(\alpha p_1 + \beta p_2, \alpha q_1 + \beta q_2), (\alpha J_n(p_1, q_1), \beta(p_2, q_2))/2) \\ &= \mu(\alpha \nu(\mathcal{V})(p_1, q_1) + \beta \nu(\mathcal{V})(p_2, q_2), (y_1, x_2)/2 - (x_1, y_2)/2). \end{aligned} \quad (5.34)$$

A necessary condition such that (5.34) holds for all $\alpha, \beta, p_1, p_2, q_1$ and q_2 is given by the linearity of $\nu(\mathcal{V})$ for all $\mathcal{V} \in G$. Consequently, $\nu(\mathcal{V}) : \mathbb{R}^{2n} \rightarrow \mathbb{R}^{2n}$ is a homomorphism, that satisfies

$$\mathcal{V}^* \mu(p, q, t) \mathcal{V} = \mu(\nu(\mathcal{V})(p, q), t). \quad (5.35)$$

Using this relation we can show, that $\nu(\mathcal{V})$ is also injective. To do this, we assume $\nu(\mathcal{V})g = 0$, or equivalently $\mu(g, t)\mathcal{V} = \mu(0, t)$. Then

$$\mu(g, t) = \mathcal{V} \mu(0, t) \mathcal{V}^* = \mu(0, t),$$

which yields $g = 0$.

Furthermore, ν satisfies

$$\mu(\nu(C\mathcal{V})(p, q), t) = (\overline{C}\mathcal{V}^*)\mu(p, q, t)(C\mathcal{V}) = \mu(\nu(\mathcal{V})(p, q), t)$$

and

$$\begin{aligned} \mu(\nu(\mathcal{V}_1 \mathcal{V}_2)(p, q), t) &= \mathcal{V}_2^* (\mathcal{V}_1^* \mu(p, q, t) \mathcal{V}_1) \mathcal{V}_2 = \mathcal{V}_2^* \mu(\nu(\mathcal{V}_1)(p, q), t) \mathcal{V}_2 \\ &= \mu(\nu(\mathcal{V}_2) \nu(\mathcal{V}_1)(p, q), t), \end{aligned}$$

for all $\mathcal{V}_1, \mathcal{V}_2 \in U(L^2(\mathbb{R}^n))$ and $|C| = 1$. In the following lemma we deal with some other properties of the mapping ν .

Lemma 5.2.1 *Let G be the subgroup of $U(L^2(\mathbb{R}^n))$ as defined in (5.33) and let ν be the mapping as defined in (5.35). Then ν is a continuous mapping from G onto $Sp(n)$ in the subspace topology of $G \subset U(L^2(\mathbb{R}^n))$. The kernel of ν is given by $\text{Ker } \nu = \{CI \mid |C| = 1\}$.*

Proof

Since p' and q' are uniquely determined in (5.33) it follows that $\nu(\mathcal{V})$ is a non-singular mapping on \mathbb{R}^{2n} , or equivalently $\nu(\mathcal{V}) \in GL(2n)$ for all $\mathcal{V} \in G$. To show that $\nu(\mathcal{V}) \in Sp(n)$, we take $T = \nu(\mathcal{V})$ and $p_1, p_2, q_1, q_2 \in \mathbb{R}^n$. Then by (5.34) we get for $\alpha = 1$ and $\beta = 1$

$$\begin{aligned} & \mu(T(p_1 + p_2, q_1 + q_2), (J_n(p_1, q_1), (p_2, q_2))/2) \\ &= \mu(T(p_1 + p_2, q_1 + q_2), (J_n(x_1, y_1), (x_2, y_2))/2) \\ &= \mu(T(p_1 + p_2, q_1 + q_2), (T^T J_n T(p_1, q_1), (p_2, q_2))/2). \end{aligned}$$

This result must hold for all $p_1, p_2, q_1, q_2 \in \mathbb{R}^n$. This implies that $T^T J_n T J_n^T = I$, which is equivalent with the condition in (5.32).

To compute the kernel of ν we take \mathcal{V} such that $\nu(\mathcal{V}) = I$. This yields $\mathcal{V}\mu\mathcal{V}^* = \mu$. Since μ is irreducible, we get from this equation $\mathcal{V} = CI$, with $|C| = 1$.

To complete this proof we show the continuity of the mapping. Let $\mathcal{V}_1, \mathcal{V}_2 \in G$ and $\mathcal{W} = \mathcal{V}_2 - \mathcal{V}_1$. Then for all $p, q \in \mathbb{R}^n$

$$\begin{aligned} \mu((\nu(\mathcal{V}_2) - \nu(\mathcal{V}_1))(p, q), t) &= \mu(\nu(\mathcal{V}_2)(p, q), 0) \mu(\nu(\mathcal{V}_1)(-p, -q), 0) \\ &= \mathcal{V}_2^* \mu(p, q, 0) (\mathcal{W} + \mathcal{V}_1) \mathcal{V}_1^* \mu(-p, -q, 0) \mathcal{V}_1 \\ &= I - \mathcal{V}_2^* \mathcal{W} + \mathcal{V}_2^* \mu(p, q, 0) \mathcal{W} \mathcal{V}_1^* \mu(-p, -q, 0) \mathcal{V}_1, \end{aligned}$$

with $t = -(\nu(\mathcal{V}_1)^T J_n \nu(\mathcal{V}_2)(p, q), (p, q))$. Consequently,

$$\forall \varepsilon > 0 \exists \delta > 0 \forall p, q \in \mathbb{R}^n : \|\mathcal{V}_2 - \mathcal{V}_1\|_2 < \delta \implies \|\mu((\nu(\mathcal{V}_2) - \nu(\mathcal{V}_1))(p, q), t) - \mu(0, 0, 0)\|_2 < \varepsilon.$$

It can be shown, see e.g. [100], that $\|\mu(p, q, t) - \mu(0, 0, 0)\|_2 \rightarrow 0$ implies $(p, q, t) \rightarrow (0, 0, 0)$. Since the latter result must hold for all $p, q \in \mathbb{R}^n$, we get $\|\nu(\mathcal{V}_2) - \nu(\mathcal{V}_1)\|_2 \rightarrow 0$. This condition is not only necessary to obtain $\|\mu(x, y, t) - \mu(0, 0, 0)\|_2 \rightarrow 0$. It is also sufficient, since $t \rightarrow -(\nu(\mathcal{V}_1)^T J_n \nu(\mathcal{V}_1)(p, q), (p, q)) = -(J_n(p, q), (p, q)) = 0$, if $\nu(\mathcal{V}_2) \rightarrow \nu(\mathcal{V}_1)$. \square

For solving our original problem, namely to find unitary operators on $L^2(\mathbb{R}^n)$ that act like affine transformations in the Wigner plane, we combine (5.30), (5.31) and Lemma 5.2.1. This results into the following theorem.

Theorem 5.2.2 Let \mathcal{V} be a unitary operator on $L^2(\mathbb{R}^n)$ and A a linear transformation on \mathbb{R}^{2n} . Then

$$\mathcal{W}\mathcal{V}[\mathcal{V}f](x, \omega) = \mathcal{W}\mathcal{V}[f](A(x, \omega)). \quad (5.36)$$

if and only if

- (i) $\mathcal{V} \in G$, with G as defined in (5.33),
- (ii) $A \in Sp(n)$,
- (iii) $A = \nu(\mathcal{V})^{-T}$, with ν the continuous mapping from G onto $Sp(n)$ as defined in (5.35).

Theorem 5.2.2 tells us under which conditions unitary operators on $L^2(\mathbb{R}^n)$ act like affine transformations in the Wigner plane, namely if they belong to G . However, Theorem 5.2.2 does not tell us explicitly which unitary operators satisfy (5.36), e.g. by means of a representation formula for such operators. In the following examples we revisit three operators, that have been considered in the beginning of this section. We show that these three operators are elements of G and we compute $\nu(\mathcal{V})$. These three operators will give us some insight in the type of operators, that G consists of. In Section 5.3 we will present a representation formula that gives us an explicit formula for all operators in G .

Example 5.2.3 The first unitary operator we consider is the Fourier transform on $L^2(\mathbb{R}^n)$. We derive

$$\begin{aligned} (\mathcal{F}^* \mu(p, q, t) \mathcal{F})[f](x) &= \int_{\mathbb{R}^n} \hat{f}(\omega + q) e^{i((p, \omega) + (x, \omega) + (p, q)/2 + t)} d\omega \\ &= \int_{\mathbb{R}^n} \hat{f}(\omega) e^{i((p, \omega) + (x, \omega) - (p, q)/2 - (q, x) + t)} d\omega \\ &= e^{i((-q, x) + (-q, p)/2 + t)} \int_{\mathbb{R}^n} \hat{f}(\omega) e^{i(x + p, \omega)} d\omega \\ &= \mu(-q, p, t)[f](x), \end{aligned}$$

for all $f \in L^2(\mathbb{R}^n)$. Consequently, $\mathcal{F} \in G$ and

$$\nu(\mathcal{F}) = J_n^T. \quad (5.37)$$

According to Theorem 5.2.2 the symplectic transformation in the Wigner plane corresponding to the Fourier transform is given by

$$A = \nu(\mathcal{F})^{-T} = (J_n^T)^{-T} = J_n^T,$$

which corresponds with (5.24).

Example 5.2.4 The second unitary operator we consider is the dilation operator \mathcal{D}_B on $L^2(\mathbb{R}^n)$, with $B \in \mathbb{R}^{n \times n}$ and $\det B \neq 0$. We derive

$$\begin{aligned} (\mathcal{D}_B^* \mu(p, q, t) \mathcal{D}_B)[f](x) &= e^{i(p, Bx)} e^{i(t + (p, q)/2)} f(x + B^{-1}q) \\ &= e^{i(B^T p, x)} e^{i(t + (B^T p, B^{-1}q)/2)} f(x + B^{-1}q) \\ &= \mu(B^T p, B^{-1}q, t)[f](x). \end{aligned}$$

this shows that also $\mathcal{D}_B \in G$ for $B \in GL(n)$. Moreover, we have

$$\nu(\mathcal{D}_B) = \begin{pmatrix} B^T & 0 \\ 0 & B^{-1} \end{pmatrix}. \quad (5.38)$$

Now, Theorem 5.2.2 states that the action of the dilation operator in the Wigner plane is given by

$$A = \nu(\mathcal{D}_B)^{-T} = \begin{pmatrix} B^T & 0 \\ 0 & B^{-1} \end{pmatrix}^{-T} = \begin{pmatrix} B^{-1} & 0 \\ 0 & B^T \end{pmatrix}.$$

We observe that this result corresponds to the linear transformation that we derived in (5.27).

Example 5.2.5 The last unitary operator we consider here is the operator \mathcal{C}_S with $S \in \mathbb{R}^{n \times n}$ symmetric, as defined in (5.28). We have already seen

$$(\mathcal{C}_S^* \mu(p, q, t) \mathcal{C}_S)[f](x) = \mu(p + Sq, q, t)[f](x),$$

for $t = 0$. A straightforward computation shows that this result also holds for $t \neq 0$. This result yields that $\mathcal{C}_S \in G$ for $S \in \mathbb{R}^{n \times n}$ symmetric. Furthermore, we have

$$\nu(\mathcal{C}_S) = \begin{pmatrix} I & S \\ 0 & I \end{pmatrix}. \quad (5.39)$$

Theorem 5.2.2 can also be applied to this operator. This yields

$$A = \nu(\mathcal{C}_S)^{-T} = \begin{pmatrix} I & S \\ 0 & I \end{pmatrix}^{-T} = \begin{pmatrix} I & 0 \\ -S & I \end{pmatrix},$$

which is the same result we derived in (5.29).

We observe that the fractional Fourier transform on $L^2(\mathbb{R}^n)$ is a combination of the three unitary operators discussed in the previous examples. We have for $0 < |\alpha_i| < \pi$, $i = 1, \dots, n$,

$$\mathcal{F}_{\alpha_1, \dots, \alpha_n} = C_{\alpha_1} \cdots C_{\alpha_n} \mathcal{C}_{S(\underline{\alpha})} \mathcal{D}_{B(\underline{\alpha})} \mathcal{F} \mathcal{C}_{S(\underline{\alpha})}, \quad (5.40)$$

with

$$S(\underline{\alpha}) = \text{diag}(\cot \alpha_1, \dots, \cot \alpha_n) \text{ and } B(\underline{\alpha}) = \text{diag}(\sin \alpha_1, \dots, \sin \alpha_n).$$

Starting from (5.40) a limit process determines the FRFT with $\alpha_i = 0$ or $\alpha_i = \pi$, for some $i = 1, \dots, N$.

The following theorem classifies all possible elements of $Sp(n)$. A proof of this result can be found in [29, 99].

Theorem 5.2.6 (Bruhat Decomposition) *Let G be the group as defined in (5.33) and let ν be the anti-homomorphism from G onto $Sp(n)$ as defined in (5.35). Then ν is surjective. Moreover, let J_n , $\nu(\mathcal{D}_B)$ and $\nu(\mathcal{C}_S)$ be the real valued $(n \times n)$ matrices as given in (5.25), (5.38) and (5.39) and let*

$$G_1 = \{\nu(\mathcal{C}_S) \mid S \in \mathbb{R}^{n \times n}, S^T = S\}$$

and

$$G_2 = \{\nu(\mathcal{D}_B) \mid B \in \mathbb{R}^{n \times n}, \det B \neq 0\},$$

then $Sp(n)$ is generated by $G_1 \cup G_2 \cup \{J_n\}$.

This result is a corollary of the generalized Bruhat decomposition with respect to a suitable maximal parabolic subgroup [102].

The next corollary combines Theorem 5.2.2 and Theorem 5.2.6. It characterizes all unitary operators on $L^2(\mathbb{R}^n)$ that correspond to linear transformations in the Wigner plane.

Corollary 5.2.7 *Let $f, g \in L^2(\mathbb{R}^n)$. Then*

$$\mathcal{WV}[g](x, \omega) = \mathcal{WV}[f](T(x, \omega)),$$

for some $T \in Sp(n)$ if and only if

$$g = C \mathcal{U}_1 \cdots \mathcal{U}_N f,$$

with $|C| = 1$ and $\mathcal{U}_i = \mathcal{C}_S$, $\mathcal{U}_i = \mathcal{D}_B$ or $\mathcal{U}_i = \mathcal{F}$, with $S \in \mathbb{R}^{n \times n}$ symmetric and $B \in \mathbb{R}^{n \times n}$ non-singular, for $i = 1, \dots, N$, and $N \in \mathbb{N}$.

We omit the proof of this corollary since it follows immediately from Theorem 5.2.2 and Theorem 5.2.6 by observing that $\nu(\mathcal{F})^{-T} = \nu(\mathcal{F})$, $\nu(\mathcal{D}_B)^{-T} = \nu(\mathcal{D}_{B^{-T}})$ and $\nu(\mathcal{C}_S)^{-T} = J_n^T \nu(\mathcal{C}_S) J_n = \nu(\mathcal{F} \mathcal{C}_S \mathcal{F}^*)$.

The classification presented in Corollary 5.2.7 also holds for the mixed Wigner distribution. For a unitary operator \mathcal{V} on $L^2(\mathbb{R}^n)$ that corresponds to a linear transformation A in the Wigner plane we also have

$$\mathcal{WV}[\mathcal{V}f, \mathcal{V}g](x, \omega) = \mathcal{WV}[f, g](A(x, \omega)), \quad (5.41)$$

with $A \in Sp(n)$. This relation holds by polarization, i.e.,

$$\begin{aligned}
 & \mathcal{WV}[\mathcal{V}f, \mathcal{V}g](x, \omega) \\
 &= (\mathcal{WV}[\mathcal{V}f](x, \omega) + \mathcal{WV}[\mathcal{V}g](x, \omega) - \mathcal{WV}[\mathcal{V}(f+g)](x, \omega)) / 2 \\
 &= (\mathcal{WV}[f](A(x, \omega)) + \mathcal{WV}[g](A(x, \omega)) - \mathcal{WV}[f+g](A(x, \omega))) / 2 \\
 &= \mathcal{WV}[f, g](A(x, \omega)),
 \end{aligned}$$

for real-valued $f, g \in L^2(\mathbb{R}^n)$. For complex-valued functions we have to deal with the real and complex part separately.

In Section 5.3 this relation is used to come to a representation formula for the unitary operators as discussed in Corollary 5.2.7.

5.2.2 The FRFT Generalized

As we have seen in (5.40) the fractional Fourier transform on $L^2(\mathbb{R}^n)$ can be decomposed into four unitary operators, namely a chirp multiplication, the Fourier transform, a dilation and again a chirp multiplication. Both the chirp multiplications and the dilation depend on a set of parameters $\alpha_1, \dots, \alpha_n$, that determine the FRFT. Therefore, a natural generalization of the FRFT is given by

$$\mathcal{F}_{\Gamma, \Delta} = C \mathcal{C}_{\Gamma} \mathcal{D}_{\Delta} \mathcal{F} \mathcal{C}_{\Gamma}, \quad (5.42)$$

for some $|C| = 1$, $\Gamma, \Delta \in \mathbb{R}^{n \times n}$, both symmetric and with Δ non-singular. We observe, that Δ is not required to be symmetric in (5.26). Here we require the symmetry of Δ to obtain a symmetrical representation formula for the generalized FRFT.

We observe, that (5.42) generalizes the multi-dimensional FRFT, which was introduced in Section 5.1.2. Indeed, by taking

$$\Gamma = \text{diag}(\cot \alpha_1, \dots, \cot \alpha_n) \text{ and } \Delta = \text{diag}(\sin \alpha_1, \dots, \sin \alpha_n) \quad (5.43)$$

the generalized FRFT with the definition of the multi-dimensional FRFT.

As a consequence of Corollary 5.2.7, we have for all operators $\mathcal{F}_{\Gamma, \Delta}$

$$\mathcal{WV}[\mathcal{F}_{\Gamma, \Delta} f](x, \omega) = \mathcal{WV}[f](A(x, \omega)),$$

for some $A \in Sp(n)$. Using (5.37), (5.38) and (5.39) we compute straightforwardly

$$A = \nu(\mathcal{C}_{\Gamma} \mathcal{D}_{\Delta} \mathcal{F} \mathcal{C}_{\Gamma})^{-T} = \begin{pmatrix} \Delta \Gamma & -\Delta \\ -\Gamma \Delta \Gamma + \Delta^{-1} & \Gamma \Delta \end{pmatrix}. \quad (5.44)$$

Taking Γ and Δ as in (5.43) we arrive at the matrix A as given in (5.20).

A special property of the FRFT is that for its corresponding transformation in the Wigner plane we have $A \in Sp(n) \cap SO(2n)$, the orthonormal symplectic group. One may ask whether the generalized FRFT is also related to an orthogonal transformation in the Wigner plane. The answer to this question is given in the following lemma.

Lemma 5.2.8 *Let $\mathcal{F}_{\Gamma, \Delta}$ be the generalized FRFT as defined in (5.42), for certain symmetric real valued $(n \times n)$ matrices Γ and Δ . Then A as given by (5.44) is orthogonal if and only if*

- (i) $\Delta^{-2} - \Gamma^2 = I$,
- (ii) $\Gamma\Delta^{-1}$ is symmetric.

Proof

We compute

$$A^T A = \begin{pmatrix} X & Y \\ Y^T & Z \end{pmatrix},$$

with

$$\begin{aligned} X &= \Gamma\Delta\Gamma - \Gamma\Delta\Gamma^2\Delta\Gamma + \Delta^{-2} - \Delta^{-1}\Gamma\Delta\Gamma - \Gamma\Delta\Gamma\Delta^{-1}, \\ Y &= \Delta^{-1}\Gamma\Delta - \Gamma\Delta^2 - \Gamma\Delta\Gamma^2\Delta, \\ Z &= \Delta + \Delta\Gamma^2\Delta. \end{aligned}$$

For orthonormal A we should have $X = Z = I$ and $Y = 0$. The condition $Z = I$ yields $\Delta^{-1}Z\Delta^{-1} = \Delta^{-2}$, which equals (i). Obviously, Condition (i) is also sufficient to guarantee $Z = I$. Substituting (i) into the matrix Y yields

$$Y = 0 \iff \Gamma\Delta^{-1} = \Delta^{-1}\Gamma \iff \Gamma\Delta^{-1} = (\Gamma\Delta^{-1})^T.$$

After substituting Condition (i) and (ii) in the matrix X we get $X = I$. So for the equation $X = I$ no further conditions are required. \square

We observe that Conditions (i) and (ii) in Lemma 5.2.8 are equivalent with

$$(\Delta^{-1} + \Gamma)(\Delta^{-1} - \Gamma) = I.$$

It follows from this relation, that we have $n^2/2 + n$ degrees of freedom for choosing symmetric matrices Γ and Δ , such that the matrix A corresponding to $\mathcal{F}_{\Gamma, \Delta}$ is orthogonal. Therefore, for higher dimensional function spaces we may expect more variety in the class of operators $\mathcal{F}_{\Gamma, \Delta}$ that yield orthogonal symplectic transformations in the Wigner plane. For the one-dimensional case the one-parameter family of the FRFT turns out to be the only transformation up to a constant, that is in the class of generalized FRFT and that acts like an orthogonal transform in the Wigner plane.

Lemma 5.2.9 Let $\mathcal{F}_{\Gamma, \Delta}$ be the unitary operator on $L^2(\mathbb{R})$ as given in (5.42), with $\Gamma, \Delta \in \mathbb{R}$. Then $A = \nu(\mathcal{F}_{\Gamma, \Delta})^{-T}$ is orthonormal if and only if $\mathcal{F}_{\Gamma, \Delta} = C \mathcal{F}_\alpha$, for some $\alpha \in \mathbb{R}$ and C with $|C| = 1$.

Proof

In the case that Γ and Δ are scalars, the conditions in Lemma 5.2.8 reduce to

$$\Delta^{-2} = 1 + \Gamma^2.$$

This equation can be parameterized by taking $\Gamma = \cot \alpha$ and $\Delta = \sin \alpha$, for some $\alpha \in \mathbb{R}$. Substituting this parameterization into (5.42) leaves the FRFT \mathcal{F}_α up to a constant of absolute value 1, which does not affect A . \square

As we expected from the considerations before Lemma 5.2.9, this lemma cannot be extended in a canonical way to higher dimensions. This is shown by the following example for $n = 2$. Moreover, by extending the example to higher dimensions in a natural way it follows that the preceding lemma can only hold for $\mathcal{F}_{\Gamma, \Delta} \in U(L^2(\mathbb{R}))$.

Example 5.2.10 We consider $\mathcal{F}_{\Gamma, \Delta}$ on $L^2(\mathbb{R}^2)$, with

$$\Gamma = \begin{pmatrix} r_1^2 \cos^2 \alpha + r_2^2 \sin^2 \alpha & (r_1 - r_2) \cos \alpha \sin \alpha \\ (r_1 - r_2) \cos \alpha \sin \alpha & r_1^2 \sin^2 \alpha + r_2^2 \cos^2 \alpha \end{pmatrix} \quad \text{and}$$

$$\Delta = \begin{pmatrix} \rho_1^2 \cos^2 \alpha + \rho_2^2 \sin^2 \alpha & (\rho_1 - \rho_2) \cos \alpha \sin \alpha \\ (\rho_1 - \rho_2) \cos \alpha \sin \alpha & \rho_1^2 \sin^2 \alpha + \rho_2^2 \cos^2 \alpha \end{pmatrix}^{-1},$$

with $\alpha \in \mathbb{R}$ and $\rho_i^2 = 1 + r_i^2$, $i = 1, 2$. Then

$$\Delta^{-2} - \Gamma^2 = \begin{pmatrix} \rho_1^2 - r_1^2 & 0 \\ 0 & \rho_2^2 - r_2^2 \end{pmatrix} = I,$$

and

$$\Gamma \Delta^{-1} = \begin{pmatrix} r_1^2 \rho_1^2 \cos^2 \alpha + r_2^2 \rho_2^2 \sin^2 \alpha & (r_1 \rho_1 - r_2 \rho_2) \cos \alpha \sin \alpha \\ (r_1 \rho_1 - r_2 \rho_2) \cos \alpha \sin \alpha & r_1^2 \rho_1^2 \sin^2 \alpha + r_2^2 \rho_2^2 \cos^2 \alpha \end{pmatrix} = (\Gamma \Delta^{-1})^T.$$

Consequently, the matrices Γ and Δ satisfy the conditions in Lemma 5.2.8. The orthogonal symplectic transformation in the Wigner plane, that corresponds to $\mathcal{F}_{\Gamma, \Delta}$ is now given by $A = U(\alpha)^T M U(\alpha)$, with

$$M = \begin{pmatrix} -r_1/\rho_1 & 0 & -1/\rho_1 & 0 \\ 0 & -r_2/\rho_2 & 0 & -1/\rho_2 \\ 1/\rho_1 & 0 & -r_1/\rho_1 & 0 \\ 0 & 1/\rho_2 & 0 & -r_2/\rho_2 \end{pmatrix}$$

and

$$U(\alpha) = \begin{pmatrix} \cos \alpha & \sin \alpha & \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha & -\sin \alpha & \cos \alpha \\ \cos \alpha & \sin \alpha & \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha & -\sin \alpha & \cos \alpha \end{pmatrix}.$$

Resuming, we have extended the FRFT to a unitary transformation on $L^2(\mathbb{R}^n)$ given by $\mathcal{F}_{\Gamma, \Delta}$, where $\Gamma, \Delta \in \mathbb{R}^{n \times n}$, both symmetric and Δ non-singular. So the set of all generalizations of the FRFT on $L^2(\mathbb{R}^n)$ of this kind are given by the set

$$V_n = \{\mathcal{F}_{\Gamma, \Delta} \mid \Gamma, \Delta \in \mathbb{R}^{n \times n} \text{ symmetric, } \det \Delta \neq 0\}.$$

Furthermore, a subset of V_n is defined consisting of all $\mathcal{F}_{\Gamma, \Delta} \in V_n$ that act like orthogonal transformations in the Wigner plane. This subset is given by

$$W_n = \{\mathcal{F}_{\Gamma, \Delta} \in V_n \mid \Delta^{-2} - \Gamma^2 = I, \Gamma\Delta = (\Gamma\Delta)^T\}.$$

For the FRFT we have $\mathcal{F}_{\alpha_1, \dots, \alpha_n} \in W_n \subset V_n$. Moreover, for the one-dimensional case we have

$$W_1 = \{C \mathcal{F}_\alpha \mid \alpha \in \mathbb{R}, |C| = 1\}$$

and

$$W_n \setminus \{\mu \mathcal{F}_{\alpha_1, \dots, \alpha_n} \mid \alpha_1, \dots, \alpha_n \in \mathbb{R}, |\mu| = 1\} \neq \emptyset,$$

for $n \geq 2$.

5.3 A Representation Formula

In this section we present a representation formula for all unitary operators \mathcal{V} on $L^2(\mathbb{R}^n)$ for which there exists a transformation A on \mathbb{R}^{2n} such that

$$\mathcal{W}\mathcal{V}[\mathcal{V}f, \mathcal{V}g](x, \omega) = \mathcal{W}\mathcal{V}[f, g](A(x, \omega)). \quad (5.45)$$

We observe, that for the particular choice $f = g$, (5.45) coincides with (5.23). We have already shown that (5.45) can only be realized for symplectic transformations A . Therefore, we start with some properties of symplectic matrices.

Given a matrix $A \in Sp(n)$, then we can represent A by its 2×2 block decomposition

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}. \quad (5.46)$$

Since A is symplectic, it has to satisfy (5.32). This yields for the block decomposition

$$A^{-1} = \begin{pmatrix} A_{22}^T & -A_{12}^T \\ -A_{21}^T & A_{11}^T \end{pmatrix}, \quad (5.47)$$

or equivalently

$$A_{22}^T A_{11} - A_{12}^T A_{21} = I, \quad (5.48)$$

$$A_{11}^T A_{21} - A_{21}^T A_{11} = 0, \quad (5.49)$$

$$A_{22}^T A_{12} - A_{12}^T A_{22} = 0. \quad (5.50)$$

Using these relation we prove the following less known properties of symplectic matrices.

Lemma 5.3.1 *Let $A \in Sp(n)$ be given by its 2×2 block decomposition (5.46). Then the following relations hold*

$$(i) (A_{22}^T)^{\leftarrow}(\text{Ran}(A_{12}^T)) = \text{Ran}(A_{12}),$$

$$(ii) \dim A_{22}(\text{Ker}(A_{12})) = \dim \text{Ker}(A_{12}),$$

$$(iii) A_{22}(\text{Ker}(A_{12})) = (\text{Ran}(A_{12}))^{\perp},$$

with $\text{Ker}(B)$ and $\text{Ran}(B)$ denoting respectively the null space and range of a linear transformation B and with $B^{\leftarrow}(W)$ denoting the inverse image of a subspace W under the linear transformation B .

Proof

Let $v \in (A_{22}^T)^{\leftarrow}(\text{Ran}(A_{12}^T))$. Then there exists an $u \in \mathbb{R}^n$ such that $A_{22}^T v + A_{12}^T u = 0$. Hence,

$$A^T \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} A_{11}^T & A_{21}^T \\ A_{12}^T & A_{22}^T \end{pmatrix} \begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} A_{11}^T u + A_{21}^T v \\ 0 \end{pmatrix}.$$

Since A is symplectic, we can apply (5.47). This yields

$$\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} A_{22} & -A_{21} \\ -A_{12} & A_{11} \end{pmatrix} \begin{pmatrix} A_{21}^T v + A_{11}^T u \\ 0 \end{pmatrix}.$$

Consequently, $v = -A_{12}(A_{21}^T v + A_{11}^T u) \in \text{Ran}(A_{12})$. On the other hand, if $v \in \text{Ran}(A_{12})$, then there exists a $w \in \mathbb{R}^n$ such that $v = A_{12} w$. Using (5.50) we derive

$$A_{22}^T v = A_{22}^T A_{12} w = A_{12}^T A_{22} w \in \text{Ran}(A_{12}^T),$$

which proves Property (i).

In order to prove (ii), it is sufficient to show that, if $A_{22} u = 0$, for $u \in \text{Ker}(A_{12})$, then $u = 0$. Using (5.48), this follows from

$$u = I_n u = A_{11}^T A_{22} u - A_{21}^T A_{12} u = 0.$$

In [63] we have shown that, given a linear transformation B in \mathbb{R}^n and a linear subspace V in \mathbb{R}^n , we have

$$\dim(B^T(V^\perp)) = \dim V^\perp \implies B^T(V^\perp) = (B^\leftarrow(V))^\perp.$$

Now, replacing B by A_{22}^T and V by $\text{Ran}(A_{12}^T)$ yields

$$(A_{22}(\text{Ker}(A_{12})))^\perp = (A_{22}^T)^\leftarrow(\text{Ran}(A_{12}^T)) = \text{Ran}(A_{12}),$$

which proves Property (iii). \square

For deriving a representation formula we also need the following result.

Lemma 5.3.2 *Let W be a subspace of \mathbb{R}^n and let B be a linear transformation on \mathbb{R}^n , such that $\dim(B(W)) = \dim(W) = d$. Then*

$$\int_W f(Bx) dx = \frac{1}{q_W(B)} \int_{B(W)} f(x) dx, \quad \forall f \in S(\mathbb{R}^n), \quad (5.51)$$

with $q_W(B)$ the d -dimensional volume of the simplex generated by $B e_1, \dots, B e_d$, with e_1, \dots, e_d an orthonormal basis in W .

The proof of this lemma is omitted, since it is straightforward. We observe, that $q_W(B)$ is positive. Furthermore, if W is the null space and B is non-singular, then by setting $q_W(B) = 1$ the definition of $q_W(B)$ is extended in a consistent way.

The last lemma we need to derive our representation formula is as follows.

Lemma 5.3.3 *Let $f \in S(\mathbb{R}^n)$ and $A \in Sp(n)$ with block decomposition (5.46). Furthermore, let $\dim \text{Ran}(A_{12}) = d > 0$. Then,*

$$\int_{\text{Ker}(A_{12})} \int_{\mathbb{R}^n} f(u) e^{i(v, A_{22}^T u)} du dv = \frac{(2\pi)^{n-d}}{q_{\text{Ker}(A_{12})}(A_{22})} \int_{\text{Ran}(A_{12})} f(v) dv. \quad (5.52)$$

Proof

Since $\dim A_{22}(\text{Ker}(A_{12})) = \dim \text{Ker}(A_{12}) = n - d$, cf. Property (ii) of Lemma 5.3.1, we may apply Lemma 5.3.2. This yields

$$\begin{aligned} \int_{\text{Ker}(A_{12})} \left(\int_{\mathbb{R}^n} f(u) e^{i(v, A_{22}^T u)} du \right) dv &= (2\pi)^{n/2} \int_{\text{Ker}(A_{12})} \hat{f}(A_{22} v) dv = \\ &= \frac{(2\pi)^{n/2}}{q_{\text{Ker}(A_{12})}(A_{22})} \int_{A_{22}(\text{Ker}(A_{12}))} \hat{f}(v) dv. \end{aligned} \quad (5.53)$$

for all $f \in S(\mathbb{R}^n)$ and linear subspaces W of \mathbb{R}^n . By taking $W = A_{22}(\text{Ker}(A_{12}))$ this result becomes

$$\int_{A_{22}(\text{Ker}(A_{12}))} \hat{f}(v) dv = (2\pi)^{n/2-d} \int_{A_{22}(\text{Ker}(A_{12}))^\perp} f(v) dv.$$

Since $A_{22}(\text{Ker}(A_{12}))^\perp = \text{Ran}(A_{12})$, we have, cf. Property (iii) of Lemma 5.3.1,

$$\int_{A_{22}(\text{Ker}(A_{12}))} \hat{f}(v) dv = (2\pi)^{n/2-d} \int_{\text{Ran}(A_{12})} f(v) dv.$$

In combination with (5.53) the latter result establishes the proof. \square

The starting point for the derivation of our representation formula is the characteristic function of the Wigner distribution (2.35). For the n -dimensional mixed Wigner distribution, we can also define a characteristic function by

$$M[f, g](\theta, t) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(u + t/2) \overline{g(u - t/2)} e^{i(u, \theta)} du,$$

or equivalently

$$M[f, g](\theta, t) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(u + t) \overline{g(u)} e^{i(u+t/2, \theta)} du, \quad (5.54)$$

with $f, g \in L^2(\mathbb{R}^n)$. By the inverse Fourier transform we have

$$f(x) \overline{g(y)} = (2\pi)^{-n/2} \int_{\mathbb{R}^n} M[f, g](\theta, x - y) e^{-i(\theta, x+y)/2} d\theta. \quad (5.55)$$

For the n -dimensional mixed Wigner distribution we have

$$\mathcal{WV}[f](x, \omega) = (2\pi)^{-3n/2} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} M[f](\theta, t) e^{-i(\theta, x)} e^{-i(t, \omega)} d\theta dt. \quad (5.56)$$

Now, let \mathcal{V} be a unitary operator satisfying (5.45). It follows from (5.56) together with (5.45) that

$$M[\mathcal{V}f, \mathcal{V}g] = M[f, g] \circ (A^{-1})^T. \quad (5.57)$$

Combining (5.57) with (5.47) and (5.55) we arrive at

$$\mathcal{V}[f](x) \overline{\mathcal{V}[g](y)} = (2\pi)^{-n/2} \int_{\mathbb{R}^n} M[f, g]((A^{-1})^T(\theta, x - y)) e^{-i(\theta, x+y)/2} d\theta$$

Combining (5.57) with (5.47) and (5.55) we arrive at

$$\begin{aligned} \mathcal{V}[f](x) \overline{\mathcal{V}[g](y)} &= (2\pi)^{-n/2} \int_{\mathbb{R}^n} M[f, g]((A^{-1})^T(\theta, x - y)) e^{-i(\theta, x+y)/2} d\theta \\ &= (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(u - A_{12}\theta/2 + A_{11}(x - y)/2) \times \\ &\quad \overline{g(u + A_{12}\theta/2 - A_{11}(x - y)/2)} E_0(u, \theta, x, y) du d\theta, \text{ a.e..} \end{aligned}$$

for all f and g in $L^2(\mathbb{R}^n)$, with

$$E_0(u, \theta, x, y) = \exp(i(A_{22}\theta - A_{21}(x - y), u) - i(\theta, x + y)/2).$$

This last relation only holds formally for general $f, g \in L^2(\mathbb{R}^n)$, but it holds rigorously for $f, g \in S(\mathbb{R}^n)$. Therefore, we assume $f, g \in S(\mathbb{R}^n)$ from now on. After this derivation, we will show that the representation formula also hold for $f \in L^2(\mathbb{R}^n)$.

By taking $v = u - A_{11}(x + y)/2$ in the previous result, we have

$$\begin{aligned} \mathcal{V}[f](x) \overline{\mathcal{V}[g](y)} &= \\ (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(v - A_{12}\theta/2 + A_{11}x) \overline{g(v + A_{12}\theta/2 + A_{11}y)} \times \\ &\quad \exp(i E_1(v, \theta, x, y)) dv d\theta, \end{aligned}$$

with $E_1(v, \theta, x, y) = (A_{22}\theta - A_{21}(x - y), v + A_{11}(x + y)/2) - (\theta, x + y)/2$. Using Relations (5.48) - (5.50), we can write E_1 as

$$\begin{aligned} E_1(v, \theta, x, y) &= (A_{22}\theta - A_{21}(x - y), v) + (A_{12}\theta, A_{21}(x + y))/2 - \\ &\quad (A_{21}x, A_{11}x)/2 + (A_{21}y, A_{11}y)/2. \end{aligned}$$

Hence, $\mathcal{V}[f](x) \overline{\mathcal{V}[g](y)}$ can be rewritten as

$$\mathcal{V}[f](x) \overline{\mathcal{V}[g](y)} = e^{-i(A_{21}x, A_{11}x)/2} e^{i(A_{21}y, A_{11}y)/2} \mathcal{H}[f, g](x, y), \quad (5.58)$$

with

$$\begin{aligned} \mathcal{H}[f, g](x, y) &= (2\pi)^{-n} \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} f(v - A_{12}\theta/2 + A_{11}x) \overline{g(v + A_{12}\theta/2 + A_{11}y)} \\ &\quad \times e^{i(A_{12}\theta, A_{21}(x+y))/2} e^{i(A_{22}\theta - A_{21}(x-y), v)} dv d\theta. \end{aligned}$$

Our aim is now to write \mathcal{H} in a possible degenerate form. If this is established, then the representation formula for $\mathcal{V}f$ can be read off from this form. To come to such a form we

We are now in the position to apply Lemma 5.3.3 with respect to the function

$$v \mapsto f(v - A_{12} \theta_1/2 + A_{11} x) \overline{g(v + A_{12} \theta_1/2 + A_{11} y)} e^{i(A_{22} \theta_1 - A_{21}(x-y), v)}.$$

By applying this lemma, we arrive at

$$\begin{aligned} \mathcal{H}[f, g](x, y) = & \frac{(2\pi)^{-d}}{q_{\text{Ker}(A_{12})}(A_{22})} \int_{\text{Ran}(A_{12}^T)} \int_{\text{Ran}(A_{12})} f(v - A_{12} \theta_1/2 + A_{11} x) \times \\ & \overline{g(v + A_{12} \theta_1/2 + A_{11} y)} \times \\ & e^{i((A_{12} \theta_1, A_{21}(x+y))/2 + (A_{22} \theta_1 - A_{21}(x-y), v))} dv d\theta_1, \end{aligned}$$

with $d = \dim \text{Ran}(A_{12})$. Since $v \in \text{Ran}(A_{12})$, we may substitute $v = A_{12} w$ with $w \in \text{Ran}(A_{12}^T)$, since A_{12} restricted to $\text{Ran}(A_{12}^T)$ is a linear bijection onto $\text{Ran}(A_{12})$. We obtain

$$\begin{aligned} \mathcal{H}[f, g](x, y) = & C_A^2 \int_{\text{Ran}(A_{12}^T)} \int_{\text{Ran}(A_{12}^T)} f(A_{12} w - A_{12} \theta_1/2 + A_{11} x) \times \\ & \overline{g(A_{12} w + A_{12} \theta_1/2 + A_{11} y)} \times \\ & e^{i((A_{12} \theta_1, A_{21}(x+y))/2 + (A_{22} \theta_1 - A_{21}(x-y), A_{12} w))} dw d\theta_1, \end{aligned}$$

with

$$C_A = \sqrt{\frac{s(A_{12})}{(2\pi)^d q_{\text{Ker}(A_{12})}(A_{22})}}. \quad (5.59)$$

Here $s(A_{12})$ denotes the product of the nonzero singular values of A_{12} , or equivalently

$$s(A_{12}) = q_{\text{Ran}(A_{12}^T)}(A_{12}).$$

Our next step is to substitute $t_1 = w - \theta_1/2$ and $t_2 = w + \theta_1/2$. Then, by using (5.48) - (5.50) one has

$$\begin{aligned} (A_{12} \theta_1, A_{21}(x+y))/2 + (A_{22} \theta_1 - A_{21}(x-y), A_{12} w) = & \\ (A_{12}(t_2 - t_1), A_{21}(x+y))/2 + (A_{22}(t_2 - t_1), A_{12}(t_1 + t_2))/2 - & \\ (A_{21}(x-y), A_{12}(t_1 + t_2))/2 = & \\ -(A_{22} t_1, A_{12} t_1)/2 + (A_{22} t_2, A_{12} t_2)/2 - (A_{12} t_1, A_{21} x) + (A_{21} y, A_{12} t_2). & \end{aligned}$$

With this result we can rewrite $\mathcal{H}[f, g](x, y)$ in the degenerate form

$$\mathcal{H}[f, g](x, y) = C_A^2 \mathcal{H}_0[f](x) \overline{\mathcal{H}_0[g](y)}, \quad (5.60)$$

$$\begin{aligned} & (A_{12}(t_2 - t_1), A_{21}(x + y))/2 + (A_{22}(t_2 - t_1), A_{12}(t_1 + t_2))/2 - \\ & (A_{21}(x - y), A_{12}(t_1 + t_2))/2 = \\ & -(A_{22}t_1, A_{12}t_1)/2 + (A_{22}t_2, A_{12}t_2)/2 - (A_{12}t_1, A_{21}x) + (A_{21}y, A_{12}t_2). \end{aligned}$$

With this result we can rewrite $\mathcal{H}[f, g](x, y)$ in the degenerate form

$$\mathcal{H}[f, g](x, y) = C_A^2 \mathcal{H}_0[f](x) \overline{\mathcal{H}_0[g](y)}, \quad (5.60)$$

with

$$\mathcal{H}_0[f](x) = \int_{\text{Ran}(A_{12}^T)} f(A_{12}t + A_{11}x) e^{-i((A_{22}t, A_{12}t)/2 + (A_{12}t, A_{21}x))} dt.$$

Finally, combining (5.58) and (5.60) yields the degenerate form for $\mathcal{V}[f](x) \overline{\mathcal{V}[g](y)}$

$$\mathcal{V}[f](x) \overline{\mathcal{V}[g](y)} = C_A^2 \mathcal{H}_0[f](x) \overline{\mathcal{H}_0[g](y)}. \quad (5.61)$$

In a natural way this derivation results into the definition of an operator \mathcal{F}_A that satisfies (5.45). We will define this operator on $L^2(\mathbb{R}^n)$ and show that it indeed corresponds to the unitary operator we have been searching for.

Definition 5.3.4 Let $A \in Sp(n)$ with block decomposition (5.46). Then the linear operator \mathcal{F}_A on $L^2(\mathbb{R}^n)$ is defined as follows. If $\dim(\text{Ran}(A_{12})) > 0$, then

$$\begin{aligned} \mathcal{F}_A[f](x) &= C_A e^{-i(A_{11}^T A_{21} x, x)/2} \times \\ &\int_{\text{Ran}(A_{12}^T)} f(A_{12}t + A_{11}x) e^{-i(A_{12}^T A_{22} t, t)/2 - i(t, A_{12}^T A_{21} x)} dt, \end{aligned} \quad (5.62)$$

for all $f \in L^2(\mathbb{R}^n)$ and with C_A as given in (5.59). Furthermore, if $\dim(\text{Ran}(A_{12})) = 0$ then

$$\mathcal{F}_A[f](x) = \sqrt{|\det A_{11}|} e^{-i(A_{11}^T A_{21} x, x)/2} f(A_{11}x), \quad (5.63)$$

for all $f \in L^2(\mathbb{R}^n)$.

The main theorem of this section can be stated as follows.

Theorem 5.3.5 Let $A \in Sp(n)$ and \mathcal{F}_A be given as in Definition 5.3.4. Then

$$\mathcal{WV}[\mathcal{F}_A f, \mathcal{F}_A g](x, \omega) = \mathcal{W}[f, g](A(x, \omega)),$$

for all $f, g \in L^2(\mathbb{R}^n)$.

for all $f \in L^2(\mathbb{R}^n)$. The proof for $\dim(\text{Ran}(A_{12})) > 0$ is completed by assuming, that \mathcal{V} satisfies (5.45).

If $\dim(\text{Ran}(A_{12})) = 0$, we have $A_{12} = 0$. Then (5.48) and (5.49) yield, that A_{11} is non-singular and that $A_{11}^{-1} = A_{22}^T$. Moreover, $A_{11}^T A_{21}$ is symmetric. Using these observations, we compute the mixed Wigner distribution of $\mathcal{F}_A f$ and $\mathcal{F}_A g$ as follows.

$$\begin{aligned} W[\mathcal{F}_A f, \mathcal{F}_A g](x, \omega) &= \frac{|\det A_{11}|}{(2\pi)^n} \int_{\mathbb{R}^n} f(A_{11}x + A_{11}t/2) \times \\ &\quad \overline{g(A_{11}x - A_{11}t/2)} e^{-i(A_{11}^T A_{21}x, t)} e^{-i(t, \omega)} dt = \\ &= (2\pi)^{-n} \int_{\mathbb{R}^n} f(A_{11}x + t/2) \overline{g(A_{11}x - t/2)} e^{-i((A_{11}^T A_{21}x, A_{11}^{-1}t) + (A_{11}^{-1}t, \omega))} dt. \end{aligned}$$

Hence,

$$\mathcal{WV}[\mathcal{F}_A f, \mathcal{F}_A g](x, \omega) = \mathcal{WV}[f, g](A_{11}x, A_{21}x + A_{22}\omega).$$

This establishes the proof for $\dim(\text{Ran}(A_{12})) = 0$. \square

At the end of this section, we present two well-known examples of unitary operators, that satisfy (5.45).

Example 5.3.6 We recall, that for a set of parameters $\alpha_1, \dots, \alpha_n \in (0, \pi)$ the n -dimensional fractional Fourier transform is given by

$$\mathcal{F}_{\alpha_1, \dots, \alpha_n}[f](x) = \frac{C_{\underline{\alpha}} e^{i(Bx, x)/2}}{\sqrt{(2\pi)^n |\sin \alpha_1 \cdots \sin \alpha_n|}} \int_{\mathbb{R}^n} f(u) e^{i((Bu, u)/2 - (Cx, u))} du, \quad (5.64)$$

with $B = \text{diag}(\cot \alpha_1, \dots, \cot \alpha_n)$, $C = \text{diag}(\csc \alpha_1, \dots, \csc \alpha_n)$ and $C_{\underline{\alpha}} = C_{\alpha_1} \cdots C_{\alpha_n}$, where C_{α_n} is given by (5.2). The symplectic matrix, that corresponds to this transform in the Wigner plane is given by the rotation matrix $R_{\alpha_1, \dots, \alpha_n}$ as given in (5.20). We observe, that in this particular case A_{12} is non-singular. This yields $q_{\text{Ker}(A_{12})}(A_{22}) = 1$ and $s(A_{12}) = \det(A_{12})$. Using these simplifications and the substitution $u = A_{12}t + A_{11}x$, Formula (5.62) simplifies to

$$\mathcal{F}_A[f](x) = \frac{e^{-i(A_{12}^{-1} A_{11}x, x)/2}}{(2\pi)^{n/2} \sqrt{|\det A_{12}|}} \int_{\mathbb{R}^n} f(u) e^{-i((A_{22} A_{12}^{-1} u, u)/2 - (x, A_{12}^{-1} u))} du.$$

Taking $A_{11} = A_{22} = \text{diag}(\cos \alpha_1, \dots, \cos \alpha_n)$ and $A_{12} = \text{diag}(-\sin \alpha_1, \dots, -\sin \alpha_n)$, the latter representation formula turns into the n -dimensional FRFT as given in (5.64).

$\det(A_{12})$. Using these simplifications and the substitution $u = A_{12}t + A_{11}x$, Formula (5.62) simplifies to

$$\mathcal{F}_A[f](x) = \frac{e^{-i(A_{12}^{-1} A_{11} x, x)/2}}{(2\pi)^{n/2} \sqrt{|\det A_{12}|}} \int_{\mathbb{R}^n} f(u) e^{-i((A_{22} A_{12}^{-1} u, u)/2 - (x, A_{12}^{-1} u))} du.$$

Taking $A_{11} = A_{22} = \text{diag}(\cos \alpha_1, \dots, \cos \alpha_n)$ and $A_{12} = \text{diag}(-\sin \alpha_1, \dots, -\sin \alpha_n)$, the latter representation formula turns into the n -dimensional FRFT as given in (5.64).

Example 5.3.7 The second example is the unitary operator on $L^2(\mathbb{R}^2)$, which corresponds in the Wigner plane to the symplectic matrix

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

Remark, that all matrices in the block decomposition of A are singular.

It can be verified in a straightforward way, that $q_{\text{Ker}(A_{12})}(A_{22}) = 1$ and $s(A_{12}) = 1$. By substituting the block matrices of A into (5.62), the unitary operator, we are dealing with, reads

$$\mathcal{F}_A[f](x_1, x_2) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} f(x_1, \xi) e^{-i\xi x_2} d\xi,$$

which is the one-dimensional Fourier-transform of $f(x_1, \cdot)$. We observe, that this operator can also be derived from (5.64) by taking $\alpha_1 \rightarrow 0$ and $\alpha_2 \rightarrow \pi/2$.

We observe that in [29] and [33] also a representation formula is presented for unitary operators that correspond to symplectic transformations in the Wigner plane. However, both references do not give a formula that can also handle symplectic transformations with a block decomposition, that consists of four singular block matrices, which is the case in the second example.

Chapter 6

Localization Problems in Phase Space

A celebrated problem in signal processing is the problem of maximizing energy in both time and frequency. This problem already has received much attention in the literature, see e.g. [20, 28, 41, 58]. This chapter consists of two parts that are also devoted to this problem.

In the first part we discuss two classical problems. The first problem concerns the maximization of energy of time-limited signals within a frequency band, i.e. finite interval in the Fourier domain. For this problem we revisit a series of papers by Slepian and co-workers, [57, 80, 93]. Furthermore, we give a rigorous proof of a conjecture by Slepian [92]. The second problem concerns the maximization of energy within a disk in the Wigner plane, i.e. the phase space related to the Wigner distribution. Although this problem is discussed in several papers [20, 28, 29, 46], we also present alternative proofs and additional results in this section.

The second part of this chapter is devoted to the FRFT, which we generalized in the previous chapter. Using this generalization we are able to relate several classes of energy maximization problems in phase space to the two classical problems as discussed in the first part of this chapter. For this, we discuss the Weyl correspondence, see e.g. [29, 107].

This chapter is based on [73, 75].

6.1 Slepian's Energy Problem

The first problem to be considered in this part of the chapter is the concentration of energy in a certain frequency band of a time-limited signal. So we consider for time limited signals f the ratio

$$E_f(\omega_0) = \frac{\int_{-\omega_0}^{\omega_0} |\hat{f}(\omega)|^2 d\omega}{\int_{\mathbb{R}} |\hat{f}(\omega)|^2 d\omega}, \quad (6.1)$$

with $[-\omega_0, \omega_0]$ the frequency band we are looking at in this problem. Obviously, $E_f(\omega_0) \geq 0$, for all $f \in L^2(\mathbb{R})$. Moreover, Corollary 2.1.13 yields $E_f(\omega_0) < 1$.

Since $E_f(\omega_0) < 1$ for all $f \in L^2(\mathbb{R})$, the problem arises of maximizing this energy ratio over all $f \in L^2([-x_0, x_0])$, for some fixed $x_0 > 0$.

For solving this problem we introduce two operators. The first operator we discuss is the integral operator $\mathcal{B}(\omega_0) : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$. For $\omega_0 > 0$ fixed, this operator is given by

$$\mathcal{B}(\omega_0)[f](x) = \sqrt{\frac{2}{\pi}} \int_{\mathbb{R}} \frac{\sin(\omega_0(x-u))}{(x-u)} f(u) du, \quad (6.2)$$

for all $f \in L^2(\mathbb{R})$. We observe that

$$\mathcal{F}^{-1}[\chi_{[-\omega_0, \omega_0]}](x) = \sqrt{\frac{2}{\pi}} \frac{\sin(\omega_0 x)}{x}.$$

According to Lemma 2.1.8 the latter result yields

$$\mathcal{F}\mathcal{B}(\omega_0)f = \chi_{[-\omega_0, \omega_0]} \cdot \mathcal{F}f \text{ a.e. on } \mathbb{R}. \quad (6.3)$$

Hence $\mathcal{B}(\omega_0)$ is a Hermitian projection operator; in fact it is an orthoprojector.

The second operator we introduce in relation to the energy localization problem is the projection $\mathcal{P}(x_0) : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$. For $x_0 > 0$ fixed, this operator is defined by

$$\mathcal{P}(x_0)[f](x) = \begin{cases} f(x), & \text{if } |x| \leq x_0, \\ 0, & \text{if } |x| > x_0. \end{cases} \quad (6.4)$$

By combining the introduced operators we arrive at

$$\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)[f](x) = \sqrt{\frac{2}{\pi}} \int_{-x_0}^{x_0} \frac{\sin(\omega_0(x-u))}{(x-u)} f(u) du, \quad |x| \leq x_0, \quad (6.5)$$

for all $f \in L^2(\mathbb{R})$. Since the integral kernel in (6.5) is in $L^2([-x_0, x_0]^2)$, we have that $\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)$ is a Hilbert-Schmidt operator. Hence, $\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)$ is a compact operator. Also $\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)$ is a positive definite operator on $L^2([-x_0, x_0])$, which is shown as follows. Using (6.3) we derive

$$\begin{aligned} (\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)f, f)_2 &= (\mathcal{B}(\omega_0)\mathcal{P}(x_0)f, \mathcal{P}(x_0)f)_2 \\ &= (\mathcal{F}\mathcal{B}(\omega_0)\mathcal{P}(x_0)f, \mathcal{F}\mathcal{P}(x_0)f)_2 \\ &= (\chi_{[-\omega_0, \omega_0]} \cdot \mathcal{F}\mathcal{P}(x_0)f, \mathcal{F}\mathcal{P}(x_0)f)_2 \\ &= (\chi_{[-\omega_0, \omega_0]} \cdot \mathcal{F}\mathcal{P}(x_0)f, \chi_{[-\omega_0, \omega_0]} \cdot \mathcal{F}\mathcal{P}(x_0)f)_2. \end{aligned}$$

We assume

$$(\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)f, f)_2 = 0,$$

for some $f \in L^2(\mathbb{R})$. Then $\mathcal{F}\mathcal{P}(x_0)[f](\omega) = 0$, for almost all $\omega \in [-\omega_0, \omega_0]$. However, $\mathcal{F}\mathcal{P}(x_0)f$ is holomorphic by Theorem 2.1.12. This yields in combination with the latter result $\mathcal{F}\mathcal{P}(x_0)f = 0$, or equivalently $f(x) = 0$ for almost all $|x| < x_0$. For $f \in L^2([-x_0, x_0])$ this yields

$$(\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)f, f)_2 = 0 \implies f = 0 \text{ a.e. on } \mathbb{R}.$$

Following Pollack and Slepian [79, 91], we consider possible solutions $\mathcal{P}(x_0)f_{\max}$, with $f_{\max} \in L^2(\mathbb{R})$, that maximize (6.1). Then

$$E_{\mathcal{P}(x_0)f_{\max}}(\omega_0) \cdot (\mathcal{F}\mathcal{P}(x_0)f_{\max}, \mathcal{F}\mathcal{P}(x_0)f_{\max})_2 = (\chi_{[-\omega_0, \omega_0]} \cdot \mathcal{F}\mathcal{P}(x_0)f_{\max}, \mathcal{F}\mathcal{P}(x_0)f_{\max})_2.$$

Equivalently, using Parseval's theorem and (6.3),

$$E_{\mathcal{P}(x_0)f_{\max}}(\omega_0) \cdot (\mathcal{P}(x_0)f_{\max}, \mathcal{P}(x_0)f_{\max})_2 = (\mathcal{B}(\omega_0)\mathcal{P}(x_0)f_{\max}, \mathcal{P}(x_0)f_{\max})_2.$$

Since f_{\max} is a stationary solution of this equation, it must satisfy

$$\mathcal{B}(\omega_0)\mathcal{P}(x_0)f_{\max} = \lambda\mathcal{P}(x_0)f_{\max}, \quad (6.6)$$

a homogeneous Fredholm equation of the first kind.

We recall that $\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)$ is compact. Furthermore, it is a positive definite operator on $L^2([-x_0, x_0])$. These considerations yield that solutions $\mathcal{P}(x_0)f$ for equation (6.6) only exist for a discrete set of real positive values of λ , with the properties that

$$1 > \lambda_0 > \lambda_1 > \lambda_2 > \dots$$

and $\lim_{k \rightarrow \infty} \lambda_k = 0$. In general, the eigenvalues of a compact Hermitian operator are not necessarily distinct. However, for this particular Fredholm operator, Pollack and Slepian have shown in [79], that its eigenvalues are distinct. Also Slepian showed, see [91], that the kernel of the integral operator $\mathcal{B}(\omega_0)$ commutes with the second order differential operator

$$\mathcal{D}(x_0\omega_0) = \frac{d}{dx}(1-x^2)\frac{d}{dx} - (x_0\omega_0)^2x^2. \quad (6.7)$$

Since both operators have the same spectrum, they must have the same eigenvectors.

Differential operator (6.7) is a well-known operator. It arises on separating the 3-dimensional scalar wave equation in a prolate spheroidal coordinate system. Its real-valued eigenfunctions $\psi_0, \psi_1, \psi_2, \dots$ are known as prolate spheroidal wave functions (PSWF), see [27]. We observe, that the concentration of energy problem is solved by $\mathcal{P}(x_0)\psi_0$.

Some useful properties of the PSWF have been derived in the past. We present some of them in the following lemma. For a proof of these properties we refer to [57, 80, 93].

Lemma 6.1.1 *Let $\psi_0, \psi_1, \psi_2, \dots$ be the eigenfunctions of $\mathcal{P}(x_0)^* \mathcal{B}(\omega_0) \mathcal{P}(x_0)$ and let their corresponding eigenvalues be given by $\lambda_0, \lambda_1, \lambda_2, \dots$. Then*

- (i) $\hat{\psi}_k \in L^2([-\omega_0, \omega_0]) \quad \forall k \in \mathbb{N}$,
- (ii) $\int_{-x_0}^{x_0} \psi_k(x) \psi_n(x) dx = \lambda_k \delta_{k,n}$,
- (iii) $\int_{\mathbb{R}} \psi_k(x) \psi_n(x) dx = \delta_{k,n}$.

Other properties for the PSWF follow from this lemma, e.g. Theorem 2.1.12 and (i) yield that ψ_k is holomorphic. However, this lemma does not provide us with an explicit expression for ψ_k and consequently for λ_k . More insight in the behaviour of the eigenvalues λ_k is given by a conjecture of Slepian, which can be proven rigorously by using the following classical result, that is due to Landau and Widom, see [58].

Lemma 6.1.2 *Let $\mathcal{H}(x_0\omega_0) : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R})$ be given by*

$$\mathcal{H}(x_0\omega_0) = \mathcal{P}(x_0\omega_0) \mathcal{B}(1) \mathcal{P}(x_0\omega_0).$$

Furthermore, let $N(\mathcal{H}(x_0\omega_0), p)$, $0 < p < 1$, denote the number of eigenvalues of $\mathcal{H}(x_0\omega_0)$ which are greater than or equal to p . Then

$$N(\mathcal{H}(x_0\omega_0), p) = \frac{2x_0\omega_0}{\pi} + \frac{1}{\pi^2} \log\left(\frac{1-p}{p}\right) \log(x_0\omega_0) + R(x_0\omega_0), \quad (6.8)$$

with $R(x)$ of order $o(\log(x))$ as $x \rightarrow \infty$. Using this result we prove Slepian's conjecture in a rigorous way.

Theorem 6.1.3 (Slepian's conjecture) *Let $\mathcal{P}(x_0) \mathcal{B}(\omega_0) \mathcal{P}(x_0)$ be as defined in (6.5) and let λ_k , $k \in \mathbb{N}$, be its eigenvalues. Then for all $\delta, \epsilon \in (0, 1)$ there exists an $M \in \mathbb{N}$ such that*

- (i) $\lambda_k < \epsilon$, if $k \geq (1 + \delta) \frac{2x_0\omega_0}{\pi}$, for $x_0\omega_0 > M$,
- (ii) $1 - \lambda_k < \epsilon$, if $1 \leq k \leq (1 - \delta) \frac{2x_0\omega_0}{\pi}$, for $x_0\omega_0 > M$.

Moreover, for all $\varepsilon > 0$ and $\theta \in \mathbb{R}$, there exist $\delta > 0$ and $M \in \mathbb{N}$ such that

$$(iii) \quad |\lambda_k - (1 + e^{\pi\theta})^{-1}| < \varepsilon, \text{ for } |k - \frac{2x_0\omega_0}{\pi} - \frac{\theta}{\pi} \log(x_0\omega_0)| < \delta \log(x_0\omega_0) \text{ and } x_0\omega_0 > M.$$

Proof

We define $\phi_k(x) = \psi_k(x/\omega_0)$. Then, for $|x| < x_0\omega_0$, we derive

$$\begin{aligned} \lambda_k \phi_k(x) &= \sqrt{\frac{2}{\pi}} \int_{-x_0}^{x_0} \frac{\omega_0 \sin(x - u\omega_0)}{(x - u\omega_0)} \phi_k(u\omega_0) du \\ &= \sqrt{\frac{2}{\pi}} \int_{-x_0\omega_0}^{x_0\omega_0} \frac{\sin(x - v)}{(x - v)} \phi_k(v) dv \end{aligned}$$

or equivalently

$$\mathcal{H}(x_0\omega_0)\phi_k = \lambda_k \phi_k \quad \forall k \in \mathbb{N} \setminus \{0\}.$$

Consequently, Lemma 6.1.2 can also be applied on the eigenvalues of $\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)$.

Let $0 < \varepsilon < 1$ and $0 < \delta < 1$. We take $M > 0$ such that

$$\delta > \frac{\log\left(\frac{1-\varepsilon}{\varepsilon}\right) \log x}{2\pi x} + \frac{\pi R(x)}{2x}, \quad (6.9)$$

for $x > M$. Then

$$\begin{aligned} N(\mathcal{H}(x_0\omega_0), \varepsilon) &= \frac{2x_0\omega_0}{\pi} + \frac{1}{\pi^2} \log\left(\frac{1-\varepsilon}{\varepsilon}\right) \log(x_0\omega_0) + R(x_0\omega_0) \\ &< (1 + \delta) \frac{2x_0\omega_0}{\pi}, \end{aligned}$$

for $x_0\omega_0 > M$. Consequently, if $k \geq (1 + \delta) \frac{2x_0\omega_0}{\pi}$, then $N(\mathcal{H}(x_0\omega_0), \varepsilon) < k$. This result yields $\lambda_k < \varepsilon$.

For proving Property (ii) we also take $M > 0$ such that (6.9) holds. Then

$$\begin{aligned} N(\mathcal{H}(x_0\omega_0), 1 - \varepsilon) &= \frac{2x_0\omega_0}{\pi} - \frac{1}{\pi^2} \log\left(\frac{1-\varepsilon}{\varepsilon}\right) \log(x_0\omega_0) + R(x_0\omega_0) \\ &> (1 - \delta) \frac{2x_0\omega_0}{\pi}, \end{aligned}$$

for $x_0\omega_0 > M$. Therefore, if $1 \leq k \leq (1 - \delta) \frac{2x_0\omega_0}{\pi}$, then $N(\mathcal{H}(x_0\omega_0), 1 - \varepsilon) > k$, which leads to $1 - \lambda_k < \varepsilon$.

Finally, let $\varepsilon > 0$ and $\theta \in \mathbb{R}$. Furthermore, take $\delta > 0$ and $M \in \mathbb{N}$ such that

$$\delta < \frac{1}{\pi^2} \log\left(\frac{1 + \varepsilon + \varepsilon e^{\pi\theta}}{1 - \varepsilon - \varepsilon e^{-\pi\theta}}\right) - \frac{R(x)}{\log x},$$

for $x > M$. Then we have

$$\begin{aligned}
 & N(\mathcal{H}(x_0\omega_0), (1 + e^{\pi\theta})^{-1} + \varepsilon) = \\
 & \frac{2x_0\omega_0}{\pi} + \frac{1}{\pi^2} \log \left(e^{\pi\theta} \cdot \frac{1 - \varepsilon - \varepsilon e^{-\pi\theta}}{1 + \varepsilon + \varepsilon e^{\pi\theta}} \right) \log(x_0\omega_0) + R(x_0\omega_0) = \\
 & \frac{2x_0\omega_0}{\pi} + \frac{\theta}{\pi} \log(x_0\omega_0) - \\
 & \frac{1}{\pi^2} \log \left(\frac{1 + \varepsilon + \varepsilon e^{\pi\theta}}{1 - \varepsilon - \varepsilon e^{-\pi\theta}} \right) \log(x_0\omega_0) + R(x_0\omega_0) < \\
 & \frac{2x_0\omega_0}{\pi} + \frac{\theta}{\pi} \log(x_0\omega_0) - \delta \log(x_0\omega_0)
 \end{aligned}$$

for $x_0\omega_0 > M$. Consequently, if

$$k > \frac{2x_0\omega_0}{\pi} + \frac{\theta}{\pi} \log(x_0\omega_0) - \delta \log(x_0\omega_0),$$

or equivalently, if

$$\frac{2x_0\omega_0}{\pi} + \frac{\theta}{\pi} \log(x_0\omega_0) - k < \delta \log(x_0\omega_0),$$

then $\lambda_k - (1 + e^{\pi\theta})^{-1} < \varepsilon$.

In the same way, we derive

$$\begin{aligned}
 & N(\mathcal{H}(x_0\omega_0), (1 + e^{\pi\theta})^{-1} - \varepsilon) = \\
 & \frac{2x_0\omega_0}{\pi} + \frac{\theta}{\pi} \log(x_0\omega_0) + \\
 & \frac{1}{\pi^2} \log \left(\frac{1 + \varepsilon + \varepsilon e^{\pi\theta}}{1 - \varepsilon - \varepsilon e^{-\pi\theta}} \right) \log(x_0\omega_0) + R(x_0\omega_0) > \\
 & \frac{2x_0\omega_0}{\pi} + \frac{\theta}{\pi} \log(x_0\omega_0) + \delta \log(x_0\omega_0)
 \end{aligned}$$

for $x_0\omega_0 > M$. Therefore, if

$$k < \frac{2x_0\omega_0}{\pi} + \frac{\theta}{\pi} \log(x_0\omega_0) + \delta \log(x_0\omega_0),$$

or equivalently, if

$$k - \frac{2x_0\omega_0}{\pi} - \frac{\theta}{\pi} \log(x_0\omega_0) < \delta \log(x_0\omega_0),$$

then $\lambda_k - (1 + e^{\pi\theta})^{-1} > -\varepsilon$. Combining these two results establishes the proof of Property (iii). \square

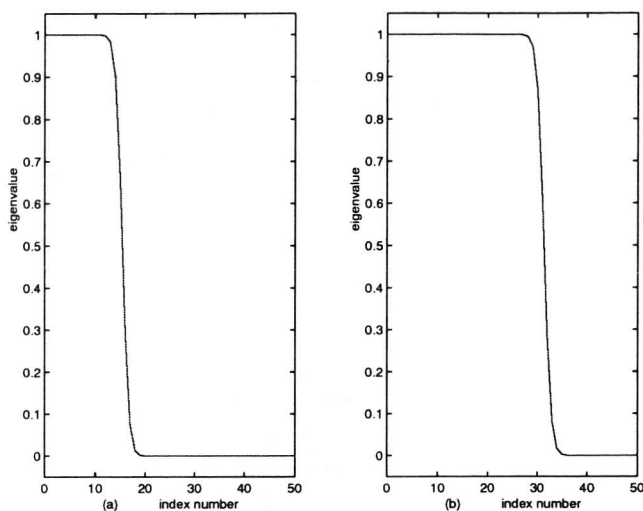


Figure 6.1: Eigenvalues corresponding to the PSWF for a) $x_0\omega_0 = 25$, b) $x_0\omega_0 = 50$.

From this theorem it follows, that for large $x_0\omega_0$ approximately the first $2x_0\omega_0/\pi$ eigenvalues that correspond to the PSWF attain a value close to unity. For index numbers in a region around $2x_0\omega_0/\pi$ the eigenvalues plunge to zero and attain values close to zero afterwards. The number of eigenvalues in the region where the eigenvalues decrease from close to one to close to zero is proportional to $\log x_0\omega_0$. Remark, that the eigenvalues depend on the product $x_0\omega_0$.

In Figure 6.1 the eigenvalues of $\mathcal{H}(x_0\omega_0)$ are depicted for a) $x_0\omega_0 = 25$ and b) $x_0\omega_0 = 50$ respectively. We observe that in both figures the number of eigenvalues close to unity is given by $2x_0\omega_0/\pi$. For $x_0\omega_0 = 25$, approximately the first 16 eigenvalues are close to unity. For $x_0\omega_0 = 50$, this number is approximately 32. The number of eigenvalues in the plunge region in Figure 6.1.b is approximately 1.25 times the number of eigenvalues in this region in Figure 6.1.a. This corresponds with the observation we have made after Theorem 6.1.3, namely that the multiplication factor is approximately given by $\log 32 / \log 16 = 5/4$.

6.2 Energy Concentration on a Circle in the Wigner Plane

The second problem to be considered is the concentration of energy in a circular region in the Wigner plane. So we consider a region

$$C_R = \{(x, \omega) \in \mathbb{R}^2 \mid x^2 + \omega^2 \leq R\} \quad (6.10)$$

and search for functions $f \in L^2(\mathbf{R})$ for which

$$E_f(R) = \int_{C_R} \mathcal{WV}[f](x, \omega) dx d\omega / \|f\|_2^2 \quad (6.11)$$

is maximized. An upperbound for $E_f(R)$ follows from an upperbound for $\mathcal{WV}[f]$ which can be derived from (5.15) in the following way

$$|\mathcal{WV}[f](x, \omega)| = |(\mathcal{M}_{-\omega} \mathcal{T}_{-x} f, \mathcal{M}_{\omega} \mathcal{T}_x f)| / \pi \leq \|f\|_2^2 / \pi.$$

This result yields

$$E_f(R) \leq R^2.$$

Of course a better and more natural upperbound for $E_f(R)$ would be given by 1, i.e., if $E_f(R)$ is the total amount of energy of f . A conjecture of Flandrin states that such an upperbound indeed exists, not only for integrals over circular regions, but in general for integrals over convex regions, see [28]. As far as we know, a proof of this conjecture has not been given yet. For non-convex regions this conjecture does not hold, which follows from various examples in [81].

We observe that from (2.41) it follows that

$$E_f(R) \rightarrow 1 \quad (R \rightarrow \infty),$$

if also $f \in L^1(\mathbf{R})$ or $\hat{f} \in L^1(\mathbf{R})$. Since the Wigner distribution can attain both positive and negative values, this result is not sufficient to prove Flandrin's conjecture.

In order to solve this energy localization problem, we introduce the localization operator $\mathcal{L}(\sigma)$ on $L^2(\mathbf{R})$, associated with a bounded symbol on \mathbf{R}^2 , by

$$(\mathcal{L}(\sigma)f, g)_2 = \int_{\mathbf{R}} \int_{\mathbf{R}} \sigma(x, \omega) W[f, g](x, \omega) dx d\omega, \quad (6.12)$$

for all $f, g \in L^2(\mathbf{R})$ and with $\mathcal{WV}[f, g]$ the mixed Wigner distribution of f and g . Then

$$E_f(R) = (\mathcal{L}(\sigma)f, f)_2 / (f, f)_2,$$

with $\sigma = \chi_{C_R}$. Furthermore, we observe that $\mathcal{L}(\sigma)$ is a Weyl transform with symbol $\sigma \in L^2(\mathbf{R}^2)$, see [107].

It can be proved, see e.g. [107], that $\mathcal{L}(\sigma)$ is compact for $\sigma \in L^p(\mathbf{R}^2)$, $1 \leq p \leq 2$. Moreover, Flandrin showed in [28] that $\mathcal{L}(\sigma)$ is self-adjoint for σ real-valued. This means that $\mathcal{L}(\sigma)$ is a compact Hermitian operator on $L^2(\mathbf{R})$ for real-valued $\sigma \in L^p(\mathbf{R}^2)$, $1 \leq p \leq 2$. Consequently, the eigenvectors of $\mathcal{L}(\sigma)$ can be chosen to form an orthonormal basis for $L^2(\mathbf{R})$, the

set of real-valued eigenvalues is countable and the only possible accumulation point is 0.

These considerations yield that the function f_{\max} , that maximizes $E_f(R)$ is given by the eigenvector ϕ_0 of $\mathcal{L}(\chi_{C_R})$ corresponding to the largest eigenvalue λ_0 of $\mathcal{L}(\chi_{C_R})$. Moreover, $E_{f_{\max}}(R)$ is given by λ_0 .

The eigenvectors of $\mathcal{L}(\chi_{C_R})$ are given by the Hermite functions h_k , $k \in \mathbb{N}$, as introduced in (5.10). This result was already given by Janssen in [46]. In the following lemma we come to the same result using a proof based on a property of the fractional Fourier transform.

Lemma 6.2.1 *Let $C_R = \{(x, \omega) \in \mathbb{R}^2 \mid x^2 + \omega^2 \leq R\}$ and $\mathcal{L}(\chi_{C_R})$ as defined in (6.12). Then the eigenvectors of $\mathcal{L}(\chi_{C_R})$ are given by*

$$\{h_k \mid k \in \mathbb{N}\}$$

with h_k the Hermite functions as defined in (5.10).

Proof

Since χ_{C_R} is rotation invariant, we have for all $\alpha \in [0, 2\pi)$

$$\begin{aligned} (\mathcal{L}(\chi_{C_R})\mathcal{F}_\alpha f, \mathcal{F}_\alpha g)_2 &= \int_{C_R} \mathcal{WV}[\mathcal{F}_\alpha f, \mathcal{F}_\alpha g](x, \omega) dx d\omega \\ &= \int_{C_R} \mathcal{WV}[f, g](R_\alpha(x, \omega)) dx d\omega \\ &= \int_{C_R} \mathcal{WV}[f, g](x, \omega) dx d\omega = (\mathcal{L}(\chi_{C_R})f, g)_2, \end{aligned}$$

with R_α the rotation matrix as given in (5.17). Consequently, we have for all $\alpha \in [0, 2\pi)$

$$\mathcal{F}_\alpha \mathcal{L}(\chi_{C_R}) = \mathcal{L}(\chi_{C_R}) \mathcal{F}_\alpha.$$

Let now ϕ_k be an eigenvector of $\mathcal{L}(\chi_{C_R})$ and λ_k its corresponding eigenvalue. Then

$$\mathcal{L}(\chi_{C_R})\mathcal{F}_\alpha \phi_k = \mathcal{F}_\alpha \mathcal{L}(\chi_{C_R})\phi_k = \lambda_k \mathcal{F}_\alpha \phi_k.$$

This shows, that if ϕ_k is an eigenvector of $\mathcal{L}(\chi_{C_R})$, then also $\mathcal{F}_\alpha \phi_k$ is an eigenvector of $\mathcal{L}(\chi_{C_R})$ for all $\alpha \in [0, 2\pi)$. Since $\mathcal{L}(\chi_{C_R})$ is compact, the set of eigenvectors

$$\{\mathcal{F}_\alpha \phi_k \mid \alpha \in [0, 2\pi)\}$$

should be finite or countable. This can only be realized if ϕ_k is an eigenvector of \mathcal{F}_α for all $\alpha \in [0, 2\pi)$, i.e., ϕ_k is a Hermite function following (5.12). \square

The eigenvalues λ_k of $\mathcal{L}(\chi_{C_R})$ can be expressed in terms of Laguerre polynomials L_k given by

$$L_k(x) = \frac{1}{k!} e^x \left(\frac{d}{dx} \right)^k (e^{-x} x^k). \quad (6.13)$$

In the following lemma we present a recurrence relation involving Laguerre polynomials that we shall use to compute the eigenvalues λ_k .

Lemma 6.2.2 Define $I_n(y) = \int_0^y e^{-x/2} L_n(x) dx$. Then

$$I_{n+1}(y) = -I_n(y) + 2e^{-y/2} (L_n(y) - L_{n+1}(y)). \quad (6.14)$$

Proof

First we observe that $L'_n(x) = L'_{n+1}(x) + L_n(x)$, which follows from the recurrence relations for Laguerre polynomials, and $L_n(0) = 1$, see e.g. [96]. Integration by parts yields

$$\begin{aligned} I_n(y) &= 2 - 2L_n(y)e^{-y/2} + 2 \int_0^y e^{-x/2} L'_n(x) dx \\ &= 2 - 2L_n(y)e^{-y/2} + 2I_n(y) + 2 \int_0^y e^{-x/2} L'_{n+1}(x) dx. \end{aligned}$$

We conclude $2 \int_0^y e^{-x/2} L'_{n+1}(x) dx = -I_n(y) + 2L_n(y)e^{-y/2} - 2$.

Applying the same procedure on I_{n+1} yields

$$I_{n+1}(y) = 2 - 2L_{n+1}(y)e^{-y/2} + 2 \int_0^y e^{-x/2} L'_{n+1}(x) dx,$$

or equivalently

$$2 \int_0^y e^{-x/2} L'_{n+1}(x) dx = I_{n+1}(y) + 2L_{n+1}(y)e^{-y/2} - 2.$$

Combining these two results completes the proof. \square

Using this lemma we come to the following recurrence relation for the eigenvalues of $\mathcal{L}(\chi_{C_R})$.

Theorem 6.2.3 Let $\{\lambda_k \mid k \in \mathbb{N}\}$ denote the set of eigenvalues of $\mathcal{L}(\chi_{C_R})$, with

$$C_R = \{(x, \omega) \in \mathbb{R}^2 \mid x^2 + \omega^2 \leq R\},$$

with $R > 0$. Then

- $\lambda_0 = (1 - e^{-R^2})$,
- $\lambda_{k+1} = \lambda_k - (-1)^k e^{-R^2} (L_k(2R^2) - L_{k+1}(2R^2))$, $k \in \mathbb{N} \setminus \{0\}$.

Proof

The Wigner distribution $\mathcal{WV}[h_k](x, \omega)$ can be expressed in terms of Laguerre polynomials, see e.g. [106]. This relation with Laguerre polynomials is given by

$$\mathcal{WV}[h_k](x, \omega) = 2(-1)^k (2\pi)^{-1} L_k(2(x^2 + \omega^2)) e^{-(x^2 + \omega^2)}.$$

Using polar coordinates we get

$$\begin{aligned} \lambda_k &= (\mathcal{L}(\chi_{C_R})h_k, h_k)_2 \\ &= \int_{C_R} \mathcal{WV}[h_k](x, \omega) dx d\omega \\ &= 2(-1)^k \int_0^R \rho L_k(2\rho^2) e^{-\rho^2} d\rho \\ &= \frac{(-1)^k}{2} \int_0^{2R^2} e^{-x/2} L_k(x) dx = (-1)^k I_k(2R^2)/2. \end{aligned}$$

Consequently, we have

$$\lambda_0 = I_0(2R^2)/2 = 1/2 \int_0^{2R^2} e^{-x/2} dx = (1 - e^{-R^2}).$$

Moreover, Lemma 6.2.2 yields

$$\begin{aligned} \lambda_{k+1} &= (-1)^{k+1} I_{k+1}(2R^2)/2 \\ &= (-1)^k I_k(2R^2)/2 + (-1)^{k+1} e^{-R^2} (L_k(2R^2) - L_{k+1}(2R^2)) \\ &= \lambda_k - (-1)^k e^{-R^2} (L_k(2R^2) - L_{k+1}(2R^2)). \end{aligned}$$

This gives the recurrence relation for the eigenvalues. □

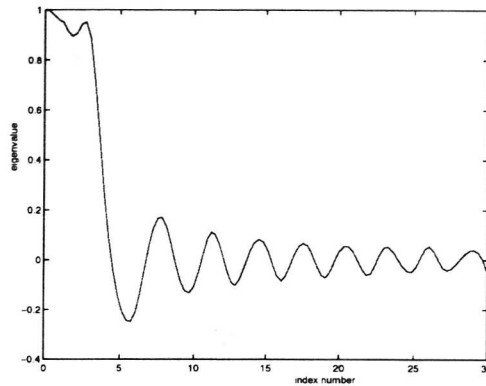


Figure 6.2: Eigenvalue behavior of the energy localization problem on a disk with radius $R = \sqrt{3}$.

In Figure 6.2 the first 30 eigenvalues as given in Theorem 6.2.3 are depicted for $R = \sqrt{3}$. To emphasize the eigenvalue behavior a spline interpolation function is used in this figure. As we have seen before for the eigenvalues Theorem 6.1.3, the first eigenvalues are close to λ_0 . Later the values plunge down towards zero and remain close to zero for larger index numbers. For the Wigner distribution, the eigenvalues can be negative, which can be observed in Figure 6.2 as well. Moreover, starting from a certain index number the eigenvalues alternate around zero.

6.3 Localization Problems and the Generalized FRFT

In this section we return to the fractional Fourier transform as introduced in Section 5.2.2. This generalized FRFT is used to solve two classes of energy localization problems that are related to the two problems, which we discussed in the previous sections. These two classes of localization problems are related to the discussed problems via the Weyl correspondence.

Although the problems we discuss concern signals in $L^2(\mathbb{R})$ we consider first localization problems for signals in $L^2(\mathbb{R}^n)$. For this we generalize the Weyl correspondence (6.12) to higher dimensions. Then a bounded symbol σ on \mathbb{R}^{2n} is associated with the localization operator $\mathcal{L}(\sigma)$ on $L^2(\mathbb{R}^n)$ by

$$(\mathcal{L}(\sigma)f, g)_2 = \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \sigma(x, \omega) W[f, g](x, \omega) dx d\omega, \quad (6.15)$$

for all $f, g \in L^2(\mathbb{R}^n)$. Consequently, if $\sigma = \chi_\Omega$, with $\Omega \subset \mathbb{R}^{2n}$, then

$$(\mathcal{L}(\sigma)f, f)_2 = \int_{\Omega} W[f](x, \omega) dx d\omega$$

represents the energy of f in the Wigner plane within the region Ω .

Using the generalized FRFT $\mathcal{F}_{\Gamma, \Delta}$ as introduced in (5.42) we compute

$$\begin{aligned} (\mathcal{F}_{\Gamma, \Delta} \mathcal{L}(\sigma) \mathcal{F}_{\Gamma, \Delta}^* f, g)_2 &= (\mathcal{L}(\sigma) \mathcal{F}_{\Gamma, \Delta}^* f, \mathcal{F}_{\Gamma, \Delta}^* g) \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \sigma(x, \omega) W[\mathcal{F}_{\Gamma, \Delta}^* f, \mathcal{F}_{\Gamma, \Delta}^* g](x, \omega) dx d\omega \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \sigma(x, \omega) W[f, g](A^{-1}(x, \omega)) dx d\omega \\ &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \sigma(A(x, \omega)) W[f, g](x, \omega) dx d\omega \\ &= (\mathcal{L}(\sigma_A) f, g)_2, \end{aligned}$$

with $\sigma_A(x, \omega) = \sigma(A(x, \omega))$ and A as given in (5.44). Now, assume $\{\phi_k \mid k \in \mathbb{N}\}$ is the set of eigenvectors of $\mathcal{L}(\sigma)$ and $\{\lambda_k \mid k \in \mathbb{N}\}$ the set of corresponding eigenvalues. Then

$$\begin{aligned} \mathcal{L}(\sigma_A) \mathcal{F}_{\Gamma, \Delta} \phi_k &= (\mathcal{F}_{\Gamma, \Delta} \mathcal{L}(\sigma) \mathcal{F}_{\Gamma, \Delta}^*) \mathcal{F}_{\Gamma, \Delta} \phi_k \\ &= \mathcal{F}_{\Gamma, \Delta} \mathcal{L}(\sigma) \phi_k = \lambda_k \mathcal{F}_{\Gamma, \Delta} \phi_k. \end{aligned} \quad (6.16)$$

Consequently, the eigenvectors and eigenvalues of $\mathcal{L}(\sigma_A)$ are given by

$$\{\mathcal{F}_{\Gamma, \Delta} \phi_k \mid k \in \mathbb{N}\} \text{ and } \{\lambda_k \mid k \in \mathbb{N}\}$$

respectively. If $\mathcal{L}(\sigma)$ is a compact operator, both the eigenvectors ϕ_k and $\mathcal{F}_{\Gamma, \Delta} \phi_k$ form an orthonormal set in $L^2(\mathbb{R}^n)$.

6.3.1 The Rectangle/Parallelogram Case and the Rihaczek Distribution

The first problem we consider is to maximize

$$(\mathcal{L}(\sigma)f, f)_2 / (f, f)_2 \quad (6.17)$$

for $f \in L^2(\mathbb{R})$, with $\sigma = \chi_{[-x_0, x_0] \times [-\omega_0, \omega_0]}$.

This problem may seem to be similar to Slepian's energy problem in Section 6.1. However, results presented for Slepian's energy problem cannot be related to the problem of localizing

the energy on a rectangle in the Wigner plane.

The two problems can only be related to each other if (6.17) is maximized over absolutely integrable $f \in L^2_{comp}(\mathbb{R})$, with $\text{supp}(f) = [-x_0, x_0]$. Using these constraints (6.17) is equal to (6.1), which follows straightforwardly from Theorem 2.3.4. If we do not require these constraints on the maximizing function f , we are only provided with some asymptotical results on the eigenvalues of $\mathcal{L}(\sigma)$, see [41, 81].

A less trivial relation with Slepian's energy problem is given for

$$\sigma = \chi_{[-x_0, x_0] \times [-\omega_0, \omega_0]} * \varphi, \quad (6.18)$$

for some $x_0, \omega_0 \in \mathbb{R}^+$ and where φ is given by

$$\varphi(x, \omega) = e^{-2ix\omega}.$$

Using Young's inequality we have

$$\|\sigma\|_\infty = \|\chi_{[-x_0, x_0] \times [-\omega_0, \omega_0]} * \varphi\|_\infty \leq 2x_0\omega_0/\pi,$$

and so $\sigma \in L^\infty(\mathbb{R}^2)$.

The following lemma shows that the localization operator $\mathcal{L}(\sigma)$, with σ as in (6.18), can be rewritten as an energy density operator related to the Rihaczek distribution, see [82].

Lemma 6.3.1 *Let $\mathcal{L}(\sigma)$ be the localization operator as defined in (6.15), with σ the symbol as given in (6.18). Then for all $f, g \in L^2(\mathbb{R})$*

$$(\mathcal{L}(\sigma)f, g)_2 = \int_{\mathbb{R}} \int_{\mathbb{R}} \chi_{[-x_0, x_0] \times [-\omega_0, \omega_0]}(x, \omega) \mathcal{R}[f, g](x, \omega),$$

with $\mathcal{R}[f, g]$ the mixed Rihaczek distribution given by

$$\mathcal{R}[f, g](x, \omega) = f(x) \overline{\hat{g}(\omega)} e^{-i\omega x} / \sqrt{2\pi}. \quad (6.19)$$

Proof

We observe that

$$(\mathcal{L}(\sigma)f, g)_2 = (\sigma_0 * \varphi, \mathcal{WV}[f, g])_2 = (\sigma_0, \varphi * \mathcal{WV}[f, g])_2,$$

with $\sigma_0 = \chi_{[-x_0, x_0] \times [-\omega_0, \omega_0]}$. This expression can be rewritten by

$$\begin{aligned} & (\varphi * \mathcal{WV}[f, g])(x, \omega) \\ &= \frac{1}{2\pi^2} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \varphi(p, q) f(x - p + t) \overline{g(x - p - t)} e^{-2it(\omega - q)} dt dp dq \end{aligned}$$

$$\begin{aligned}
 &= \frac{1}{4\pi^2} \int_{\mathbf{R}} \int_{\mathbf{R}} \int_{\mathbf{R}} e^{-iqx} f(u) \overline{g(v)} e^{-iu(\omega-q)} e^{iv\omega} du dv dq \\
 &= \frac{1}{2\pi} \int_{\mathbf{R}} e^{-iqx} \hat{f}(\omega-q) \overline{\hat{g}(\omega)} dq = \frac{1}{2\pi} e^{-i\omega x} \overline{\hat{g}(\omega)} \int_{\mathbf{R}} e^{-i\omega x} \hat{f}(q) e^{iqx} dq \\
 &= f(x) \overline{\hat{g}(\omega)} e^{-i\omega x} / \sqrt{2\pi}.
 \end{aligned}$$

This yields $(\mathcal{L}(\sigma)f, g)_2 = \frac{1}{\sqrt{2\pi}} \int_{\mathbf{R}} \int_{\mathbf{R}} \sigma_0(x, \omega) f(x) \overline{\hat{g}(\omega)} e^{-i\omega x} dx d\omega$ □

Using this lemma we prove the following theorem, that relates $\mathcal{L}(\sigma)$, with σ as in (6.18), with the localization operator of Slepian's energy problem.

Theorem 6.3.2 *Let $\mathcal{L}(\sigma)$ be the operator as in (6.15), with σ the symbol as in (6.18). Then*

$$\mathcal{L}(\sigma)^* \mathcal{L}(\sigma) = \mathcal{P}(x_0) \mathcal{B}(\omega_0) \mathcal{P}(x_0),$$

with $\mathcal{B}(\omega_0)$ and $\mathcal{P}(x_0)$ as defined in (6.2) and (6.4) respectively.

Proof

From the preceding lemma it follows immediately that

$$\begin{aligned}
 \mathcal{L}(\sigma)^*[g](x) &= \chi_{[-x_0, x_0]}(x) \cdot \frac{1}{\sqrt{2\pi}} \int_{-\omega_0}^{\omega_0} \hat{g}(\omega) e^{i\omega x} d\omega \\
 &= \mathcal{P}(x_0) \mathcal{B}(\omega_0).
 \end{aligned}$$

Since both $\mathcal{B}(\omega_0)$ is a projection operators, we have

$$\mathcal{L}(\sigma)^* \mathcal{L}(\sigma) = \mathcal{P}(x_0) \mathcal{B}(\omega_0) \mathcal{P}(x_0).$$

□

Remark, that although $\sigma \in L^\infty(\mathbf{R})$, $\mathcal{L}(\sigma)$ is compact for σ as in (6.18). This follows from the fact that $\mathcal{L}(\sigma)^* \mathcal{L}(\sigma)$ is compact. Furthermore, we observe that the result of Theorem 6.3.2 was already given in [28]. However, our aim is not to investigate existing time-frequency distributions, but to consider the generalized FRFT acting on these distributions. In this context, we return to the first part of this section.

We have seen that the eigenvalues of $\mathcal{L}(\sigma)$ and $\mathcal{L}(\sigma_A)$ coincide. In a direct way, we can also show that the eigenvalues of $\mathcal{L}(\sigma)^* \mathcal{L}(\sigma)$ and $\mathcal{L}(\sigma_A)^* \mathcal{L}(\sigma_A)$ coincide. This yields that the singular values of $\mathcal{L}(\sigma)$ and $\mathcal{L}(\sigma_A)$ are the same. These singular values are given by

$$s_k = \sqrt{\lambda_k},$$

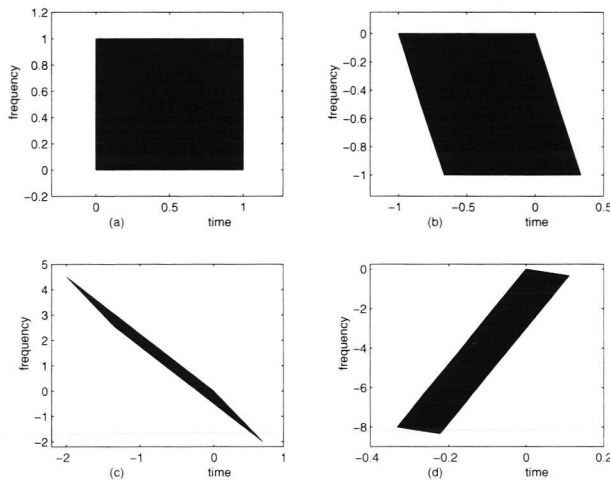


Figure 6.3: Localisation on a rectangle/parallelogram: fig. a) $\sigma = 1$ on $[0, 1] \times [0, 1]$ and fig. b, c, d) σ_A with $\Delta = -1/\Gamma$, $\Delta = -2/\Gamma$ and $\Delta = -1/\Gamma^2$ respectively.

where λ_k denote the eigenvalues of the operator $\mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)$. Since these λ_k satisfy Theorem 6.1.3, a similar result holds for the singular values. Moreover, the asymptotical behavior of s_k and λ_k is similar.

The eigenvectors of $\mathcal{L}(\sigma)$ do not follow from Theorem 6.3.2. The eigenvectors of $\mathcal{L}(\sigma)^*\mathcal{L}(\sigma)$ are known, namely the prolate spheroidal wave functions ψ_k . As before we can also show that the eigenvectors of $\mathcal{L}(\sigma_A)^*\mathcal{L}(\sigma_A)$ are then given by $\mathcal{F}_{\Gamma,\Delta}\psi_k$. They can be computed as the eigenvectors of the operator

$$\mathcal{D}'(x_0\omega_0) = \mathcal{F}_{\Gamma,\Delta}\mathcal{D}(x_0\omega_0)\mathcal{F}_{\Gamma,\Delta}^*,$$

which is also a second order differential operator that commutes with $\mathcal{L}(\sigma_A)^*\mathcal{L}(\sigma_A)$.

In Figure 6.3.b,c and d the domain of σ_A is depicted instead of σ , with the substitutions $\Delta = -1/\Gamma$, $\Delta = -2/\Gamma$ and $\Delta = -1/\Gamma^2$ and with $\Gamma = 3$ in (5.44). We observe that with these substitutions $\mathcal{L}(\sigma_A)$ represents the energy of the Rihaczek distribution within differently orientated parallelograms in phase space. The singular values of $\mathcal{L}(\sigma_A)$ for all A related to these parallelograms are the same and are given by $\sqrt{\lambda_k}$, with λ_k as in Theorem 6.1.3.

6.3.2 The Circle/Ellipse Case

In Section 6.2 we already discussed the energy localization problem on a circle. Moreover, we studied the operator $\mathcal{L}(\chi_{C_R})$, with C_R a circle in the Wigner plane concentrated around

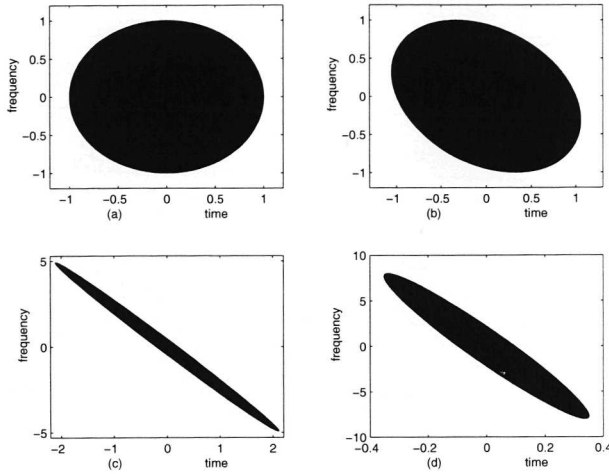


Figure 6.4: Localisation on a circle/ellipse: fig. a) $\sigma = 1$ on $\{(x, \omega) \in \mathbb{R}^2 \mid x^2 + \omega^2 \leq 1\}$, fig. b, c, d) σ_A with $\Delta = -1/\Gamma$, $\Delta = -2/\Gamma$ and $\Delta = -1/\Gamma^2$ respectively.

the origin and with radius $R > 0$. It turned out that its eigenvectors are given by the Hermite functions h_k , defined by (5.11), and that the corresponding eigenvalues are given by Theorem 6.2.3.

It follows from (6.16), that the eigenvectors of $\mathcal{L}(\sigma_A)$, with A as given in (5.44), are given by $\mathcal{F}_{\Gamma, \Delta} h_k$, $k \in \mathbb{N}$. The eigenvalues of $\mathcal{L}(\sigma_A)$ are given by the recurrence relation in Theorem 6.2.3.

In Figure 6.4.b,c and d the domain of σ_A is depicted with σ the characteristic function of C_R , with the substitutions $\Delta = -1/\Gamma$, $\Delta = -2/\Gamma$ and $\Delta = -1/\Gamma^2$ and with $\Gamma = 3$. With these substitutions $\mathcal{L}(\sigma_A)$ represents the energy in the Wigner plane within differently orientated ellipses. The energy localization problem for each of these ellipsoidal areas is now solved by the eigenvectors $\mathcal{F}_{\Gamma, \Delta} h_k$, using the corresponding substitutions, and the eigenvalues λ_k .

Chapter 7

A Seismic Problem: Automatic S-Phase Detection with the DWT

In this chapter we consider a seismic problem, which has been dealt with while the author was seconded to the seismic department of the Royal Dutch Meteorological Institute (KNMI). For this problem an algorithm based on the DWT has been derived. Implementation and test results of the algorithm are also included in this chapter. These were established at KNMI by R. Sleeman and T. van Eck. This chapter is mainly based on [70].

7.1 Introduction and Approach of the Seismic Problem

A seismic earthquake signal recorded by a seismic station (seismogram) is built up by several different seismic waves (phases), which characterize the type of the signal. Amongst others, significant phases that appear in a seismogram are the P-phase (primary phase) and the S-phase (secondary or shear phase), which we consider here. The problem we are dealing with, is to detect automatically the S-phase and to determine its arrival time, once the P-phase arrival time is known with high accuracy as in [3, 6, 90]. This arrival time is defined as the time sample in the seismogram at which the P-phase appears for the first time. An accurate estimate of these arrival times is important for determining the type and location of the seismic event.

The S-phase arrival time is determined in a three-component seismogram, representing motion on a ground detector in three mutually orthogonal directions, two in the horizontal plane (x-y plane) and one vertical direction (z-axis). An example of a three-component seismogram is depicted in Figure 7.1.

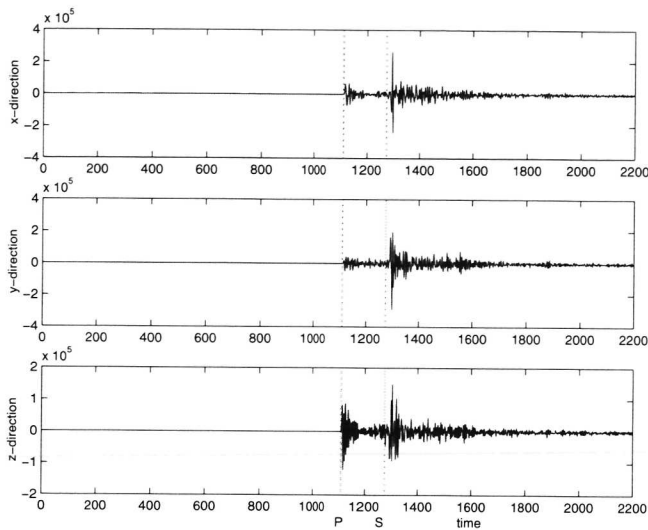


Figure 7.1: A three-component seismogram, with P- and S-phase arrival times picked by an analyst at $i_P = 1120$ and $i_S = 1275$ respectively.

The detection of the S-phase arrival time is mostly based on some physical differences between the P-phase and the S-phase, as described thoroughly in [2, 13]. For our problem the most obvious property is the difference in arrival times. The S-phase is always delayed as compared to the P-phase arrival at the seismic station. A more fundamental property is the fact that P-phases compress volumes and S-waves deform volumes. Furthermore, the S particle motion, i.e. the direction of the S-phase when it arrives at the earth's surface, is contained in a plane perpendicular to the direction of the P particle motion, called the S-plane. This property only holds if reflections at the earth's surface may be ignored or when the phases arrive in a direction almost perpendicular to the earth's surface. In our problem the latter assumption is justified. The P-phase travels along the travel direction of the seismic event, unless the medium is anisotropic. Finally, comparing the frequency spectra of both phases, the P-phase appears at higher frequencies than the S-phase.

The automatic S-phase detection algorithm that we present in this chapter is a combination of traditional methods to detect S-phases as described before and the discrete wavelet transform for $l^2(\mathbb{Z})$ as introduced in Section 3.2.3. The idea to analyse the three components of a seismogram at several scales has been described already in the literature, e.g. [5, 69]. However in these papers only the wavelet transform itself has been used as a phase detector in seismograms, whereas in our method the wavelet transform has been used in combination with traditional approaches, which are very well known in seismology and have been discussed in

the past in various papers, e.g. [13, 48, 84]. In Section 7.5 we compare results of this new algorithm with results based on ideas of Cichowicz [13].

7.2 Seismic methods for S-Phase Picking

A common strategy to detect phase arrival times is to construct one or more so-called characteristic functions. These are discrete-time functions, with some specific properties at the time sample, at which a phase appears in the seismogram. In this section we discuss some of them for detecting S-phase arrival times.

7.2.1 Characteristic functions based on a cross-power matrix

We consider a (real valued) three-component seismic signal $u \in l^2(\mathbb{Z}, \mathbb{R}^3)$. The time-dependent N -point cross-power matrix for such signal u is then defined as follows.

Definition 7.2.1 Let $N \in \mathbb{N}$ and $u \in l^2(\mathbb{Z}, \mathbb{R}^3)$. Then the N -point cross-power matrix of u at $i \in \mathbb{Z}$ is given by

$$M_{N,u}(i) = \begin{pmatrix} \langle u_1, u_1 \rangle_i & \langle u_1, u_2 \rangle_i & \langle u_1, u_3 \rangle_i \\ \langle u_2, u_1 \rangle_i & \langle u_2, u_2 \rangle_i & \langle u_2, u_3 \rangle_i \\ \langle u_3, u_1 \rangle_i & \langle u_3, u_2 \rangle_i & \langle u_3, u_3 \rangle_i \end{pmatrix}, \quad (7.1)$$

for $i \in \mathbb{Z}$, with

$$\langle u_n, u_m \rangle_i = 1/N \sum_{k=i}^{i+N-1} u_n(k) \cdot u_m(k),$$

for all $i \in \mathbb{Z}$.

We observe that $\langle \cdot, \cdot \rangle_i$ in this definition depends also on N . The window length N itself is chosen depending on the (non)-stationary character of the signal u . Generally we take $N \sim 1/\omega_0$, with ω_0 the dominating frequency of u . Furthermore, before computing the cross-power matrix of a signal at a certain time sample, we first create a signal u , with zero mean at each component. This is done by subtracting the means of the components from the seismic signal. The reason for doing this is to neglect possible offsets without seismic cause. These can be generated by the measurement equipment.

Starting with $M_{N,u}(i)$ we can analyse the signal using the eigenvalues and eigenvectors of $M_{N,u}(i)$. Note that $M_{N,u}(i)$ is the Gram matrix of the set $\{\chi_{[i, i+N-1]} u_1, \chi_{[i, i+N-1]} u_2, \chi_{[i, i+N-1]} u_3\}$, with χ_X the characteristic function on \mathbb{Z} of a discrete set X . So $M_{N,u}(i)$ is a positive semi-definite matrix. Therefore the eigenvalues of $M_{N,u}(i)$ are real and positive,

$$\lambda_1(i) \geq \lambda_2(i) \geq \lambda_3(i) \geq 0,$$

and the eigenvectors $v_1(i)$, $v_2(i)$ and $v_3(i)$ can be chosen to form an orthonormal basis in \mathbb{R}^3 . The following two characteristic functions are based on these eigenvalues and eigenvectors. Hereby we assume $M_{N,u}(i) \neq 0$, for any $i \in \mathbb{Z}$, which is quite a realistic assumption. We observe that $M_{N,u}(i) = 0$ is equivalent with $u_1(k) = u_2(k) = u_3(k) = 0$, for $k = i, \dots, i + N - 1$.

Deflection angle:

Let $v_1(i)$ denote the eigenvector of $M_{N,u}(i)$ corresponding to $\lambda_1(i)$. This eigenvector $v_1(i)$ represents the direction of the particle motion at time i with most seismic energy. Let i_P be the P-phase arrival time. Then $v_1(i_P)$ is the direction of the P particle motion. The deflection angle is defined by

$$\kappa_1(i) = \frac{2}{\pi} \arccos \left(\frac{|(v_1(i), v_1(i_P))|}{\|v_1(i)\| \cdot \|v_1(i_P)\|} \right). \quad (7.2)$$

Note that $\kappa_1(i_P) = 0$. Furthermore, since the direction of the S-phase particle motion $v_1(i_S)$ is perpendicular to $v_1(i_P)$, κ_1 attains its maximum 1 at the S-phase arrival time i_S .

Degree of polarization:

Following [85], the degree of polarization is defined by

$$\kappa_2(i) = \frac{(\lambda_1(i) - \lambda_2(i))^2 + (\lambda_1(i) - \lambda_3(i))^2 + (\lambda_2(i) - \lambda_3(i))^2}{2 \cdot (\lambda_1(i) + \lambda_2(i) + \lambda_3(i))^2}. \quad (7.3)$$

This characteristic function can be used both for detecting P-phase arrival times and S-phase arrival times, since all types of seismic polarization, i.e. at all different phase arrival times seismic energy is concentrated along one single direction. In practice such an arrival yields $\lambda_1(i) \gg \lambda_2(i)$, which means that we may expect maxima for κ_2 at both i_P and i_S , the S-phase arrival time.

Remark that κ_1 and κ_2 depend on $M_{N,u}$, and so they depend on the window length N . In Section 7.4.2 we will also discuss window lengths that depend on the frequency behaviour of the signal u . Furthermore, we observe that for all characteristic functions κ introduced in this section we have

$$0 \leq \kappa(i) \leq 1 \quad \forall i \in \mathbb{Z} \quad (7.4)$$

In the sequel we will use a combination of three characteristic functions $\kappa_1, \kappa_2, \kappa_3$ such that

$$\prod_{m=1}^3 \kappa_m^2(i) = \max_{n \in \mathbb{Z}} \prod_{m=1}^3 \kappa_m^2(n) \iff i = i_S. \quad (7.5)$$

The square product of the characteristic functions will be used to emphasize the maximum values attained in each function at i_S and to reduce other (local) maxima, related to features in the signal other than the S-phase arrival. We have already met two candidates to be used in this product of functions. In the following subsection we introduce the third function that can be used in (7.5).

7.2.2 Rotation and the Energy Ratio

We transform the three-component seismic signal representing motion on a ground detector into a three-component signal representing motion in the longitudinal direction and in two transversal directions. The longitudinal direction is the direction of the P particle motion ($v_1(i_P)$). The transversal directions are mutually orthogonal and are chosen in the plane perpendicular to the longitudinal direction ($\text{span}\{v_2(i_P), v_3(i_P)\}$). This transversal plane is also called the S-plane, since the direction of the S particle motion is in the S-plane.

The seismic signal is transformed into the basis $\{v_1(i_P), v_2(i_P), v_3(i_P)\}$ by

$$\begin{pmatrix} u_L(i) \\ u_Q(i) \\ u_T(i) \end{pmatrix} = V(i_P) \begin{pmatrix} u_1(i) \\ u_2(i) \\ u_3(i) \end{pmatrix}, \quad (7.6)$$

with

$$V(i_P) = (v_1(i_P) \mid v_2(i_P) \mid v_3(i_P))^T. \quad (7.7)$$

The third characteristic function we use in (7.5) is the fraction of energy in the S-plane to the total amount of energy in the signal, given by

$$\kappa_3(i) = \frac{\sum_{n=i}^{i+N-1} (u_Q(n)^2 + u_T(n)^2)}{\sum_{n=i}^{i+N-1} (u_L(n)^2 + u_Q(n)^2 + u_T(n)^2)}, \quad (7.8)$$

for some $N \in \mathbb{N}$. This definition can also be rewritten as

$$\kappa_3(i) = 1 - \frac{\sum_{n=i}^{i+N-1} (v_1(i_P), u(n))^2}{\sum_{n=i}^{i+N-1} (u(n), u(n))}, \quad (7.9)$$

which shows how κ_3 depends on i_P . Note that we may expect a minimum $\kappa_3(i_P) = 0$ and a maximum $\kappa_3(i_S) = 1$. Furthermore, κ_3 satisfies (7.4).

7.3 On the Use of Characteristic Functions

In this section we discuss the stability of the characteristic functions that we considered in the previous section. Also we discuss problems we have to deal with when using the characteristic functions in practice.

7.3.1 Error Analysis of the Characteristic Functions

We consider three kind of errors that can appear in the characteristic functions κ . Amongst others, an incorrect κ can be the result of computational and truncation errors in the matrices $M_{N,u}$, measurement errors in the signal u , and the determination of an incorrect P-phase arrival time i'_P . We will show that κ depends continuously on the errors as described above.

First we consider stability with respect to computational and truncation errors in $M_{N,u}$ and measurement errors in u . To prove the stability of the characteristic functions we present some auxiliary results from linear algebra.

Theorem 7.3.1 *Let $M = A + \Delta$, with M, A and Δ ($n \times n$) matrices with $\|\Delta\|_2 \ll 1$. Furthermore, let $\alpha_1, \dots, \alpha_n$ be the eigenvalues of A and let u_1, \dots, u_n be their corresponding eigenvectors. Finally, we assume $\lambda_k \neq \lambda_j$, $k \neq j$, with j fixed. Then a first order approximation of the eigenvector v_j is given by*

$$v_j = u_j + \sum_{l \neq j} \frac{(\Delta u_j, w_l)}{(\alpha_j - \alpha_l)(u_l, w_l)} u_l, \quad (7.10)$$

with w_1, \dots, w_n the eigenvectors corresponding to the eigenvalues $\bar{\alpha}_1, \dots, \bar{\alpha}_n$ of A^H .

Proof

Cf. [32].

□

With this theorem we can prove the following corollary.

Corollary 7.3.2 *Let H be a Hermitian ($n \times n$) matrix, $n > 1$, with eigenvectors u_1, \dots, u_n , $\|u_k\| = 1$, $k = 1, \dots, n$. Furthermore, we assume $\lambda_k \neq \lambda_j$, $k \neq j$, with j fixed. Then the function g_j , given by*

$$g_j(H) = u_j$$

is continuous.

Proof

Let $0 < \varepsilon \ll 1$ and let Δ be a ($n \times n$) matrix, $n > 1$, with $\|\Delta\|_2 < \varepsilon \cdot \alpha / (n - 1)$ with

$\alpha = \min_{k \neq j} |\alpha_j - \alpha_k|$. Here $\alpha_1, \dots, \alpha_n$ are the eigenvalues of H . Then by (7.10)

$$\begin{aligned} \|g_k(H + \Delta) - g_k(H)\|_2 &= \left\| \sum_{k \neq j} \frac{(\Delta u_j, u_k)}{\alpha_j - \alpha_k} u_k \right\|_2 \leq \sum_{k \neq j} \left| \frac{(\Delta u_j, u_k)}{\alpha_j - \alpha_k} \right| \\ &\leq (n-1) \|\Delta\|_2 / \alpha < \varepsilon. \end{aligned}$$

□

We follow the definitions of κ_1 and κ_3 in (7.2) and (7.9) respectively. Now, Corollary 7.3.2 immediately shows that errors in κ_1 and κ_3 depend continuously on errors in $M_{N,u}(i)$ under the condition that $\lambda_1(i_P) > \lambda_2(i_P)$. This condition will be satisfied in practice, since $\lambda_1(i) \gg \lambda_2(i)$ for all phase arrival times i .

From [32] we take the following result to prove the stability of κ_2 .

Theorem 7.3.3 *Let $M = A + \Delta$, with M, A and Δ Hermitian $(n \times n)$ matrices. Let further $\mu_1 \geq \dots \geq \mu_n$, $\alpha_1 \geq \dots \geq \alpha_n$ and $\delta_1 \geq \dots \geq \delta_n$ be the eigenvalues of M, A and Δ respectively. Then*

$$\alpha_k + \delta_n \leq \mu_k \leq \alpha_k + \delta_1. \quad (7.11)$$

Using this theorem we arrive at the following corollary.

Corollary 7.3.4 *Let K_n denote the set of all Hermitian $(n \times n)$ matrices and let the mapping $\lambda : K_n \rightarrow \mathbb{R}^n$ be given by $\lambda(H) = (\lambda_1, \dots, \lambda_n)$, with $\lambda_1 \geq \dots \geq \lambda_n \geq 0$ the eigenvalues of $H \in K_n$. Then λ is continuous on K_n .*

Proof

Let $\varepsilon > 0$ and let Δ be a Hermitian $(n \times n)$ matrix, with eigenvalues $\delta_1 \geq \dots \geq \delta_n$ and such that $\|\Delta\|_2 < \varepsilon$. Furthermore, let $\lambda'_1 \geq \dots \geq \lambda'_n$ denote the eigenvalues of $H + \Delta$. Then by (7.11)

$$|\lambda'_k - \lambda_k| \leq \max\{|\delta_1|, |\delta_n|\} \leq \|\Delta\|_2 < \varepsilon.$$

Therefore $\|\lambda(H + \Delta) - \lambda(H)\|_2 < \varepsilon$. □

Corollary 7.3.4 yields immediately

Corollary 7.3.5 *Let $f : \mathbb{R}^n \rightarrow [0, 1]$ be continuous and let K_n denote the set of all Hermitian $(n \times n)$ matrices. Then $\kappa = f \circ \lambda$ is continuous on K_n , with λ as defined in Corollary 7.3.4.*

Since κ_2 can be written as in Corollary 7.3.5, also errors in κ_2 depend continuously on errors in $M_{N,u}$. Resuming, we have proved

$$\forall \varepsilon > 0 \exists \delta > 0 \quad \|M_{N,u}(i) - M'_{N,u}(i)\|_2 < \delta \implies |\kappa_n(i) - \kappa'_n(i)| < \varepsilon, \quad n = 1, 2, \quad (7.12)$$

for any fixed $i \in \mathbb{Z}$ and with κ'_n the characteristic function associated with the perturbed matrices $M'_{N,u}$. The only restriction we have to make is that $\lambda_1(i_P) > \lambda_2(i_P)$, which is always the case when the signal is polarized. We will now show that matrices $M_{N,u}$ depend continuously on u in the following sense:

Lemma 7.3.6 *Let $u, w \in l^2(\mathbb{Z}, \mathbb{R}^3)$ and let $M_{N,u}$ and $M_{N,w}$ be the cross-power matrices associated with u and w respectively. Then*

$$\forall \varepsilon > 0 \exists \delta > 0 \forall i \in \mathbb{Z} : \|u - w\|_\infty < \delta \implies \|M_{N,u}(i) - M_{N,w}(i)\|_2 < \varepsilon. \quad (7.13)$$

Proof

Let $\gamma \in l^2(\mathbb{Z}, \mathbb{R}^3)$ be defined as $\gamma(n) = u(n) - w(n)$ and let $\Delta(i) \in \mathbb{R}^{3 \times 3}$ be defined by $\Delta(i) = M_{N,u}(i) - M_{N,w}(i)$. With a straightforward calculation we get

$$\begin{aligned} |\Delta_{k,l}(i)| &= | \langle u_k, u_l \rangle_i - \langle w_k, w_l \rangle_i | \\ &= | \langle u_k, u_l \rangle_i - \langle u_k - \gamma_k, u_l - \gamma_l \rangle_i | \\ &= | \langle \gamma_k, u_l \rangle_i + \langle u_k, \gamma_l \rangle_i - \langle \gamma_k, \gamma_l \rangle_i | \\ &\leq 2N \|\gamma\|_\infty \cdot \|u\|_\infty + N \|\gamma\|_\infty^2. \end{aligned}$$

So $\|\Delta(i)\|_\infty \rightarrow 0$ if $\|u - w\|_\infty \rightarrow 0$. The proof is completed by the equivalence of matrix norms. \square

Resuming, for all characteristic functions of Section 7.2 we proved stability with respect to computational and truncation errors in $M_{N,u}$ and measurement errors in u .

The last kind of error in the characteristic functions we discuss here is the error due to an incorrectly determined P-phase arrival time i'_P .

With a straightforward calculation we get $M_{N,Vu}(i) = VM_{N,u}(i)V^T$ for all (3×3) matrices V . For orthonormal matrices V this relation yields $\sigma(M_{N,Vu}(i)) = \sigma(M_{N,u}(i))$, with $\sigma(A)$ denoting the spectrum of an $(n \times n)$ matrix A . In particular we have $\sigma(M_{N,V(i_P)u}(i)) = \sigma(M_{N,V(i'_P)u}(i))$ yielding that κ_2 is invariant under any orthogonal transformation of u . So an incorrect P-phase arrival time i'_P will not affect κ_2 .

The deflection angle κ_1 is affected by an incorrect i_P . An expression for the error in κ_1 due to an incorrect i_P is given in the following lemma.

Lemma 7.3.7 *Let $u \in l^2(\mathbb{Z}, \mathbb{R}^3)$ and let \mathcal{R} denote the shift on $l^2(\mathbb{Z}, \mathbb{R}^3)$ given by*

$$(\mathcal{R}y)(k) = y(k-1).$$

Furthermore, let $M_{N,u}$ be the cross-power matrix associated with u and assume $\lambda_1(i_P) > \lambda_2(i_P)$. Then

$$\forall \varepsilon > 0 \exists \delta > 0 : \|u - \mathcal{R}^{i'_P - i_P} u\|_\infty < \delta \implies \|\kappa_1 - \kappa'_1\|_\infty < \varepsilon, \quad (7.14)$$

with κ_1 as given in (7.2) and with

$$\kappa'_1 = 2 \arccos(|(v_1(i), v_1(i'_P)))| / \pi.$$

Proof

We define $\delta(i) = |(v_1(i), v_1(i'_P))| - |(v_1(i), v_1(i_P))|$. Assume $\|u - \mathcal{R}^{i'_P - i_P} u\|_\infty \rightarrow 0$, then Lemma 7.3.6 yields with $w = \mathcal{R}^{i'_P - i_P} u$,

$$\|M_{N,u}(i_P) - M_{N,u}(i'_P)\|_2 \rightarrow 0.$$

Now, by Corollary 7.3.2 and the assumption $\lambda_1(i_P) > \lambda_2(i_P)$ we get

$$\|v_1(i'_P) - v_1(i_P)\|_2 \rightarrow 0.$$

This yields

$$\begin{aligned} \delta(i) &= |(v_1(i), v_1(i'_P))| - |(v_1(i), v_1(i_P))| \\ &\leq |v_1(i), v_1(i'_P) - v_1(i_P)| \rightarrow 0, \end{aligned}$$

for all $i \in \mathbb{Z}$. For $|\kappa_1 - \kappa'_1|$ we derive an expression using a trigonometric formula.

$$\begin{aligned} |\arccos a - \arccos b| &= |\arccos(ab + \sqrt{1-a^2}\sqrt{1-b^2})| \\ &= |\arccos(a^2 + ad + \sqrt{1-a^2}\sqrt{1-(a+d)^2})| \\ &= |\arccos(a^2 + ad + \sqrt{(1-a^2-ad)^2 - d^2})|, \end{aligned}$$

with $d = b - a$. After substituting $a = |(v_1(i), v_1(i_P))|$, $b = |(v_1(i), v_1(i'_P))|$ and $d = \delta(i)$ we obtain

$$|\kappa_1(i) - \kappa'_1(i)| = 2 |\arccos(x(i) + \sqrt{(1-x(i))^2 - \delta(i)^2})| / \pi,$$

with $x(i) = |(V(i_P)v_1(i))_1|^2 + \delta(i)|V(i_P)v_1(i)|_1$, for all $i \in \mathbb{Z}$. So we get

$$\begin{aligned} |\kappa_1(i) - \kappa'_1(i)| &\rightarrow 2 |\arccos(x(i) + \sqrt{(1-x(i))^2})| / \pi \\ &= 2 \arccos(1) / \pi = 0, \end{aligned}$$

which completes the proof, since this result holds for all $i \in \mathbb{Z}$. \square

We conclude from Lemma 7.3.7, that not only $i'_P - i_P$ has to be small in order to get small errors in κ_1 , but also $\lambda_1(i_P) > \lambda_2(i_P)$ and u has to be a signal of bounded variation, i.e., u should satisfy the sufficient condition in (7.14). Note that only the last condition holds automatically due to the constraints within the problem.

Finally we derive an expression for the error in κ_3 due to an incorrectly determined P-phase arrival time.

Lemma 7.3.8 Let κ'_3 be as defined in (7.8) with u'_L, u'_Q and u'_T substituted for u_L, u_Q and u_T respectively. Furthermore, we define $\gamma = v_1(i_P) - v_1(i'_P)$. Then

$$\|\kappa_3 - \kappa'_3\|_\infty \leq \|\gamma v_1(i_P)^T + v_1(i_P) \gamma^T + \gamma \gamma^T\|_2. \quad (7.15)$$

Proof

Following (7.9) we write

$$\kappa_3(i) = \sum_{n=i}^{i+N-1} (V(i_P)^T Q V(i_P) u, u) / \sum_{n=i}^{i+N-1} (u, u),$$

with Q the orthogonal projection onto $\text{span}\{e_2, e_3\}$. Using this notation we derive in a straight forward way

$$\begin{aligned} & |\kappa_3(i) - \kappa'_3(i)| \\ &= \left| \sum_{n=i}^{i+N-1} (V(i_P)^T Q V(i_P) - V(i'_P)^T Q V(i'_P) u(n), u(n)) / \sum_{n=i}^{i+N-1} (u(n), u(n)) \right| \\ &\leq \|V(i_P)^T Q V(i_P) - V(i'_P)^T Q V(i'_P)\|_2 \\ &= \|V(i_P)^T P V(i_P) - V(i'_P)^T P V(i'_P)\|_2 \\ &= \|v_1(i_P) v_1(i_P)^T - v_1(i'_P) v_1(i'_P)^T\|_2, \end{aligned}$$

using $P = I - Q$. Since this upper bound holds for any $i \in \mathbb{Z}$, we can take the supremum of $|\kappa_3(i) - \kappa'_3(i)|$ over all $i \in \mathbb{Z}$. The proof is completed by substituting $v_1(i'_P) = v_1(i_P) - \gamma$. \square

We observe that the upper bound in (7.15) is sharp, so that (7.15) is a good estimation for the error in the ratio of transversal to total energy, due to an incorrect P arrival time determination.

7.3.2 Problems in Analysing Seismic Data

When analysing a three-component signal with $\kappa = \kappa_1 \cdot \kappa_2 \cdot \kappa_3$ we have to deal with several problems. First of all the window length N for the matrices $M_{N,u}$ and that in κ_3 has to be determined. Obviously N is related to the frequency spectrum of the signal. By fixing N we do not take into account that a signal can consist of a broad range of frequency contents. Moreover, the window lengths N that are not related to the frequency contents of the signal will introduce undesired spikes in the graph of κ , which has been depicted in Figure 7.2. In this figure we see the seismic signal of Figure 7.1 ($i_P = 1120$, $i_S = 1275$) analysed by κ using two different window lengths, namely $N = 15$ and $N = 30$. Due to an incorrect window length we see in the upper graph a lot of local maxima that are not related to any features of interest in the signal.

Another problem we have to deal with is the following. Ideally κ should attain its global maximum at i_S . Moreover, this maximal value should be close to 1. However in practice

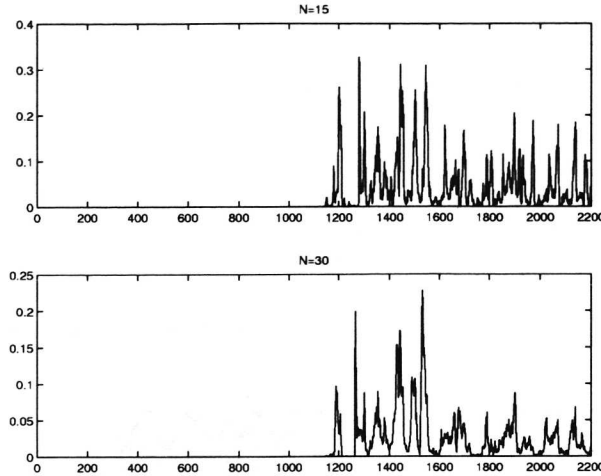


Figure 7.2: The characteristic function κ for the seismic signal in Figure 7.1 ($i_P = 1120$, $i_S = 1275$) for $N = 15$ (top) and $N = 30$ (bottom).

a seismic signal does not only consist of a P-phase and an S-phase, but also other waves, which we did not consider here, appear in the seismogram. Besides, the signal is generally measured with both background noise and signal generated noise. Due to these facts κ will generally not reach a value close to 1 at i_S and even the maximum of κ at i_S can turn out to be a local maximum instead of a global maximum. Therefore we take a threshold value that has to be attained by κ at the S-phase arrival time i_S . In practice it turns out that choosing this threshold value is very difficult. To illustrate this phenomenon we may have another look at Figure 7.2. In both pictures we notice that the global maximum of κ is much less than its ideal value 1. Furthermore, we see that in the second picture κ attains its global maximum at $i = 1540$ while $i_S = 1290$.

7.4 The Wavelet Method

The reliable method Cichowicz introduced in his paper [13] to determine S-phase arrival times uses the product function κ as given in (7.5) to determine i_S as described in Section 7.2. The problems he has to deal with in his paper are exactly the same problems as described in the previous section. To overcome these problems we introduce an algorithm based on both Cichowicz's traditional method and the discrete wavelet transform for $l^2(\mathbb{Z})$.

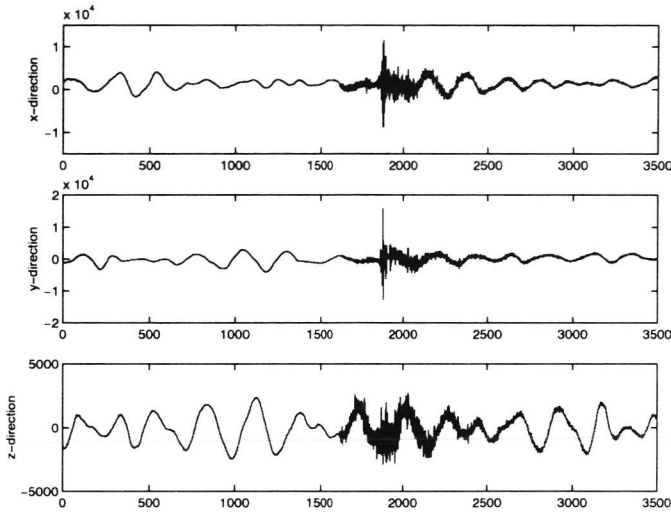


Figure 7.3: A three-component seismogram of a local event containing micro seismic noise.

7.4.1 Characteristic Functions and the DWT

The idea to detect S-phase arrival times using the DWT is as follows. By making a decomposition of the three-component seismic signal $u \in l^2(\mathbb{Z}, \mathbb{R}^3)$ into signals at several scales, it can be possible to separate the S-phase from other phases, that appear at other scales. This decomposition is made by taking the l^2 -DWT of each component of u after rotating the signal into longitudinal and transverse directions, i.e.

$$u^{(m)} = u_L^{(m)} v_1(i_P) + u_Q^{(m)} v_2(i_P) + u_T^{(m)} v_3(i_P).$$

We already know that the frequency spectrum of the P-phase appears at higher frequencies than the S-phase. Therefore generally the S-phase will appear at higher scales than the P-phase. Our aim is now to find a set J_S , representing the collection of scaling levels at which most of the S-phase appears. To make this choice more explicit, we compute the energy in the l^2 -DWT of a segment u^+ of u just after i_P . This will yield information about J_S . The measured seismic signal can also consist of background noise and micro seismic noise, which has been depicted in Figure 7.3. Therefore we have to subtract the energy of the noise. This is done by using a wavelet soft thresholding method similar to Donoho's method [24].

For denoising we assume $u = s + r$ (s : signal consisting of several phase, r : several types of noise). By definition $s(i) = 0$, $i < i_P$. We construct

$$u^+(i) = u(i_P + i), \quad i = 0, \dots, T - 1,$$

$$u^-(i) = u(i_P - i - 1), \quad i = 0, \dots, T - 1,$$

with T such that $T/f_s = 20$, for the sampling frequency f_s . Also we take $u_n^+(i) = 0$ and $u_n^-(i) = 0$ for $0 < i$ and $i \geq T$. So u^+ and u^- are two segments with a duration of 20 seconds before and after the P-phase arrival time. Following [24], the wavelet coefficients $\mathcal{G}_q^*(\mathcal{G}_p^*)^{m-1}u_n^-$ of u_n^- can be used as soft thresholds for $\mathcal{G}_q^*(\mathcal{G}_p^*)^{m-1}u_n^+$. In this method we assume that the noise r is present both before and after i_P . To make this more explicit we define the soft threshold operator Θ on $l^2(\mathbb{Z})$ by

$$\Theta_\eta[\alpha](k) = \text{sgn}(\alpha(k)) \cdot \max(0, |\alpha(k)| - \eta), \quad (7.16)$$

for all $\alpha \in l^2(\mathbb{Z})$ and some $\eta \in \mathbb{R}^+$. Note that the wavelet filters p and q have not been chosen yet. In Section 7.4.2 we will deal with the question of choosing appropriate wavelet filters.

We compute $t(m) = \|\mathcal{G}_q^*(\mathcal{G}_p^*)^{m-1}u_n^-\|_\infty$ for $m \in \mathbb{N}$ and use this as soft threshold for $\mathcal{G}_q^*(\mathcal{G}_p^*)^{m-1}u_n^+$. This yields the sequences

$$w_n^{(m)} = \mathcal{G}_q \mathcal{G}_p^{m-1} \Theta_{t(m)} \mathcal{G}_q^* (\mathcal{G}_p^*)^{m-1} u_n^+, \quad n = 1, 2, 3, \quad m \in \mathbb{N},$$

representing the denoised segment u^+ at scaling level m . Experiments showed that with this denoising method the sequences $w_n^{(m)}$ can become free of micro seismic noise.

After this denoising procedure we can estimate J_S by using the energy distribution of w over all scales

$$E(m) = \sum_{n=1}^3 \|w_n^{(m)}\|_2^2. \quad (7.17)$$

Remark that E tends to zero once a certain scaling level has been reached, cf. Lemma 3.2.6. So we can stop computing the l^2 -DWT for higher scales at such a scaling level. In Figure 7.4 the relative energy distribution as a function of the scaling level has been depicted for both a local event, i.e. an event for which the distance from the source to the measurement equipment is less than 100 km, and a non-local event. In this analysis the Daubechies-4 (Db4) wavelet filter and its corresponding scaling filter have been used, see [23]. Our method also holds for wavelet filters that do not come from orthonormal wavelets, which we showed rigorously in [72].

The estimation of J_S is based on the following physical properties of local and non-local events. Experiments showed that for local events the S-phase arrival time can be noticed within the time period of 20 seconds. For non-local events the difference $i_S - i_P$ will be larger than 20 seconds in general, due to the fact that for these events the traveling distance of the phases, i.e. the distance between the seismic station and the source of the event, is much

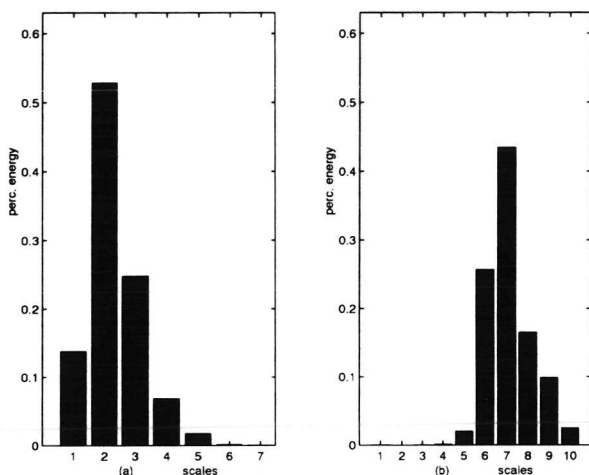


Figure 7.4: Energy distributions as functions of the scale for a) a local event, b) a non-local event.

larger. Another difference between local and non-local events is the frequency behaviour of the phases. Both the P-phase and the S-phase for local events are high frequency signals compared to the P-phase and S-phase of non-local events. Therefore most of the energy of w is found at low scaling levels for local events and at higher scaling levels for non-local events. For local events i_S is contained in u^+ , however in general $i_S - i_P > T$ for non-local events. So only the P-phase can be found in u^+ for non-local events

To analyse the energy distribution we consider the scale m_{max} for which the maximum of E is attained. In case of a local event this maximum will be related to the S-phase, which contains most of the energy in w . For non-local events m_{max} is related to the P-phase. Let us now assume that we know that most of the energy of local events can be found generally at the first m_{loc} scales. Then in general an event can be characterized as a local event if $m_{max} \leq m_{loc}$ and as a non-local event if $m_{max} > m_{loc}$. The parameter m_{loc} is determined experimentally and depends on the wavelet filters. Experiments with the Daubechies-4 wavelets yielded $m_{loc} = 4$.

After looking at the energy distribution E of a seismic event we know with which type of event we are dealing with. Once we have this knowledge, we take

$$J_S = \begin{cases} \{m_{max}, m_{max} + 1\}, & \text{if } m_{max} \leq m_{loc}, \\ \{m_{max} + 1, m_{max} + 2\}, & \text{if } m_{max} > m_{loc}. \end{cases} \quad (7.18)$$

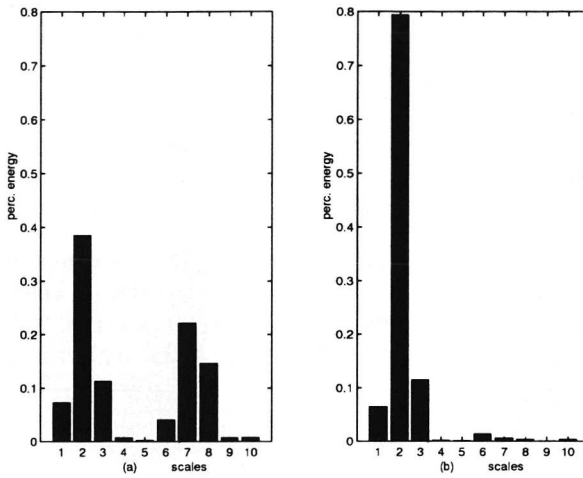


Figure 7.5: Energy distributions as functions of the scale for the local event in Figure 7.3 a) without wavelet denoising, b) with wavelet denoising.

The latter choice is justified by observing that the S-phase will appear at higher scales than the P-phase. Experiments showed that these choices for J_S only lead to scales at which most of the S-phase appears for signals free of micro seismic noise. Micro seismic noise that contains a substantial part of the total amount of the signal's energy may appear at other scales than the S-phase. This phenomenon has been depicted in Figure 7.5.a. Here we see the relative energy distribution of the signal in Figure 7.3 during a 20 second period after i_P . In Figure 7.5.b we see the relative energy distribution of the same signal during the same time period, but now using the wavelet thresholding procedure. The effect of our wavelet thresholding procedure is obvious.

Once J_S has been determined κ can be applied to $u^{(m)}$ for $m \in J_S$ and the S-phase arrival time is then given by

$$\prod_{m \in J_S} (\kappa(u^{(m)}))^2(i) = \max_{n \in N} \prod_{m \in J_S} (\kappa(u^{(m)}))^2(n) \iff i = i_S. \quad (7.19)$$

In practice we have to work with a threshold value that has to be attained at i_S . However by separating the S-phase from the P-phase it will be less difficult to find an appropriate threshold value than in the situation we discussed in Subsection 7.3.2.

Another problem discussed in Subsection 7.3.2 is the window length N which was not related to the frequency behaviour of u in Cichowicz's method. Now that we use a decomposition

of u at several scales, we can use window lengths $N(m)$, i.e. a monotone ascending function of the scaling level m . So the characteristic function κ to analyse $u^{(m)}$ will use the window length $N(m)$ in its definition. In this manner the window length is adapted automatically to the frequency behaviour of the signal. Amongst others, in the next subsection we will discuss the choice we made for $N(m)$.

7.4.2 A Set-Up for the DWT Analysis

In the previous section we discussed our approach to analyse three-component seismic data using characteristic functions and the DWT. Here we will present some choices we made for the parameters in our algorithm after testing the method on seismic data. Note that also other choices for the parameters can be made as long as they fit in the mathematical framework of our algorithm.

Wavelet filters:

The first parameter we discuss is the wavelet function and the associated wavelet filters. Generally, in order to come to an appropriate choice for these filters we induce some constraints on the corresponding wavelet functions related to the physical problem. In our problem the most important constraint is that the wavelet function should match with the seismic data, in such a way that the dilated and translated wavelets generate a good approximation of the several phases at the particular scaling levels. In experiments we minimized for a set of seismic data (also synthetic data) for which i_S is known the error in the determined i_S for a collection of candidate wavelet filters. The experimental results led to the choice of the Daubechies-4 (D_4) wavelet filters.

We observe that also other filters associated to wavelets within the Daubechies family (e.g. D_8, D_{20}) performed very well in this test. These wavelets are much smoother than D_4 , see [23], however in our problem smoothness of the wavelet does not play an important role. The most important property of the chosen wavelet is its matching property with the seismic signals, which are not smooth at all. Furthermore, the wavelet filters of D_8 and D_{20} contain respectively two and five times as many filter coefficients as the D_4 wavelet filters. Using more filter coefficients will increase the computing time of our algorithm slightly.

Window length $N(m)$:

In the previous section we already mentioned a way to adapt the window length N used in the characteristic functions. This adaptation can be related to the frequency behaviour of the seismic signals. Obviously, to analyse u at several scales $m \in \mathbb{N}$ the window length $N(m)$ has to be a monotone ascending function of the scaling level m . In order to come to an appropriate function for $N(m)$ one has to consider two facts. The signal is scaled at each level by the factor 2. Furthermore, for low and high scaling levels $N(m)$ should not become

too small or too big, since $N(m)$ is used to obtain information out of the signal in a certain neighbourhood. These considerations led to the choice

$$N(m) = \lceil 30 \cdot 2^{\max(0, (m-4)/2)} \rceil, \quad (7.20)$$

with $\lceil \cdot \rceil$ the entier function. So for the first four levels we take $N = 30$ (≈ 0.75 seconds at a 40 Hz sampling rate) and thereafter N is multiplied by powers of $\sqrt{2}$.

Threshold value:

To declare an S-phase arrival time we can use (7.19). However, as we discussed already in practice κ will not attain its maximum value κ_{\max} at i_S . Therefore we replace κ_{\max} by a threshold value η for which $0 \leq \eta \leq \kappa_{\max}$. Since κ_{\max} will vary for a set of seismic data we also want to make η adaptive. This can be done by putting

$$\eta = c \cdot \kappa_{\max}. \quad (7.21)$$

Now the problem is left to choose c . Experiments with a substantial number of seismic events showed that the algorithm performed most successfully for $c = 0.2$.

7.5 Examples and Results

In this section we demonstrate our algorithm by means of two examples. To illustrate the difference of our approach compared to Cichowicz's approach, we have plotted both the characteristic function in (7.5) and the characteristic function in (7.19) for two events. For these examples we used the set-up as discussed in the previous section.

After these two examples we present the results of a test done at KNMI. We used both approaches on a set of 313 local events to come to some conclusions.

Example 7.5.1 For the local event in Figure 7.1, we computed the energy distribution along scales (7.17) to obtain $J_S = \{2, 3\}$. For these scales we computed κ using the window length $N = 30$. Also κ was computed without the wavelet method using the window length $N = 30$. Both functions have been depicted in Figure 7.6.

Obviously, in our method we have more freedom to choose the threshold parameter c than in Cichowicz's method. Note, that for automatic phase detection we have to choose a value for c before analysing a seismic event.

For this particular event an analyst at KNMI determined manually $i_P = 1120$ and $i_S = 1275$. In our approach we get $i_S = 1280$ and for Cichowicz's approach we have $i_S = 1295$. Using a 40 Hz sampling rate these results differ 0.125 and 0.5 seconds respectively. Both results are

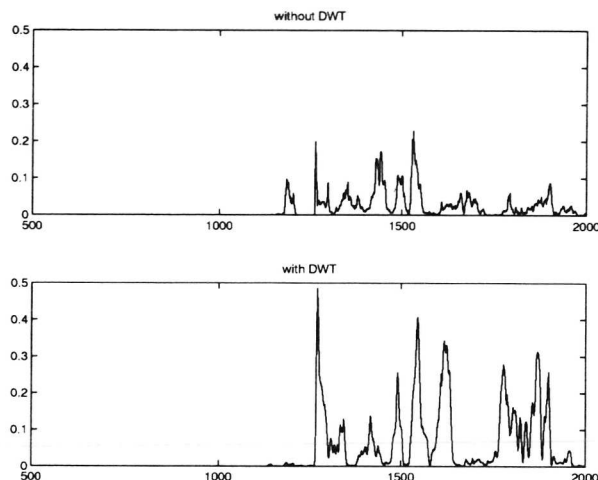


Figure 7.6: The characteristic function κ for the seismic signal in Figure 7.1 ($i_P = 1120$, $i_S = 1275$) using Cichowicz's method (top) and the DWT method (bottom).

acceptable in an automatic procedure.

Finally, we observe that for this event $i_S - i_P = 1275 - 1120 = 155$ ($= 3.875$ seconds). So $i_S - i_P < 20$ seconds which we assumed for local events.

Example 7.5.2 Also for the non-local event in Figure 7.7, we computed the energy distribution along scales (7.17) to obtain $J_S = \{8, 9\}$. We computed κ using the window length $N(8)$ and $N(9)$ at scale $j = 8$ and $j = 9$ respectively. For this event κ was also computed without the wavelet method using the window lengths $N = 60$ and $N = 120$. All three functions have been depicted in Figure 7.8.

We see that in the upper picture the S-phase arrival will not be detected unless we put $0.96 < c \leq 1$, which is not a very realistic choice for this parameter. With a larger window $N = 120$, the S-phase arrival can be detected if we take $0.55 < c \leq 1$, however in our approach we do not have to be that precise with choosing c , since for all $0.15 < c \leq 1$ the S-phase arrival will be detected. Obviously, also here we have more freedom to choose the threshold parameter c than in Cichowicz's method.

For this non-local event an analyst at KNMI determined manually $i_P = 1810$ and $i_S = 11900$. We get $i_S = 11960$ in our wavelet approach and $i_S = 11990$ in Cichowicz's approach. Since also this signal has been sampled at a 40 Hz sampling rate, these results differ

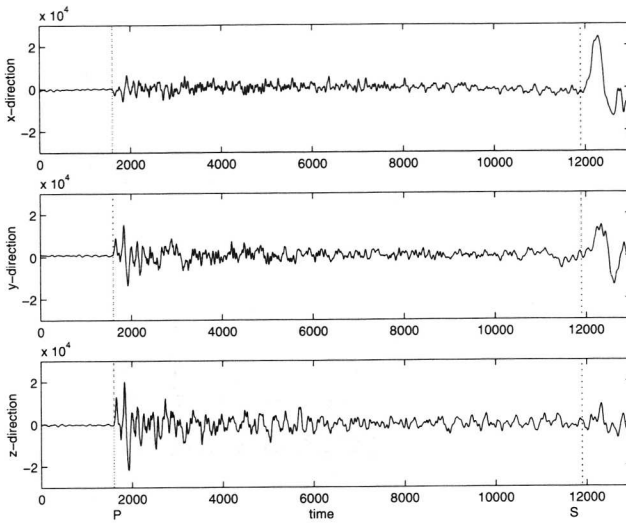


Figure 7.7: A three-component seismogram of a teleseismic event, with P and S arrival time picked by an analyst at $i_P = 1810$ and $i_S = 11900$ respectively.

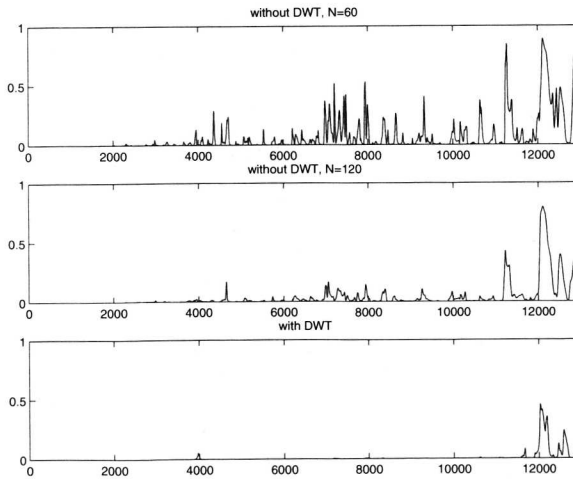


Figure 7.8: The characteristic function κ for the seismic signal in Figure 7.7 ($i_P = 1810$, $i_S = 11900$) using Cichowicz's method for $N = 60$ (top) and $N = 120$ (middle) and the DWT method (bottom) respectively.

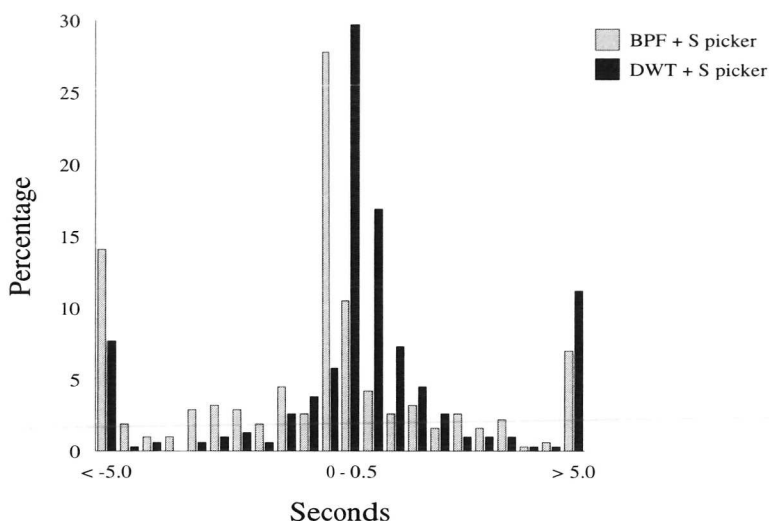


Figure 7.9: Test result for 313 local events (number of S-phase arrival picks within time deviation intervals w.r.t. i_P .)

1.5 and 2.25 seconds respectively. Also in this example our approach improves the automatic pick as compared to the traditional approach.

Finally we observe that for this event $i_S - i_P = 11900 - 1810 = 10090$ (= 252.25 seconds). So $i_S - i_P > 20$ seconds which we assumed for non-local events.

Results:

Our method has been implemented at the Royal Dutch Meteorological Institute (KNMI) and tested for set of 313 local events, using a set-up as described in Section 7.4.2. For each event a manual P-phase arrival pick by an analyst has been used within the algorithm. Furthermore, a manual S-phase arrival pick by an analyst has been used to judge the result of the test.

Also the performance of Cichowicz's method has been tested. For this test we did not use the raw seismic data, but we pre-filtered the data with a 0.6 - 6 Hz Butterworth bandpass filter, see e.g. [79]. The threshold parameter c for this method has been taken to be 0.6, which turned out to be the optimal value for c .

For each event we measured for both methods the difference in i_S determined automatically by the two algorithms compared to the manual picked i_S . The results of this test have been depicted in Figure 7.9. In this figure the number of S-phase arrival picks have been plotted

Method	Deviation i_S : $[-0.5, 0.5]$	Deviation i_S : $[-1.5, 1.5]$
BPF	38.3 %	52.2 %
DWT	46.6 %	68.2 %

Table 7.1: Number of determined S-phases (for 313 local events) for the traditional method and the wavelet based method.

versus the difference in time with respect to the manual picks. This has been done for both Cichowicz's method (light colored bars) and for our new wavelet based method (dark colored bars). Except for the first and last, all bars represent a 0.5 second time interval. The first and the last bars represent the number of picks that differ more than 5 seconds from the manual picked i_S . For both methods these bars represent $\pm 20\%$ of all tested events. These differences are grouped into intervals of 0.5 seconds.

Table 7.1 shows the percentage of S-phase picks within an 0.5 and 1.5 second time difference from the manual picked i_S . We see that for a physically acceptable difference of 1.5 seconds our new wavelet based method substantially improved Cichowicz's method (68.2 % versus 52.2 %).

Note, that we did not compare our method with other wavelet based methods like [5]. The reason for not comparing our method with these other methods is that the results established in these methods do not show substantial improvements of traditional methods in seismics. Moreover, the wavelet based methods that appear in the literature have been tested for at most 23 events. Such test results can not be used to draw serious conclusion.

With some technical adaptations the wavelet based S-phase picking algorithm is expected to become operational for picking S-phases automatically using only one seismic station of KNMI. For this the algorithm will also be tested on a set of non-local events.

Bibliography

- [1] O. Akay and G.F. Boudreaux-Bartels, "New fractional operators and their properties: A generalization of time and frequency operators", in *Proc. IEEE 2nd UK Symp. on Time-Freq. and Time-Scale Methods*, 141-144, 1997.
- [2] K. Aki and P.G. Richards, *Quantitative seismology*, Freeman and Company, San Francisco, 1980.
- [3] R.V. Allen, "Automatic earthquake recognition and timing from single traces", *Bull. Seism. Soc. Am.*, 68, 1521-1532, 1978.
- [4] L.M. Almeida, "The fractional Fourier transform and time-frequency representations", *IEEE Trans. Sig. Proc.*, 42 (11), 3084-3091, 1994.
- [5] K. Anant and F. Dowla, "Wavelet transform methods for phase identification in three-component seismograms", *Bull. Seism. Soc. Am.*, 87(6), 1598-1612, 1997.
- [6] M. Bear and U. Kradolfer, "An automatic phase picker for local and teleseismic events", *Bull. Seism. Soc. Am.*, 77(4), 1437-1445, 1987.
- [7] J.J. Benedetto and M.W. Frazier (eds.), *Wavelets: Mathematics and applications*, CRC Press, London, 1994.
- [8] G. Beylkin, R. Coifman and V. Rohklin, "Fast wavelet transforms and numerical algorithms I", *Comm. Pure Appl. Math.*, 43, 141-183, 1991.
- [9] N.G. de Bruijn, "A theory of generalized functions, with applications to Wigner distributions and Weyl correspondence", *Nieuw Archief voor Wiskunde*, (3) 21, 205-280, 1973.
- [10] A. Calderón, "Intermediate spaces and interpolation, the complex method", *Stud. Math.*, 24, 113-190.
- [11] K. Chandrasekharan, *Classical Fourier transforms*, Springer, Berlin, 1989.

- [12] C.K. Chui, *An introduction to wavelets*, Academic Press, London, 1992.
- [13] A. Cichowicz, "An automatic S-phase picker", *Bull. Seism. Soc. Am.*, 83(1), 180-189, 1993.
- [14] A. Cichowicz, R.W.E. Green and A. van Zyl Brink, "Coda polarization properties of high-frequency microseismic events", *Bull. Seism. Soc. Am.*, 78, 1297-1318, 1988.
- [15] T.A.C.M. Claasen and W.F.G. Mecklenbräuker, "The Wigner distribution - A tool for time-frequency signal analysis", *Philips J. Res.*, 35, 217-250, 1980.
- [16] A. Cohen and R. Ryan, *Wavelets and multiscale signal processing*, Chapman & Hall, London, 1995.
- [17] L. Cohen, "Generalized phase-space distribution functions", *J. Math. Phys.*, 7, 781-786, 1966.
- [18] L. Cohen, *Time-frequency analysis*, Prentice Hall, New Jersey, 1995.
- [19] J.W. Cooley and J.W. Tukey, "An algorithm for the machine calculation of complex Fourier series", *Math. Comp.*, 19, 297-301, 1965.
- [20] I. Daubechies, "Time-frequency localization operators: a geometric phase space approach", *IEEE Trans. Inf. Th.*, 34, 605-612, 1988.
- [21] I. Daubechies, "Orthonormal bases of compactly supported wavelets", *Comm. Pure Appl. Math.*, 41, 909-996, 1988.
- [22] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis", *IEEE Trans. Inf. Th.*, 36, 961-1005, 1990.
- [23] I. Daubechies, *Ten lectures on wavelets*, SIAM, Philadelphia, 1992.
- [24] D.L. Donoho, "De-noising by soft-thresholding", Tech. Rept. Dept. of Statistics, Stanford University, 1992.
- [25] R. Duffin and A. Schaefer, "A class of nonharmonic Fourier series", *Trans. Am. Math. Soc.*, 72, 341-366, 1952.
- [26] H. Dym and H.P. McKean, *Fourier series and integrals*, Academic Press, New York, 1972.
- [27] C. Flammer, *Spheroidal wave functions*, Stanford University Press, Stanford, 1957.
- [28] P. Flandrin, "Maximal signal energy concentration in a time-frequency plane", *Proc. ICASSP 88*, 2176-2179, 1988.

- [29] G. Folland, *Harmonic analysis in phase space*, Princeton University Press, Princeton, 1989.
- [30] E. Foufoula-Georgiou and P. Kumar, *Wavelets in geophysics*, Academic Press, London, 1994.
- [31] J. Fourier, *Analytical theory of heat*, Dover Publications, New York, 1955.
- [32] J. Franklin, *Matrix theory*, Prentice Hall, Englewood Cliffs, 1968.
- [33] G. Frederix, "Integral operators related to symplectic matrices", Tech. Rept. Dept. of Mathematics, Technological University Eindhoven, 1977.
- [34] D. Gabor, "Theory of communications", *J. Inst. Elec. Eng.*, 93, 429-457, 1946.
- [35] R. Goldberg, *Fourier transforms*, Cambridge University Press, Cambridge, 1961.
- [36] T.N.T. Goodman, S.L. Lee and W.S. Tang, "Wavelet bases for a set of commuting unitary operators", *Adv. Comp. Math.*, 1, 109-126, 1993.
- [37] K. Gröchenig, "Analyse multi-échelles et bases d'ondelettes", *C.R.Acad.Sc. Paris*, 305, Série I, 13-15, 1987.
- [38] A. Grossmann, J. Morlet and T. Paul, "Transforms associated to square integrable group representations I: General results", *J. Math. Phys.*, 26, 2473-2479, 1985.
- [39] A. Grossmann, J. Morlet and T. Paul, "Transforms associated to square integrable group representations II: Examples", *Ann. Inst. H. Poincaré Phys. Théor.*, 45, 293-309, 1986.
- [40] P.R. Halmos, *A Hilbert space problem book*, Springer Verlag, Berlin, 1982.
- [41] C. Heil, J. Ramanathan and P. Topiwala, "Asymptotic singular value decay of time-frequency localization operators", in: *Wavelet applications in signal and image processing II: Proc. SPIE San Diego 1994*, 2303, 15-24, 1994.
- [42] C. Heil and D. Walnut, "Continuous and discrete wavelet transforms", *SIAM Review*, 31, 4, 628-666, 1994.
- [43] C. Herley and M. Vetterli, "Wavelets and filter banks: theory and design", *IEEE Trans. Sig. Proc.*, 40, 2207-2232, 1992.
- [44] F. Hlawatsch and G.F. Boudreaux-Bartels, "Linear and quadratic time-frequency signal representations", *IEEE Sig. Proc. Mag.*, 9, 21-67, 1992.
- [45] M. Holschneider, *Wavelets: An analysis tool*, Clarendon Press, Oxford, 1995.

- [46] A.J.E.M. Janssen, "On the positivity of weighted Wigner distributions", *SIAM J. Math. Anal.*, 12, 752-758, 1981.
- [47] A.J.E.M. Janssen, "Bilinear phase-plane distribution functions and positivity", *J. Math. Phys.*, 26, 1986-1994, 1985.
- [48] D.C. Jepsen and B.L.N. Kennett, "Three-component analysis of regional seismograms", *Bull. Seism. Soc. Am.*, 80, 2032-2052, 1990.
- [49] C. Johnson, *Numerical solution of partial differential equations by the finite element method*, Cambridge University Press, Cambridge
- [50] A. Jurkevics, "Polarization analysis of three-component array data", *Bull. Seism. Soc. Am.*, 78, 1725-1743, 1988.
- [51] G. Kaiser, *A friendly guide to wavelets*, Birkhäuser, Boston, 1994.
- [52] E. Kanasewich, *Time sequence analysis in geophysics*, The University of Alberta Press, Alberta, 1981.
- [53] F.H. Kerr, "Namias' fractional Fourier transforms on L^2 and applications to differential equations", *J. Math. Anal. Appl.*, 136, 404-418, 1988.
- [54] T.H. Koornwinder (ed.), *Representations of locally compact groups with applications*, MC Syllabus 38.2, SMC, Amsterdam, 1980.
- [55] T.H. Koornwinder, in: *Wavelets: An elementary treatment of theory and applications*, T.H. Koornwinder (ed.), 27-48, World Scientific, Singapore, 1993.
- [56] P. Lancaster and M. Tismenetsky, *The theory of matrices*, Academic Press, London, 1985.
- [57] H.J. Landau & H.O. Pollack, "Prolate spheroidal wave functions, Fourier analysis and uncertainty II", *Bell System Tech. J.*, 40, 65-84, 1961.
- [58] H.J. Landau & H. Widom, "Eigenvalue distribution of time and frequency limiting", *J. Math. Anal. and Appl.*, 77, 469-481, 1980.
- [59] A. Louis, P. Maass and A. Rieder, *Wavelets: Theory and applications*, Wiley & sons, Chichester, 1997.
- [60] S. Mallat, "A theory for multiresolution signal decomposition: the wavelet representation", *IEEE Trans. Pat. Anal. & Mach. Intell.*, 11, 674-693, 1989.
- [61] A.C. McBride and F.H. Kerr, "On Namias' fractional Fourier transform", *IMA J. Appl. Math.*, 39, 159-175, 1987.

- [62] Y. Meyer, *Wavelets and operators*, Cambridge University Press, Cambridge, 1992.
- [63] H.G. ter Morsche and P.J. Oonincx, "On the integral representations for metaplectic operators", *to appear*.
- [64] P.M. Morse and H. Feshbach, *Methods of Theoretical Physics*, McGraw-Hill, London, 1953.
- [65] J.E. Moyal, "Quantum mechanics as a statistical theory", *Proc. Camb. Phil. Soc.*, 45, 99-124, 1949.
- [66] R. Murenzi, "Ondelettes multidimensionnelles et applications à l'analyse d'images", PhD Thesis, Université Cath. Louvain, 1990.
- [67] D. Mustard, "The fractional Fourier transform and the Wigner distribution", *J. Austral. Math. Soc. Ser. B*, 38, 209-219, 1996.
- [68] V. Namias, "The fractional order Fourier transform and its application to quantum mechanics", *J. Inst. Maths Appl.*, 25, 241-265, 1980.
- [69] G. Olmo and L. Lo Presti, "Applications of wavelet transform for seismic activity monitoring", in *Wavelets: Theory, Algorithms and Applications*, C. Chui, L. Montefusco & L. Puccio (eds.), Academic Press, London, 1994.
- [70] P.J. Oonincx, "A Wavelet Method for Detecting S-Waves in Seismic Data", *Computational Geosciences*, 3, 111-134, 1999.
- [71] P.J. Oonincx, "The Discrete Wavelet Transform as a Tool for Automatic Phase Pickers", *Proc. IEEE-SP Int. Symp. on Time-Freq. and Time-Scale Methods '98*, 201-204, 1998.
- [72] P.J. Oonincx, "Automatic phase detection in seismic data using the discrete wavelet transform", CWI-Report PNA-R9811, 1998.
- [73] P.J. Oonincx, "On time-frequency analysis and time-limitedness", CWI-Report PNA-R9720, 1997.
- [74] P.J. Oonincx, "A study on MRA and Riesz systems", Master's Thesis Eindhoven University of Technology, Eindhoven, 1995.
- [75] P.J. Oonincx and H.G. ter Morsche, "The fractional Fourier transform and its application to energy localisation problems", *to appear*.
- [76] P.J. Oonincx and S.J.L. Van Eijndhoven, "Frames, Riesz systems and MRA in Hilbert spaces", *Indag. Math.*, 10 (3), 369-382, 1999.

- [77] H.M. Ozaktas and D. Mendlovic, "Fourier transforms of fractional order and their optical interpretation", *Opt. Commun.*, 101, 163-169, 1993.
- [78] H.M. Ozaktas and D. Mendlovic, "Fractional Fourier transformations and their optical implementation II", *Opt. Soc. Am. A*, 10, 2522-2531, 1993.
- [79] A. Papoulis, *Signal Analysis*, McGraw-Hill, New York, 1977.
- [80] H.O. Pollack & D. Slepian, "Prolate spheroidal wave functions, Fourier analysis and uncertainty I", *Bell System Tech. J.*, 40, 43-64, 1961
- [81] J. Ramanathan and P. Topiwala, "Time-frequency localization via the Weyl correspondence", *SIAM J. Math. Anal.*, 24 (5), 1378-1393, 1993.
- [82] A.W. Rihaczek, "Signal energy distribution in time and frequency", *IEEE Trans. Inf. Th.*, 14, 369-374, 1968.
- [83] O. Rioul and M. Vetterli, "Wavelets and signal processing", *IEEE Sig. Proc.*, 8, 14-38, 1991.
- [84] W. Rudin, *Functional analysis*, McGraw-Hill, New York, 1973
- [85] J. Samson and J. Olson, "Some comments on the descriptions of the polarization states of waves", *Geophys. J. R. astr. Soc.*, 61, 115-129, 1980.
- [86] W. Schempp, "On the Wigner quasiprobability distribution function I", *C.R. Math. Rep. Acad. Sci. Canada*, 4 (6), 353-358, 1982.
- [87] W. Schempp, "On the Wigner quasiprobability distribution function II", *C.R. Math. Rep. Acad. Sci. Canada*, 5 (1), 3-8, 1983.
- [88] L. Schumaker, *Spline functions: basic theory*, Wiley, New York, 1981.
- [89] L. Schwartz, *Théorie des distributions I*, Hermann, Paris, 1950.
- [90] L. Schwartz, *Théorie des distributions II*, Hermann, Paris, 1951.
- [91] R. Sleeman and T. van Eck, "Robust automatic P-phase picking: An on-line implementation in the analysis of broad-band seismogram recordings", *Phys. Earth Planetary Interior*, 113, 45-55, 1999.
- [92] D. Slepian, "Some comments on Fourier analysis, uncertainty and modelling", *SIAM Review*, 25 (3), 379-393, 1983.
- [93] D. Slepian, "Some asymptotic expansions for prolate spheroidal wave functions", *J. Math. and Physics*, 44, 99-140, 1965.

- [94] E. Stein and G. Weiss, *Introduction to Fourier analysis on Euclidean spaces*, Princeton University Press, Princeton, 1971.
- [95] G. Strang and T. Nguyen, *Wavelets and filter banks*, Wellesley, Cambridge, 1996.
- [96] N.M. Temme, *Special functions: An introduction to the classical functions of mathematical physics*, Wiley, New York, 1996.
- [97] D.J. Thomson, "Spectrum estimation and harmonic analysis", *Proc. IEEE*, 70, 1055-1096, 1982.
- [98] V.S. Varadarajan, *Lie groups, Lie algebras, and their representations*, Springer, New York, 1984.
- [99] J. Ville, "Théorie et applications de la notion de signal analytique", *Cables et Transmissions*, 2A, 61-74, 1948.
- [100] N.R. Wallach, *Symplectic geometry and Fourier analysis*, Math Sci Press, Brookline, 1977.
- [101] J. Wang and T. Teng, "Identification and picking of S phase using an artificial neural network", *Bull. Seism. Soc. Am.*, 87(5), 1140-1149, 1997.
- [102] F.W. Warner, *Foundations of differentiable manifolds*, Springer, New York, 1983.
- [103] G. Warner, *Harmonic analysis on semi-simple Lie groups*, Springer, Berlin, 1972.
- [104] E.P. Wigner, "On the quantum correction for thermodynamic equilibrium", *Phys. Review*, 40, 749-759, 1932.
- [105] E.P. Wigner, in: *Perspectives in quantum theory*, W. Yourgrau and A. van der Merwe (eds.), 25-36, MIT Press, New York, 1971.
- [106] P. Wojtaszczyk, *A mathematical introduction to wavelets*, Cambridge University Press, Cambridge, 1997.
- [107] M.W. Wong, *Weyl transforms*, Springer, New York, 1998.
- [108] R.M. Young, *An introduction to nonharmonic Fourier series*, Academic Press, New York, 1980.
- [109] A.C. Zaanen, *Continuity, integration and Fourier theory*, Springer, Berlin, 1989.

Index

- admissibility condition, 47
- admissible, 47, 54
- affine transformation, 96

- band-limited, 24
- Bruhat decomposition, 105
- butterfly algorithm, 3

- Caldéron's reproducing formula, 9
- characteristic function (seismic), 140
 - stability, 142
- chirp multiplication, 91, 99
- Cohen's class, 7
- compression, 14
- continuous wavelet transform, 12, 45
 - range of, 53
 - reconstruction formula, 52
- convolution, 23
- cross-power matrix, 139

- Daubechies wavelet, 11
- decomposition algorithm, 59
- deflection angle, 140
- degree of polarization, 140
- denoising, 13, 148
 - wavelet method, 149
- discrete wavelet transform, 12, 55
 - for l^2 , 65, 138, 148

- energy concentration, 120, 125
- energy ratio, 141
- exponential type, 24

- fast Fourier transform, 3

- filter bank, 60
- Fourier series, 2
- Fourier transform, 3, 19
 - discrete, 3
 - on L^1 , 20
 - on L^2 , 22
- fractional Fourier transform, 9, 89, 127
 - generalized, 106, 130
- frame, 68, 84
 - bounds, 68
 - dual, 70
 - exact, 70
 - generator, 70
 - tight, 70
- Fredholm equation, 121

- Gabor transform, 6, 27
- generating sequence, 75
- group
 - affine-linear, 53
 - Heisenberg, 42, 100
 - orthonormal symplectic, 107
 - symplectic, 100

- Haar wavelet, 11, 47, 48
- Heisenberg's uncertainty relation, 4
- Hermite functions, 127, 135

- index law, 90

- Laguerre polynomial, 128
- Lie group, 42, 53

- marginals, 41

- Mexican hat, 12, 47, 49
- Meyer wavelet, 12
- mixed Wigner distribution, 33
- Moyal's formula, 39
- multiresolution analysis, 13, 56, 73
 - generator, 73
- noise
 - background, 148
 - micro seismic, 148
- P-phase, 137
 - arrival time, 140
- Paley-Wiener theorem, 25
- Parseval's formula, 23
- Plancherel's formula, 23
- prolate spheroidal wave functions, 122, 134
- pyramid algorithm, 13, 61
- reconstruction algorithm, 60
- representation, 42
 - equivalent, 43
 - irreducible, 43
 - Schrödinger, 100
 - unitary, 43
- Riemann-Lebesgue theorem, 20
- Riesz basis, 56, 72, 82
- Riesz system, 71, 77, 81
 - bounds, 71
 - dual, 72, 77
 - generator, 71
- Rihaczek distribution, 7, 132
- rotation, 95
- S-phase, 137
 - arrival time, 140
- S-plane, 138, 141
- scale relation, 57
- scale sequence, 57
- scaling function, 57
- Schwartz class, 3, 22
- seismic event
 - local, 149, 153
 - non-local, 149, 154
- seismogram, 137
- short-time Fourier transform, 25
- Slepian's conjecture, 122
- Smith normal form, 78
- spectral analysis, 3
- spectrogram, 4, 28
- time-limited, 24, 120
- topological group, 42
- wavelet, 11, 47
 - filter, 60, 152
 - frame, 55
- wavelet-Galerkin method, 15
- Weyl correspondence, 130
- Weyl transform, 126
- Wigner distribution, 6, 33
 - characteristic function of, 36, 43
- Wigner plane, 33, 125
- windowed Fourier transform, 4, 25
 - inversion formula, 30
 - range of, 33
- Young's inequality, 21

Samenvatting / Dutch Summary

Fourieranalyse is een beproefde methode om een indruk te krijgen van de aanwezige frekwenties in een tijdsafhankelijk signaal. Echter deze methode faalt bij het simultaan analyseren van een signaal in tijd en frekwentie, hetgeen voor steeds meer toepassingen wenselijk is. Het falen van de Fourieranalyse is te wijten aan het ontbreken van localisatie-eigenschappen. Zo zullen veranderingen van het signaal in een specifiek tijdsinterval uitgesmeerd worden over het totale spectrum van frekwenties.

Diverse methoden zijn in het verleden reeds aangedragen om met deze onvolkomenheid om te gaan. Een natuurlijke oplossing voor dit probleem is het gebruik van Fourieranalyse op een door een windowfunctie uitgesneden stukje signaal. Door het verschuiven van deze windowfunctie wordt er steeds op een ander stukje signaal naar de aanwezige frekwenties gekeken. Een probleem bij deze methode, de windowed Fouriertransformatie (WFT), is echter dat de windowfunctie vast gekozen dient te worden, terwijl het signaal zich wellicht beter zou lenen voor een analyse middels een combinatie van kleinere en grotere stukken van het signaal.

Een tweede veelgebruikte methode om signalen simultaan in tijd en frekwentie te kunnen bestuderen is de zogenaamde Wigner distributie. Dit is een bi-lineaire transformatie, die bestaat uit een Fourieranalyse van een auto-correlatie function van het te bestuderen signaal rondom een aan te geven tijdstip. Een nadeel van deze methode is de bi-lineariteit, waardoor de analyse van een superpositie van signalen bemoeilijkt wordt. Zowel de Wigner distributie als de WFT zijn rechtstreeks te relateren aan de Heisenberggroep. Deze en andere eigenschappen van de WFT en de Wigner distributie worden besproken in Hoofdstuk 2.

Recentelijk is een transformatie aangedragen die een signaal in tijd en schalingsgedrag analyseert, de continue wavelettransformatie (CWT). Het principe van deze transformatie lijkt op de WFT, echter de Fouriertransformatie die optreedt bij de WFT is bij de CWT vervangen door een dilatatie van de windowfunctie, wavelet genaamd. Zo kan deze wavelet geschaald worden, waardoor de transformatie zich met name lokaal beter aan het signaal kan aanpassen. Een andere kijk op deze transformatie leert ons dat de CWT een signaal ontbindt in getransleerde en gedilateerde wavelets. Een discretisatie van de CWT resulteert in de discrete wavelettransformatie (DWT). Voor het berekenen van deze DWT bestaat een snel algoritme, de pyramide-algoritme. De wavelettransformatie en bijbehorende algoritmen komen aan bod in Hoofdstuk 3.

Hoofdstuk 4 borduurt voort op de DWT. Het concept multiresolutie-analyse dat leidt tot de pyramide-algoritme wordt in dit hoofdstuk functionaalanalytisch beschreven voor 'signalen' in een algemene Hilbertruimte. Hierbij wordt het probleem van het vinden van geschikte waveletfuncties vertaald naar het vinden van basisvectoren in de ruimte $l^2(\mathbb{Z}^n)$ en in een later stadium naar het vinden van functies op de n -dimensionale eenheidscirkel.

In Hoofdstuk 5 wordt een zeer recentelijk populair geworden tijd-frekwentie operator ingevoerd, de fractionele Fouriertransformatie (FRFT). De kracht van deze FRFT blijkt uit het effect in het fasevlak. Het blijkt namelijk, dat het uitvoeren van de FRFT op een signaal, gevolgd door het bepalen van de Wigner distributie identiek is aan het bepalen van de Wigner distributie van een signaal, gevolgd door een rotatie in het fasevlak. Dit heeft geleid tot de vraag welke transformaties leiden tot affiene transformaties in het fasevlak door uitvoering van de Wigner distributie. In Hoofdstuk 5 wordt een volledige klassifikatie van dergelijke transformaties gepresenteerd.

Een deelklasse van de transformaties zoals deze beschreven zijn in Hoofdstuk 5 kunnen gezien worden als generalizaties van de FRFT. Deze worden gebruikt om een tweetal convexe geometrieën in het fasevlak, rechthoek en cirkel, d.m.v. een lineaire transformatie om te vormen. Een toepassing van deze techniek wordt gegeven door het probleem van het concentreren van energie binnen een bepaalde geometrie in het fasevlak. In Hoofdstuk 6 wordt voor beide geometrieën het klassieke probleem van energielocalizatie besproken, waarna de koppeling met andere geometrieën gemaakt wordt d.m.v. de gegeneralizeerde FRFT.

Hoofdstuk 7 bevat een praktisch probleem, waarbij de DWT gebruikt is. Het betreft het automatisch detecteren van zogenaamde S golven in seismische data. Deze S golven zijn normaliter in een seismogram vermengd met andere significante golven gecombineerd met diverse typen ruis. Een op de DWT gebaseerd algoritme wordt ter oplossing van dit probleem gepresenteerd en op stabiliteit onderzocht. Verder worden testresultaten van de nieuwe algoritme gepresenteerd. Het werk beschreven in dit hoofdstuk is uitgevoerd in nauwe samenwerking met de afdeling Seismologie van het KNMI.

Dankwoord/Acknowledgements

Diverse mensen ben ik dank verschuldigd voor hun inspiratie, hulp, samenwerking en vooral vriendschap, waarvan ik de afgelopen jaren heb mogen genieten. Een aantal personen wil ik hierbij met name noemen.

Allereerst de leider van mijn promotie-onderzoek, dr. Nico Temme, die mij niet alleen met raad en daad heeft bijgestaan, doch ook een persoonlijke steun voor me is geweest in mijn privé leven. Verder heeft de vrijheid, die ik van hem voor mijn onderzoek gekregen heb, een belangrijke bijdrage geleverd aan de tot stand gekomen onderzoeksresultaten. In dezelfde adem wil ik ook prof. dr. Tom Koornwinder noemen, die mij met name de laatste twee jaar van mijn promotietijd uitermate daadkrachtig heeft bijgestaan als promotor. Mede door zijn suggesties en opbouwende kritiek zijn de artikelen waaruit dit proefschrift is opgebouwd geworden tot wat ze zijn.

Een verder woord van dank is gericht aan Rob Zuidwijk en Paul de Zeeuw, die mij als directe collega's binnen mijn onderzoek altijd van dienst zijn geweest met suggesties en adviezen. Met name wil ik Rob bedanken voor zijn vriendschap en collegialiteit tijdens de buitenlandse bezoeken, die we samen hebben afgelegd.

Een belangrijke bijdrage bij het tot stand komen van dit proefschrift is geleverd door de mede-auteurs van een aantal artikelen, te weten Stef van Eijndhoven en Hennie ter Morsche. Niet alleen hiervoor wil ik hen bedanken, maar ook voor het mij enthousiast maken en houden voor het onderwerp wavelettheorie, nu reeds zo'n zes jaar geleden.

Voor één artikel ben ik veel dank verschuldigd aan de wetenschappers van de afdeling Seismologie van het KNMI, met name Bernard Dost, Torild van Eck en Reinoud Sleeman. De samenwerking met hen was buitengewoon aangenaam en vriendschappelijk, en heeft enkele mooie vruchten afgeworpen.

Tot slot ben ik heel veel dank verschuldigd aan mijn verloofde Birgit. Haar liefde en steun in de afgelopen jaren vormden voor mij een krachtbron die mij steeds opnieuw geïnspireerd heeft dit proefschrift te voltooien. Mijn grootste inspiratiebron heb ik ook aan haar te danken, namelijk ons onlangs geboren dochtertje Sacha.

Curriculum Vitae

Op 27 juni 1972 A.D. werd ik te Kerkrade geboren. Opgegroeid ben ik in Sittard waar ik van 1984 tot 1990 het gymnasium van het College Sittard bezocht. Na het met lof behalen van het VWO diploma, ben ik in september 1990 begonnen aan de studie technische wiskunde aan de Technische Universiteit Eindhoven. Binnen deze studie koos ik voor de afstudeerrichting functionaalanalyse in het bijzonder toegepast op wiskundige signaalverwerkingsmethoden.

Tijdens deze periode heb ik mij ook buiten het curriculum technische wiskunde verdiept in signaalverwerkingsmethoden onder meer d.m.v. een externe stage op het gebied van filterbanken bij het Instituut voor Perceptie Onderzoek te Eindhoven. Het doctoraalexamen Technische Wiskunde werd in december 1995 cum laude afgelegd met een doctoraalscriptie onder begeleiding van dr. ir. S.J.L. van Eijndhoven.

Vanaf januari 1996 ben ik werkzaam geweest als onderzoeker in opleiding aan het Centrum voor Wiskunde en Informatica (CWI) te Amsterdam. Het onderzoeksproject, dat zijn neerslag kent in deze dissertatie, betrof het analyseren en uitbreiden van signaalverwerkingsmethoden ten behoeve van geofysische signalen. Een prominente rol binnen dit onderzoek was gereserveerd voor de wavelettransformatie. Tevens is er in samenwerking met de afdeling Seismologie van het KNMI te De Bilt een algoritme ontwikkeld ten behoeve van de analyse van seismogrammen.

Het onderzoek aan het CWI is uitgevoerd onder leiding van dr. N.M. Temme. De voor u liggende dissertatie is geschreven onder supervisie van prof. dr. T.H. Koornwinder, hoogleraar analyse, in het bijzonder Lie groepen en speciale functies, aan de Universiteit van Amsterdam.

- VII -

Als het mogelijk aanwezige aardgas in de Waddenzee net zo brandbaar is als het nu explosief is voor het kabinet Kok-II, dan zit Nederland na winning van dit gas er nog vele jaren warm bij.

- VIII -

De oudste niet-commerciële activiteit in een 24-uurs economie is de zorg van een moeder voor haar pasgeboren kind.

- IX -

Hoe groter een wetenschappelijk congres is opgezet, hoe kleiner is het wetenschappelijk rendement voor de individuele deelnemer.

- X -

Het leven is geen wiskunde.

STELLINGEN

behorende bij het proefschrift

Mathematical Signal Analysis:

Wavelets, Wigner distribution and a seismic application

van

Peter J. Oonincx

Amsterdam, 9 februari 2000

- I -

Gegeven de fractionele Fourier operatoren \mathcal{F}_α en \mathcal{F}_β , met $\alpha, \beta \in [-\pi, \pi]$, dan geldt de onzekerheidsrelatie

$$\int_{\mathbb{R}} u^2 |\mathcal{F}_\alpha[f](u)|^2 du \int_{\mathbb{R}} v^2 |\mathcal{F}_\beta[f](v)|^2 dv \geq \frac{|\sin(\alpha - \beta)|}{4} \|f\|_2^4, \quad \forall f \in L^2(\mathbb{R}).$$

- II -

Zij D de diagonaalmatrix $D = \text{diag}(d_1, \dots, d_n)$, $d_i \in \mathbb{Z}$, met $\det D = N$, waarbij $N > 0$. Zij verder $\{\beta_0, \dots, \beta_{N-1}\}$ een verzameling rijen in $l^2(\mathbb{Z}^n)$, zodanig dat de verzameling

$$B_N = \{\mathcal{R}^{Dk} \beta_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\}$$

een tight frame met frame bound M is. Dan is de verzameling

$$B'_N = \{\mathcal{R}^{D'k} \beta_l \mid l = 0, \dots, N-1, k \in \mathbb{Z}^n\},$$

met $D' = \text{diag}(d'_1, \dots, d'_n)$, $d_i \in \mathbb{Z}$, en d'_i een deler van d_i , ook een tight frame met frame bound M .

Hoofdstuk 4 van dit proefschrift.

- III -

Zij A een lineaire transformatie in \mathbb{R}^n en V een lineaire deelruimte van \mathbb{R}^n . Dan geldt $\dim A^T(V^\perp) = \dim V^\perp$, dan en slechts dan als $\dim A^\perp(V) = \dim V$.

Hoofdstuk 5 van dit proefschrift.

- IV -

Gegeven de operatoren

$$\mathcal{B}(\omega_0)[f](x) = \sqrt{\frac{2}{\pi}} \int_{\mathbb{R}} \frac{\sin(\omega_0(x-u))}{(x-u)} f(u) du,$$

$$\mathcal{P}(x_0)[f](x) = f(x) \cdot \chi_{[-x_0, x_0]}(x),$$

voor alle $f \in L^2(\mathbb{R})$ en voor vaste $x_0, \omega_0 > 0$, dan geldt

$$\text{Ker}(\lambda - \mathcal{P}(x_0)\mathcal{B}(\omega_0)\mathcal{P}(x_0)) = L^2([-x_0, x_0]) \cap \left\{ \frac{1}{\lambda-1} \mathcal{B}(\omega_0) + \frac{1}{\lambda} (\mathcal{I} - \mathcal{B}(\omega_0)) \right\} L^2([-x_0, x_0])^\perp,$$

voor $\lambda \in (0, 1)$.

Hoofdstuk 6 van dit proefschrift.

- V -

De wavelettransformatie maakt het mogelijk de algorithmen van Cichowicz uit te breiden van een methode voor het automatisch bepalen van S-golven in locale events tot een methode voor het automatisch bepalen van S-golven in zowel locale als niet-locale events.

A. CICHOWICZ, "An automatic S-phase picker", *Bull. Seism. Soc. Am.*, 83, 180-189, 1993.

Hoofdstuk 7 van dit proefschrift.

- VI -

Reeds jaren voordat Candes de Ridgelettransformatie introduceerde, gebruikten Facqi, Kabir en Verschuur deze met succes als filteroperatie in de geofysica.

E.J. CANDES, "Ridgelets: Theory and Applications", PhD Thesis, Stanford University, 1998.

L. FACQI, M.M.N. KABIR EN D.J. VERSCHUUR, "Seismic processing using the wavelet transform and the Radon transform", *J. Seism. Expl.*, 4, 375-390, 1995.

