

Atomic Decompositions

Methods for the analysis and synthesis of acoustical signals

J.T. van Velthoven

Bachelor's Thesis

Institute of Sonology

May 2015

To my father, Max van Velthoven (1967-2014).

Abstract

This thesis is concerned with methods for atomic decompositions, methods that allow a representation of a signal as a linear combination of elementary functions, and their application as methods for the analysis and synthesis of acoustical signals. In this consideration both specific methods and a general framework for atomic decompositions will be treated. Examples of these specific methods for atomic decompositions are Fourier theory, Gabor theory and wavelet theory. The general framework for atomic decompositions which will be treated is frame theory. The specific methods for atomic decompositions give rise to concrete examples of methods for sound analysis and synthesis, whereas the general framework can be seen as the foundations of these specific methods for sound analysis and synthesis.

The first part of this thesis constitutes an exposition of several specific methods for atomic decompositions and its general framework. In this part of the thesis the specific methods for atomic decompositions will be introduced first. After that, these specific methods will be considered as special cases of the general framework for atomic decompositions. The second part of the thesis deals with the applications of the methods that are introduced in the first part for the analysis, synthesis and analysis-resynthesis of acoustical signals. These applications are studied from the viewpoint of the general framework for atomic decompositions.

Contents

Abstract	i
Preface	iv
Notation	v
Introduction	vi
Prologue	ix
I Methods	1
1 Fourier theory	2
1.1 Time domain	2
1.2 Frequency domain	6
1.3 Operators associated with Fourier transforms	8
2 Fourier uncertainty principles	10
2.1 The Weyl-Heisenberg inequality	10
2.2 Slepian-Pollack inequality	12
3 Gabor theory	15
3.1 Short-time Fourier transform	15
3.2 Time-frequency resolution of the short-time Fourier transform	18
3.2.1 Examples	20
3.3 Time-frequency representations based on the short-time Fourier transform	22
4 Wavelet theory	26
4.1 Wavelet transform	26
4.2 Time-frequency resolution of the wavelet transform	29
4.2.1 Examples	30
4.3 Time-frequency representations based on the wavelet transform	33
5 Frame theory	36
5.1 General frame theory	37
5.2 Gabor frames	39
5.3 Wavelet frames	43
5.4 Non-stationary Gabor frames	45
5.4.1 Non-stationary Gabor frames in time domain	45
5.4.2 Non-stationary Gabor frames in frequency domain	47

II Analysis, Synthesis and Analysis-Resynthesis	51
6 Analysis	52
6.1 Linear analysis	54
6.1.1 Linear analysis with predetermined analysing functions	54
6.1.2 Linear analysis with predetermined expansion functions	54
6.2 Examples	56
7 Synthesis	60
7.1 Linear synthesis	61
7.1.1 Linear synthesis with predetermined expansion functions	62
7.1.2 Linear synthesis with predetermined analysing functions	62
7.2 Non-linear synthesis	64
7.2.1 Algorithms for non-linear synthesis	65
7.3 Examples	67
8 Analysis-Resynthesis	70
8.1 Linear analysis-resynthesis	71
8.1.1 Combination of linear analysis and synthesis mappings	71
8.1.2 Modification of transform coefficients	72
8.2 Examples	73
Epilogue	77
Conclusion	80
References	81
A Atomic decompositions of square-integrable functions	84
B Contents of accompanying CD	97

Preface

This thesis is submitted in partial fulfilment of the requirements for obtaining the B.M. Sonology. The thesis is the result of the research I have done during the bachelor study of the Institute of Sonology, The Hague. The research was carried out from September 2011 until May 2015. The supervisor of the research was my mentor Peter Pabon.

The research that I carried out during the first year of the bachelor Sonology consisted of a small report called *Timbre: An overview of the main physical and perceptual aspects*. The main conclusion of this report was that timbre is determined by both the temporal and the spectral aspects of an acoustical signal. This conclusion was the motivation for the investigation of signal representations in which both temporal and spectral information can be represented. This investigation resulted in two papers called *Time-Frequency and Time-Scale Analysis* and *Atomic Decompositions of Acoustical Signals*, which were written during the second and the third year of the bachelor study, respectively. All the papers were accompanied by a collection of sound examples that supported the conclusions made in the papers. The topics that were treated in the papers written during the second and third year of my bachelor study form the core of this thesis. However, the results that were described during these papers are extensively expanded and revised in this thesis. Furthermore, the topics that are, in my opinion, the most interesting are treated in this thesis for the first time.

I would like to thank my mentor Peter Pabon for a lot of support and discussions. Although our opinions on the topics covered in this thesis sometimes differed greatly, I am sure that our discussions improved the result.

Notation and terminology

A mapping \mathcal{Q} from X to Y is a relation that associates with each argument $x \in X$ one value $y \in Y$, written as $y = Qx$ or $y = Q(x)$. A mapping Q will be notated as

$$Q : X \rightarrow Y, \quad x \mapsto Qx. \quad (1)$$

The set X is called the *domain* of the mapping and the set Y the *range* of the mapping. The first part of (1) denotes that \mathcal{Q} is a mapping from X to Y , that is, \mathcal{Q} maps elements of X into elements of Y . The second part of (1) denotes that \mathcal{Q} maps $x \in X$ to $Qx \in Y$, that is, the value of \mathcal{Q} at x is Qx . Note that the arrows in both parts of the definition of the mapping are different. The arrow \rightarrow is used in order to associate the mapping with its domain and range, and the arrow \mapsto is associated with the value of the mapping.

If the domain and range of a mapping are sets of numbers, then the mapping will be called a *function*. A function will be notated through a lowercase letter from the English alphabet, i.e. f, g, h . A function is thus a mapping that maps a number $x \in X$ into one, and only one, number $y \in Y$. If the domain and range of a mapping are sets of functions, then the mapping will be called an *operator*. An operator will be notated through a capital calligraphic letter, i.e. $\mathcal{F}, \mathcal{G}, \mathcal{H}$. An operator is thus a mapping that maps a function $f \in X$ into one, and only one, function $g \in Y$.

A function whose independent variable denotes time will be called a *signal*. If the domain of a signal is a continuum, e.g. the set of real numbers \mathbb{R} , then the signal will be called a *continuous-time signal*. If the domain of a signal is a discrete set, e.g. the set of natural numbers \mathbb{N} or the set of integers \mathbb{Z} , then the signal will be called a *discrete-time signal*. The quantity that the value of the signal denotes will be undefined, but could, dependent on the preferred interpretation, denote the sound pressure level or voltage.

Introduction

Representing signals as a sum of elementary functions is a common technique for both the analysis and synthesis of signals. In this case, the elementary functions can be considered as “building blocks” for the represented signals. Therefore, they deliver methods for the analysis and synthesis of signals. The main reason why such methods are used for the representation of signals is the fact that it is in general difficult, or even impossible, to analyse or synthesize a signal when it is presented in its standard form. By representing a signal as a sum of elementary functions, the analysis and synthesis of signals becomes easier. This is due to the fact that these elementary functions are in general simple, and therefore better understood than the represented signal itself. The collection of elementary functions is in general derived from a single, fixed function by elementary operations. This single, fixed function is called an *atom*. For this reason, a representation of a signal as a sum of elementary functions is called an *atomic decomposition*.

The main goal of this thesis is a systematic exposition of methods for atomic decompositions, and their possibilities for the analysis and synthesis of acoustical signals. Examples of specific methods for atomic decompositions that will be treated in this thesis are Fourier theory, Gabor theory and wavelet theory. Aside these specific methods for atomic decompositions, also the general theory of frames will be treated. This general theory offers a general framework for the study of atomic decompositions. It offers such a general framework in the sense that the condition that a specific method belongs to the theory of frames is also a guarantee that the specific method leads to an atomic decomposition. Thus, all the specific methods that belong to the theory of frames are therefore methods for atomic decompositions, and vice versa. For the sake of completeness, both the theory associated with the specific methods for atomic decompositions and their interpretation within the general theory of frames will be treated. Both treatments lead to different interpretations of the specific methods for atomic decompositions, which is useful for the analysis and synthesis of acoustical signals.

All the specific methods for atomic decompositions that are treated in this thesis are well-known in signal analysis. Most of them are also well-known in the context of electroacoustic music. Examples of these are Fourier theory, Gabor theory and wavelet theory. For example, these three methods are described in Curtis Roads’ book *Microsound*, but they are even more extensively described in the monographs *Representations of Musical Signals* and *Musical Signal Processing*, edited by De Poli, Piccialli & Roads respectively Roads, Pope, Piccialli & De Poli. A specific method for atomic decompositions that isn’t well-known in the context of electroacoustic music, but that will be treated in this thesis, is the non-stationary Gabor theory. It is remarkable that this theory isn’t well-known in the context of electroacoustic music since the original concepts of non-stationary Gabor theory were developed within this context. These concepts were namely first described in Florent Jaillet’s Ph.D. thesis called *Représentation et traitement temps-fréquence des signaux audionumriques pour des applications de design sonore* (Jaillet, 2005). Although most of the methods for atomic decompositions are thus well-known in the context of electroacoustic music, they aren’t used that often. One of the few examples of a musical composition in which methods from Gabor and wavelet theory are used extensively is Jean-Claude Risset’s *Invisible*, a piece for tape and soprano. In this piece, the methods from Gabor and wavelet theory were used for both the analysis-resynthesis and synthesis of sounds. The main reason that these methods aren’t used that often in electroacoustic music is probably a lack of available software. This can also be deduced from the following quotation:

“It is clear that the [Gabor transform] is a potent means of sound transformation. The work of the Marseilles group [a group to which Jean-Claude Risset was connected] has been especially important in adapting wavelet concepts to the [Gabor transform]. Fortunately for us, their focus is on acoustical problems. Hopefully this research will eventually be made into widely available software tools (Roads, 2004).”

The problem of a lack of available software is probably solved now methods for Fourier theory, Gabor theory, wavelet theory and even non-stationary Gabor theory are freely available in the software toolbox called *Large Time-Frequency Analysis Toolbox* (Søndergaard et al., 2012), a toolbox of which the author is one of the developers. This toolbox is a well-documented collection of routines including several methods for atomic decompositions. Even more important for electroacoustic music is the fact that the toolbox offers the possibility to connect the methods for atomic decompositions with real-time audio processing (Průša et al., 2014). For the sake of keeping this thesis to a reasonable length, the toolbox will not be discussed here. However, the interested reader is strongly recommended to have a look into the possibilities of toolbox, and is therefore referred to the aforementioned papers for a general overview of the toolbox, and for working examples to the tutorial (Van Velthoven & Søndergaard, 2014). This tutorial covers almost all the topics that are also covered in this thesis. It even covers these topics in the same order as this thesis. The tutorial can therefore easily be used aside this thesis.

Most of the specific methods for atomic decompositions are thus well-known in the context of electroacoustic music. However, there are several differences between the treatment of these methods in this thesis and in the aforementioned books. For example, in the aforementioned books, the methods for atomic decompositions are considered as methods for the analysis and analysis-resynthesis of acoustical signals only, but in this thesis these methods are also considered as methods for the synthesis of acoustical signals. However, the main difference between the treatment of these methods in this thesis and in the aforementioned books is that these methods will be treated in this thesis as special cases of frame theory. This is a difference since most of the methods are in general introduced as more or less independent methods. The notion that some of these specific methods for signal analysis and synthesis can be considered as special cases of a more general theory is, however, not entirely new in the context of electroacoustic music. One of the few examples of a study that connects the general concept of an atomic decomposition to acoustical signals is a paper by Goodwin & Vetterli (Goodwin & Vetterli, 1997). In this paper, they give as examples of specific methods for atomic decompositions Fourier series, Gabor series and granular synthesis. However, the general theory of frames isn’t mentioned in this paper. Another study in which the general concept of an atomic decomposition is connected to acoustical signals is a paper by Sturm, Roads, McLeran & Shynk (Sturm et al., 2009). In their paper, the framework of frames isn’t treated explicitly, but the so-called *dictionary-based methods*, which form the core of their paper, are closely related to frames. However, frame theory is more closely related to the general theory of orthogonal expansions, which is described in Stan Tempelaars’ book *Signal Processing, Speech and Music*. To be precise, this theory is described in the small section called *Orthogonal Functions and Signal Transforms* (Tempelaars, 1996, Section 4.5). In this section, it was shown that the elementary functions used in the Fourier and Walsh transform form an orthogonal system. For this reason, they could be treated as special cases of orthogonal expansions. Something that isn’t noted in Tempelaar’s book is that there also exist wavelet systems that are orthogonal, and therefore even some wavelet systems can be considered as special cases of orthogonal expansions. In fact, it could even have been noted that a collection of functions doesn’t necessarily have to be orthogonal in order to get expansions of which the form is similar to orthogonal expansions. It is this notion that forms the core of frame theory — see e.g. the classical papers by Daubechies, Grossmann & Meyer (Daubechies et al., 1986) and Feichtinger & Gröchenig (Feichtinger & Gröchenig, 1992). Since frame theory itself forms the core of this thesis, this thesis can be considered as a continuation and extension of the results described in the section on orthogonal functions in Tempelaars’ book.

By the treatment of the theory in this thesis, the focus is on the motivation behind the existence of the methods for atomic decompositions rather than on their technical details. The thesis is

therefore written in an informal style. However, there is also a limit to what can be said about these methods in this manner. For this reason, there is still a certain amount of mathematical knowledge needed in order to entirely appreciate the content of this thesis. The main prerequisites are the basic mathematical methods and techniques that are also used in the aforementioned book by Tempelaars (Tempelaars, 1996). This book is used in the first two years of the bachelor Sonology and it is therefore expected that most of the readers of this thesis are in the position to entirely appreciate the content of the thesis. However, even with these prerequisites, the technical details associated with atomic decompositions are far beyond the scope of this thesis. For this reason, a survey paper called *Atomic Decompositions of Square-Integrable Functions* is, for the sake to be as self-contained as possible, attached as appendix A. In this survey paper most of the statements are proved and it can therefore be seen as the foundations of the theory described in the thesis. The paper was originally written as documentation for the Large Time-Frequency Analysis Toolbox (Van Velthoven, 2015).

The main matter of this thesis comprises two parts. The first part consists of the chapters 1-5 and describes the theory associated with methods for atomic decompositions. In chapter 1, the theory associated with Fourier analysis is described. The motivation of the theory in this chapter isn't very extensive since it is extensively described in Tempelaar's book, which is the prerequisites of this thesis. Chapter 1 serves therefore mainly as a chapter in which the notations and terminology used in the rest of the thesis are introduced. In chapter 2 the theory associated with uncertainty principles is described. Uncertainty principle are the main obstructions in the construction of time-frequency representations and they are therefore inseparable related to most methods for atomic decompositions. In chapter 3 the theory associated with Gabor analysis is described. Gabor analysis can be seen as a generalization of Fourier analysis and it serves as one of the key methods for atomic decompositions. In chapter 4 the theory associated with wavelet analysis, an alternative for Gabor analysis, is described. In chapter 5 the general theory of frames is described. In this chapter, the aforementioned methods from Gabor and wavelet analysis are introduced as special cases of frames. Furthermore, the theory associated with non-stationary Gabor analysis, a generalization of both Gabor and wavelet theory, is introduced. The second part of the thesis deals with the application of the methods, introduced in the first part of the thesis, for the analysis, synthesis, and analysis-resynthesis of acoustical signals. Instead of investigating these applications for Gabor, wavelet and non-stationary Gabor theory separately, they are investigated as special cases of frame theory. This second part describes therefore the main principles behind the analysis, synthesis and analysis-resynthesis of acoustical signals through atomic decompositions. In the examples of this second part, the applications of atomic decompositions for the analysis, synthesis and analysis-resynthesis of acoustical signals are exemplified. The analysis, synthesis and analysis-resynthesis associated with atomic decompositions are treated in chapter 6, 7 and 8, respectively.

Prologue

The goal of this prologue is to give a short overview of the origins from which the methods for the analysis and synthesis of acoustical signals that are treated in the main matter of this thesis originated. The reason why these origins are important for the rest of this thesis is because they clearly show the motivation behind the use of such methods for the analysis and synthesis of acoustical signals. In the epilogue of this thesis, the methods for analysis and synthesis that are treated in the main matter of the thesis will be shortly related to these origins. This will be the only relation between the prologue and the main matter of this thesis.

In 1946, the Hungarian-British physicist Dennis Gabor (1900-1979) suggested in his paper *Theory of Communication* that any signal with time as independent variable can also be represented by a signal with time and frequency as independent variables (Gabor, 1946). The method that Gabor proposed in order to achieve such a two-dimensional representation of a signal is to expand this signal into a collection of elementary functions that have time and frequency as independent variables. Thus, according to Gabor, any signal could be represented by an appropriate combination of elementary functions that have time and frequency as independent variables.

The elementary functions that Gabor suggested for the expansion of a signal can be obtained as time-shifted and frequency-shifted versions of a single window function. Such an elementary function has thus both time and frequency as independent variables. The value of such elementary functions can be represented on a plane with time and frequency as independent axes. Such a plane is in general called a *time-frequency plane*. Each version of the elementary functions that Gabor proposed corresponds to a certain area on the time-frequency plane. Gabor called such an area a “logon”. The value of an elementary function that occupies such an area on the time-frequency plane represents what he called a “quantum of information”, which is a certain unit of information. It is nowadays known that the method for representing signals as a sum of time-shifted and frequency-shifted window functions works, under certain conditions, for all kind of window functions. However, the window function that Gabor explicitly proposed is the Gaussian function — see 1. He proposed these types of window functions since they “assure the best utilization of the information area in the sense that they possess the smallest product of effective bandwidth by effective width” (Gabor, 1946).

The expansion that Gabor suggested results thus in a representation of a signal in a domain other than the time domain, namely the time-frequency domain. In this domain every coordinate corresponds to a value of the elementary function. This representation takes the duality of time and frequency, two parameters connected to tone sensations, *ab ovo* into account and is therefore a fruitful method for the analysis of acoustical signals (Gabor, 1947). The quantum of information that such a representation of an acoustical signal posses is what Gabor called a “quantum of sound” or an “acoustical quantum”. The information that such an acoustical quantum utilizes is of significant importance for the theory of hearing (Gabor, 1947).

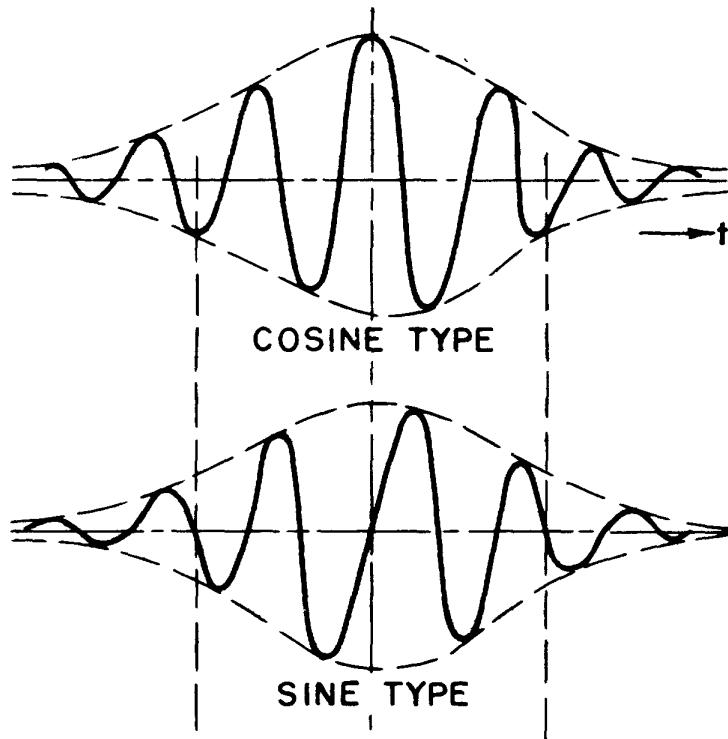


Figure 1: Representations of the elementary function proposed by Gabor (Gabor, 1952). The cosine and sine type correspond to the real respectively imaginary part of the elementary function.

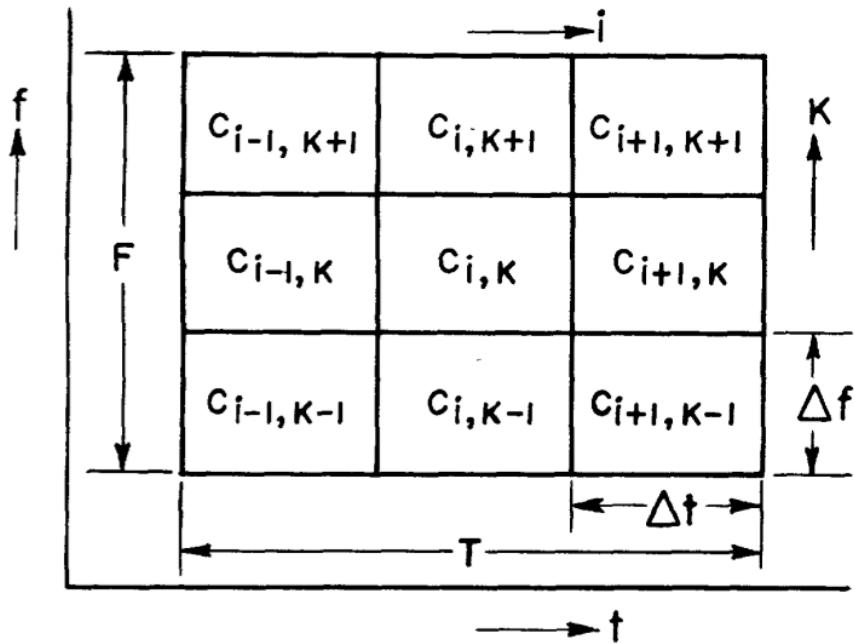


Figure 2: Representation of a signal on the time-frequency plane (Gabor, 1952). The area given by the product $\Delta t \Delta f$ is the logon associated with an elementary function. The value c is the quantum of information of the associated logon.

The Greek-French composer Iannis Xenakis (1922 - 2001) was one of the first who suggested the synthesis of acoustical signals in terms of elementary functions that have time and frequency as independent variables. In his paper *Elements of Stochastic Music*, published in 1960, he wrote “All sounds represent an integration of corpuscles, of elementary acoustic particles, of sound quanta. Each of these elementary particles possesses a double nature: the frequency and the intensity” . In his book *Musique Formelles*, published in 1963, Xenakis extended the double nature of the acoustic particles to a threefold nature with time, frequency and intensity as independent parameters (Xenakis, 1963).

The elementary functions that Xenakis suggested for the synthesis of acoustical signals were complex exponentials multiplied by a window function. Two types of window functions that Xenakis mentioned explicitly are the Gaussian and the rectangular window function. A complex exponential that is windowed by a Gaussian window function is exactly the elementary signal that Dennis Gabor proposed, and for this reason Xenakis referred to these elementary signals as signals of the “Gabor type” or “Gabor sort” — see figure 1.

For the synthesis of acoustical signals, Xenakis developed the concept of what he called a *screen*. A screen is an “intensity-frequency-plane” that can be interpreted as a grid or lattice with amplitude and frequency as independent axes — see figure 3. On such a grid, the amplitude and frequency of an elementary function can be distributed. Since the elementary functions that Xenakis proposed have, due to the windowing, a finite time duration, a screen possess a certain time duration. The time duration that a screen possesses is a constant time duration Δ_t . A screen can therefore be considered as a temporal unit — see figure 4. The fact that a screen can be considered as a temporal unit leads to the possibility to distribute signals with a certain frequency, amplitude and duration, shorter than or equal Δ_t , on a screen. Xenakis called a sequence of screens a *book* — see figure 5. Thus, a book is a sequence of temporal units in which signals with a certain amplitude, frequency and intensity can be distributed.

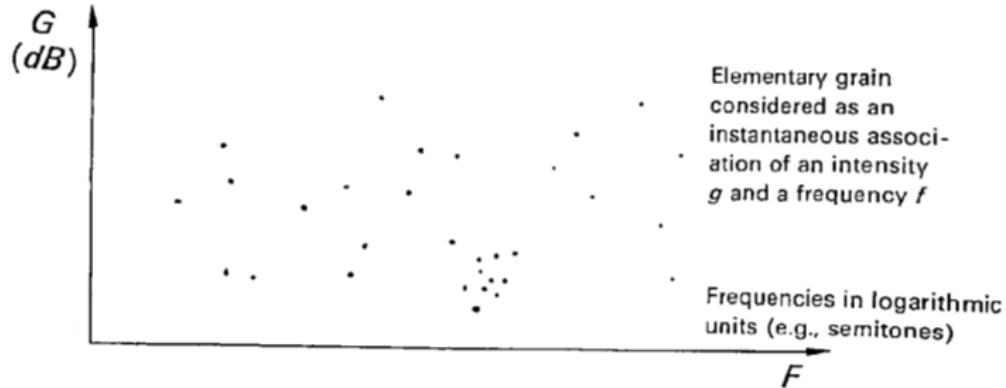


Figure 3: Screen as an intensity-frequency plane (Xenakis, 1992)

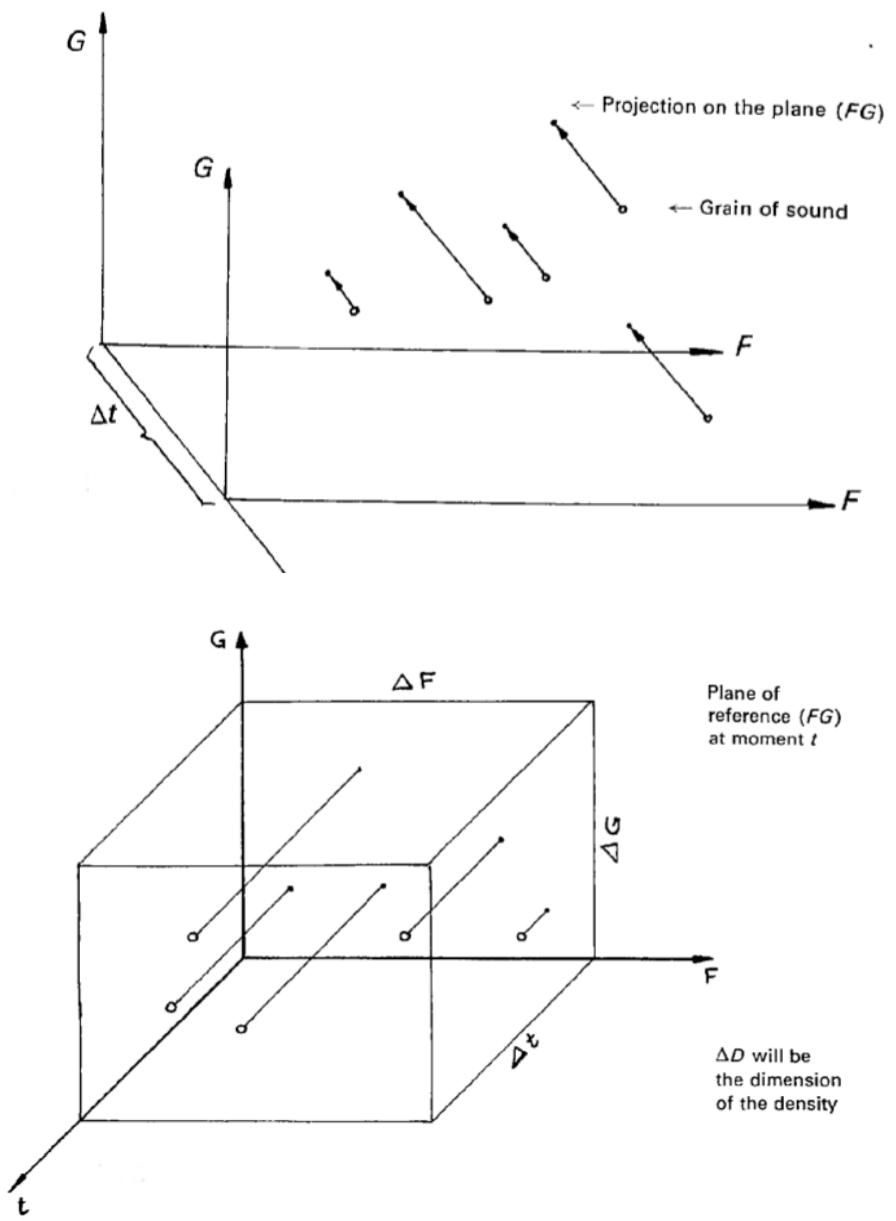


Figure 4: Screens as temporal units (Xenakis, 1992)

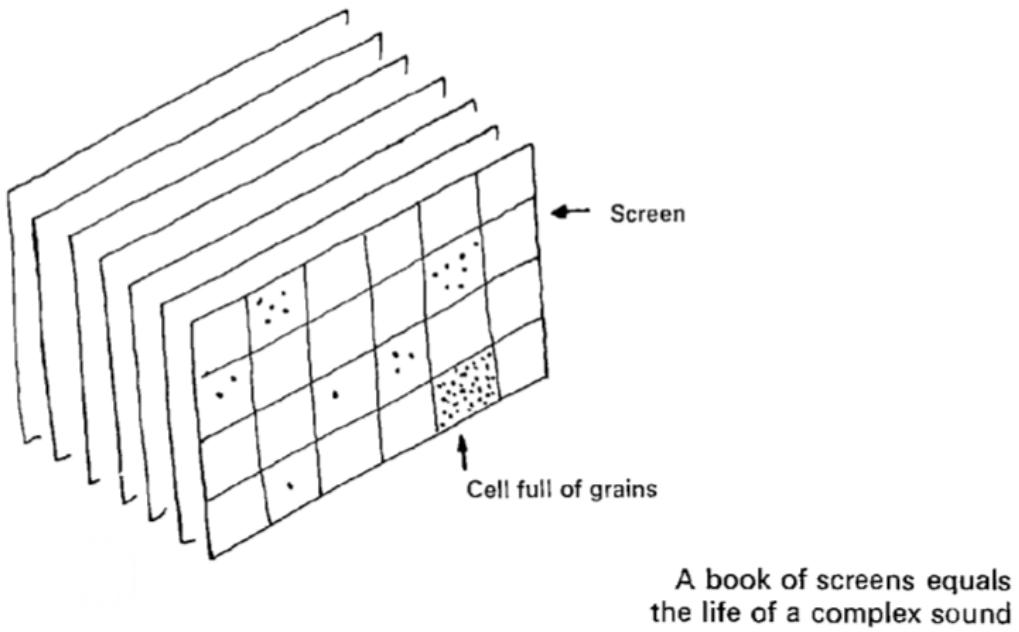


Figure 5: A book as a sequence of screens (Xenakis, 1992)

Now something on the relation between the methods for the analysis and synthesis of acoustical signals proposed by Gabor and Xenakis, respectively. Xenakis added to his temporary hypothesis regarding the threefold nature of sound the following note:

“This description of the micro-structure of acoustic signals is used as a the starting-point of the musical realization and must be understood rather as to be an intuitive representation than of scientific consistence. But it can be considered as to be a first approach towards the ideas introduced in the theory of information by Gabor (Xenakis, 1960).”

From this explicit reference to the work of Dennis Gabor it seems to be justifiable to assume that the theory developed by Gabor was the starting-point for Xenakis’ theory. However, Xenakis seems to deny any influence of Gabor’s work when he writes in 1992 about his research “This research, which I started in 1958 and wrongly attributed to Gabor, can now be pursued with much more powerful and modern means” (Xenakis, 1992). There are several possible reasons why Xenakis denies the influence of Gabor on his research. For example, it could be that Xenakis makes a sharp distinction between the analytical perspective from Gabor and his own more synthetic perspective, but it could also be that Xenakis referred to Gabor in order to get a firm foundation for his temporary hypothesis regarding the threefold nature of sound and to put this hypothesis in a well-established context. Whatever the reason is, it seems to be rather plausible that Xenakis became acquainted with Gabor’s theory through Werner Meyer-Eppler’s book *Grundlagen und Anwendungen der Informations Theorie* since Xenakis quotes the work of Gabor in (1960) through Meyer-Eppler’s book.

Part I

Methods

Chapter 1

Fourier theory

Fourier theory aims to represent functions in terms of sinusoids. These representations lead to a representation of a function in a so-called *frequency domain*. The operator that maps a function from the time domain to the frequency domain is in general called the *Fourier transform*. The Fourier transform is a method which can be used to analyse the spectral characteristics of a function. The inverse of the Fourier transform, the *inverse Fourier transform*, maps a function from the frequency domain to the time domain, and is a method which can be used to synthesize functions with certain spectral characteristics.

1.1 Time domain

A function whose independent variable denotes time is a representation of this function in the so-called *time domain*. If the variable t denotes time, then the value of a function f at time t is denoted by $f(t)$. In the case that the value $f(t)$ of f is complex-valued, that is, $f(t) \in \mathbb{C}$, then f is called a *complex-valued function*. Similarly, if the value $f(t)$ of f is real-valued, that is, $f(t) \in \mathbb{R}$ for all t , then f is called a *real-valued function*. There are several methods that can be used for the analysis or synthesis of functions that can only be applied on complex-valued functions. Aside these methods, there are several attributes of a function that are more accessible when the function is complex-valued than when it is real-valued. For this reason it is often advantageous to extend a real-valued function to a complex-valued function that corresponds to the original real-valued function.

The complex-valued function that corresponds to a real-valued function is in general called the *analytic representation* of this real-valued function. The analytic representation of a real-valued function is a complex-valued function that has the real-valued function from which it is obtained as its real part. The analytic representation of a real-valued function is thus constructed by adding to the original real-valued function a function that consists only of an imaginary part. The sum of these two functions yields then a function that is complex-valued. If f is a real-valued function, then its analytic representation is the complex-valued function z that is defined by

$$z(t) := f(t) + i(\mathcal{H}f)(t),$$

where $\mathcal{H}f$ denotes the Hilbert transform of f , which is defined by

$$(\mathcal{H}f)(t) := \frac{1}{\pi} \int_{\mathbb{R}} \frac{f(\tau)}{t - \tau} d\tau, \quad (1.1)$$

yielding

$$z(t) = f(t) + \frac{i}{\pi} \int_{\mathbb{R}} \frac{f(\tau)}{t - \tau} d\tau. \quad (1.2)$$

The Hilbert transform \mathcal{H} , as defined in (1.1), is an operator that maps the function f into the function $\mathcal{H}f$. Thus, it is this function that forms the imaginary part of the analytic representation of a real-valued function.

Two properties of a complex-valued function that cannot easily be derived from a real-valued function are the instantaneous amplitude and the instantaneous phase. If f is a complex-valued function and $t \in \mathbb{R}$ denotes time, then the value $|f(t)|$ denotes the amplitude of f at time t , and is called the *instantaneous amplitude* of f . The instantaneous amplitude of a function is thus a non-negative number that denotes the magnitude of this function at a certain time instant. The value $\arg f(t)$ of a complex-valued function f denotes the phase of f at the time instant t and is called the *instantaneous phase* of f . The instantaneous phase of a function is in general a value in the interval $] -\pi, \pi]$.

Two important properties of a function that are related to the instantaneous amplitude and the instantaneous phase of this function are the instantaneous energy and the instantaneous frequency, respectively. If $|f(t)|$ denotes the instantaneous amplitude of a function f at the time instant t , then the *instantaneous energy* of f at t is given by $|f(t)|^2$. The *total energy* of f is next given by the continuous superposition of the instantaneous energy over all $t \in \mathbb{R}$,

$$\|f\|_2^2 := \int_{\mathbb{R}} |f(t)|^2 dt.$$

A function f for which $\|f\|_2^2$ is finite, that is, $\|f\|_2^2 < \infty$, is called a *square-integrable* function. A square-integrable function can thus be interpreted as a finite-energy function. The set of all these square-integrable functions is denoted by $L^2(\mathbb{R})$. The square-integrability of a function is an assumption for almost all important operations on functions that are treated in this thesis. If a function is not square-integrable, then most operators that are treated in this thesis cannot properly be applied on this function. However, the square-integrability of a function is a condition that is in general satisfied. Trivially all functions with a “finite duration” satisfy this condition. The main operator that can be applied properly on two functions if, and only if, these functions are square-integrable is the so-called *inner product*. The inner product of two functions $f, g \in L^2(\mathbb{R})$ is defined as

$$\langle f, g \rangle := \int_{\mathbb{R}} f(t)g^*(t)dt, \quad (1.3)$$

where g^* denotes the complex conjugation of g . Intuitively, the inner product between two functions denotes the “overlap” or “similarity” between these two functions. The higher the value of the inner product, the more these two functions overlap. Conversely, the lower the value of the inner product of two functions, the less these two functions overlap. From the next calculation it follows that the inner product between a function f and itself is equal to the total energy of f ,

$$\langle f, f \rangle = \int_{\mathbb{R}} f(t)f^*(t)dt = \int_{\mathbb{R}} |f(t)|^2 dt = \|f\|_2^2. \quad (1.4)$$

From this calculation, and the definition of a square-integrable function, it follows that (1.4) can only be finite if the function f is square-integrable. In fact, it also holds that (1.3) is finite if, and only if, both f and g are square-integrable. Since the inner product is one of the main operators in this thesis, the condition that a function is square-integrable is fundamental.

The instantaneous frequency of a function is, as stated above, related to the instantaneous phase of this function. If the instantaneous phase of a function f at time t is denoted by $\phi_f(t) = \arg f(t)$, then the instantaneous frequency Φ_f of f at t is given by¹

$$\Phi_f(t) := \frac{1}{2\pi} \frac{d}{dt} \phi_f(t). \quad (1.5)$$

The instantaneous frequency of a function f corresponds therefore for each instant $t \in \mathbb{R}$ to one, and only one, value $\Phi_f(t)$. The instantaneous frequency as defined in (1.5) provides thus a value

¹The definition of the instantaneous frequency as in (1.5) is the one that is in general used. However, the instantaneous frequency can also be expressed in terms of a so-called *frequency centroid* (Guillemain & Kronland-Martinet, 1996). The expression of the instantaneous frequency in terms of the frequency centroid makes use of the *Fourier transform* of a function. Since the definition of the Fourier transform is given in section §1.2, the expression of the instantaneous frequency in terms of the frequency centroid is postponed to §1.2.

of frequency as a function of time. This value of the instantaneous frequency gives an indication of the frequency of the analysed function at a certain time instant. Since the instantaneous frequency corresponds to only one value, it only denotes the spectral content of a function that consists of just one frequency component per time instant. In the case that a function consists of more than one frequency component at a certain time instant, then the instantaneous frequency denotes a value between the frequencies of the frequency components with the highest and lowest frequency. Note that in this case the instantaneous frequency can correspond to a value of frequency that isn't present in the analysed function at all. The concept of an instantaneous frequency is therefore not as useful for functions that consist of multiple frequency components at a certain instant as it is for functions that consist of only one frequency component at a certain time instant. In order to get the entire spectral content of a function, the function should be represented in the frequency domain.

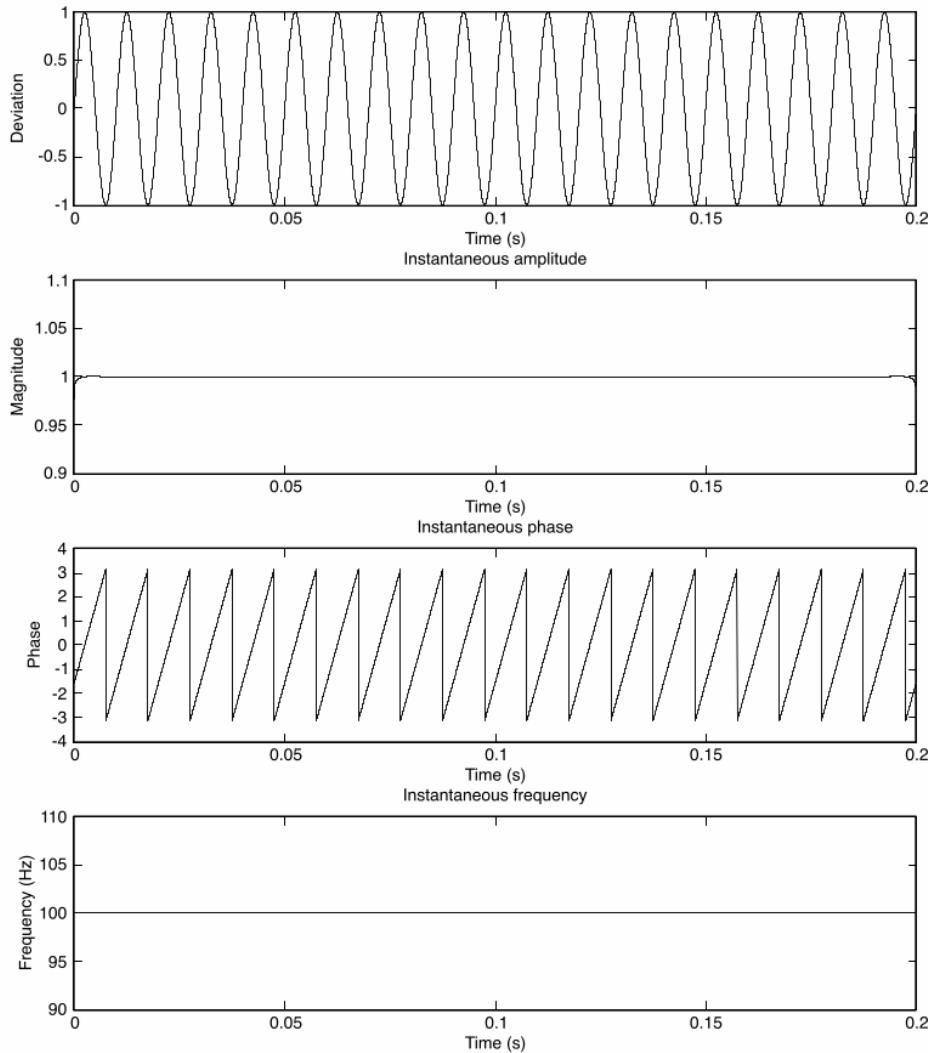
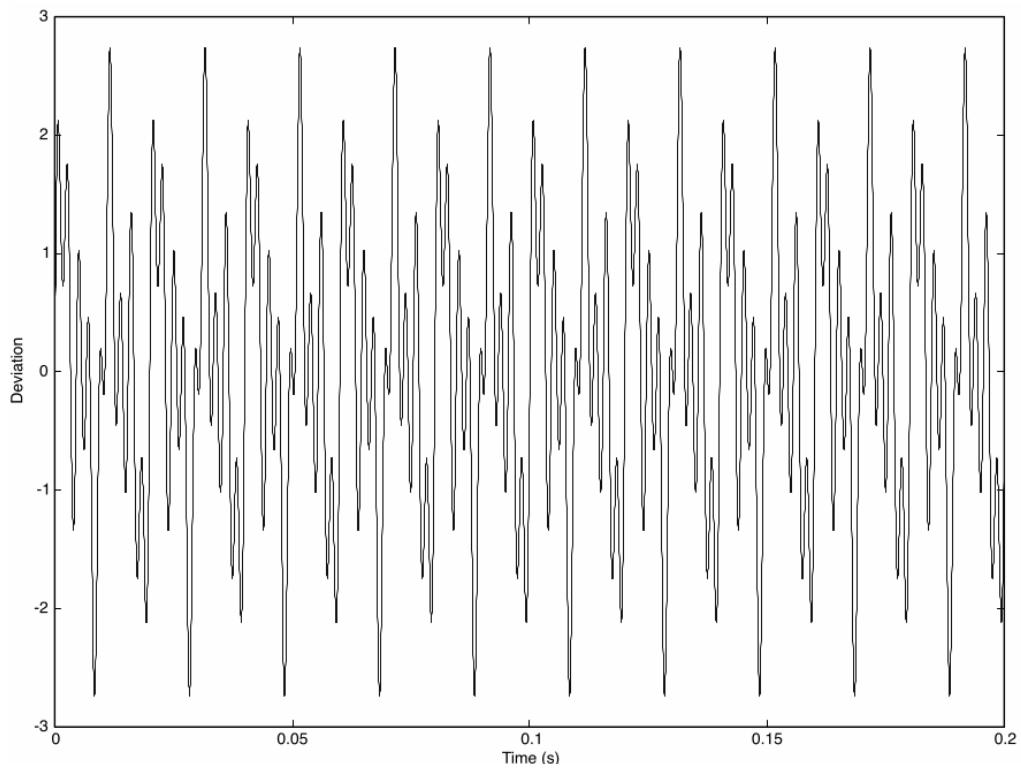
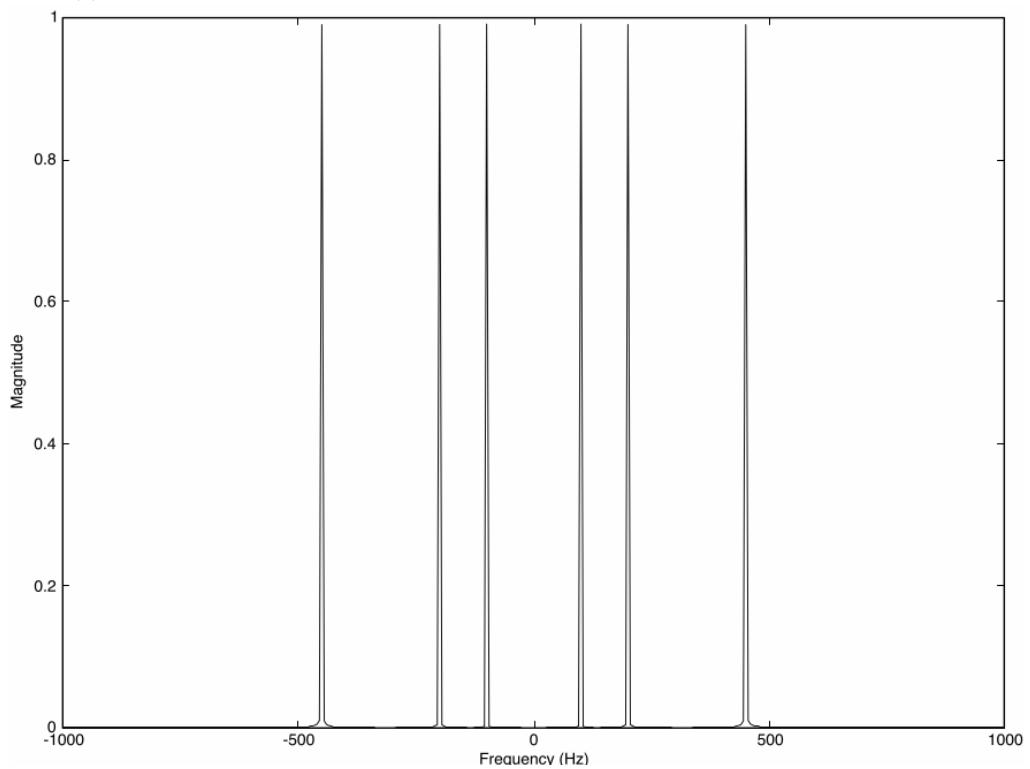


Figure 1.1: Representations of a sinusoid with constant frequency of 100 Hz in the time domain



(a) Sum of sinusoids with constant frequencies of 100, 200 and 450 Hz in the time domain



(b) Sum of sinusoids with constant frequencies of 100, 200 and 450 Hz in the frequency domain

Figure 1.2: Fourier transform pair

1.2 Frequency domain

A function whose independent variable denotes frequency is a representation of this function in the so-called *frequency domain*. A function can be mapped from the time domain to the frequency domain through the so-called *Fourier transform*. The Fourier transform \mathcal{F} of a function $f \in L^2(\mathbb{R})$ will be denoted as

$$\begin{aligned}\hat{f}(\xi) = (\mathcal{F}f)(\xi) &= \int_{\mathbb{R}} f(t)e^{-i2\pi\xi t} dt \\ &= \langle f, e^{i2\pi\xi t} \rangle.\end{aligned}$$

The Fourier transform \mathcal{F} is thus an operator that maps a function f into a function \hat{f} , called the *Fourier transform of f* . The Fourier transform \hat{f} is a function whose independent variable denotes frequency and is therefore a representation of the function f in the frequency domain. Since the Fourier transform of a function f can be written as the inner product between f and $e^{i2\pi\xi t}$, the Fourier transform describes the “overlap” or the “similarity” between f and the complex exponentials $e^{i2\pi\xi t}$, where $\xi \in \mathbb{R}$. An intuitive interpretation of the Fourier transform \hat{f} can be obtained by expressing the value $\hat{f}(\xi)$ in the form

$$\hat{f}(\xi) = a(\xi)e^{i\varphi(\xi)}.$$

Here, the value $a(\xi) = |\hat{f}(\xi)|$ of the Fourier transform \hat{f} denotes the magnitude of \hat{f} at frequency ξ , and is called the *spectral amplitude*. The spectral amplitude is a non-negative number and can be interpreted as the amplitude of the *spectral component* $e^{i2\pi\xi t}$. The value $\varphi(\xi) = \arg \hat{f}(\xi)$ of the Fourier transform \hat{f} denotes the phase of \hat{f} at frequency ξ , and is called the *spectral phase*. The spectral phase denotes the phase of the spectral component and is in general a value in the interval $]-\pi, \pi]$.

There are two important properties of a Fourier transform of a function that are related to the spectral amplitude and the spectral phase, namely the energy density spectrum respectively the group delay of this function. If $|\hat{f}(\xi)|$ denotes the spectral amplitude of the Fourier transform \hat{f} at frequency ξ , then $|\hat{f}(\xi)|^2$ denotes the energy density spectrum of \hat{f} at ξ . The energy density spectrum denotes the energy of the Fourier transform \hat{f} at the frequency ξ . The energy density spectrum can therefore be interpreted as the energy of the spectral component $e^{i2\pi\xi t}$. The total energy of \hat{f} is next given by the continuous superposition over all $\xi \in \mathbb{R}$,

$$\|\hat{f}\|_2^2 = \int_{\mathbb{R}} |\hat{f}(\xi)|^2 d\xi.$$

Since both the function f and its Fourier transform \hat{f} are representations of the same function, it is natural to assume that the energy of f and \hat{f} are equal, that is,

$$\|f\|_2^2 = \|\hat{f}\|_2^2.$$

This equality holds for all functions $f \in L^2(\mathbb{R})$, and is in general called Plancherel’s equality for the Fourier transform. From this equality it can be deduced that if a function has a finite-energy, then its Fourier transform also has a finite-energy. The inner product between the two Fourier transforms \hat{f}, \hat{g} of the functions $f, g \in L^2(\mathbb{R})$ is therefore given by

$$\langle \hat{f}, \hat{g} \rangle = \int_{\mathbb{R}} \hat{f}(\xi)\hat{g}^*(\xi) d\xi.$$

It further holds that

$$\langle f, g \rangle = \langle \hat{f}, \hat{g} \rangle, \tag{1.6}$$

which is known as Parseval’s equality for the Fourier transform. Parseval’s equality shows that the inner product of two functions in the time domain is equal to the inner product of the Fourier transforms of these functions.

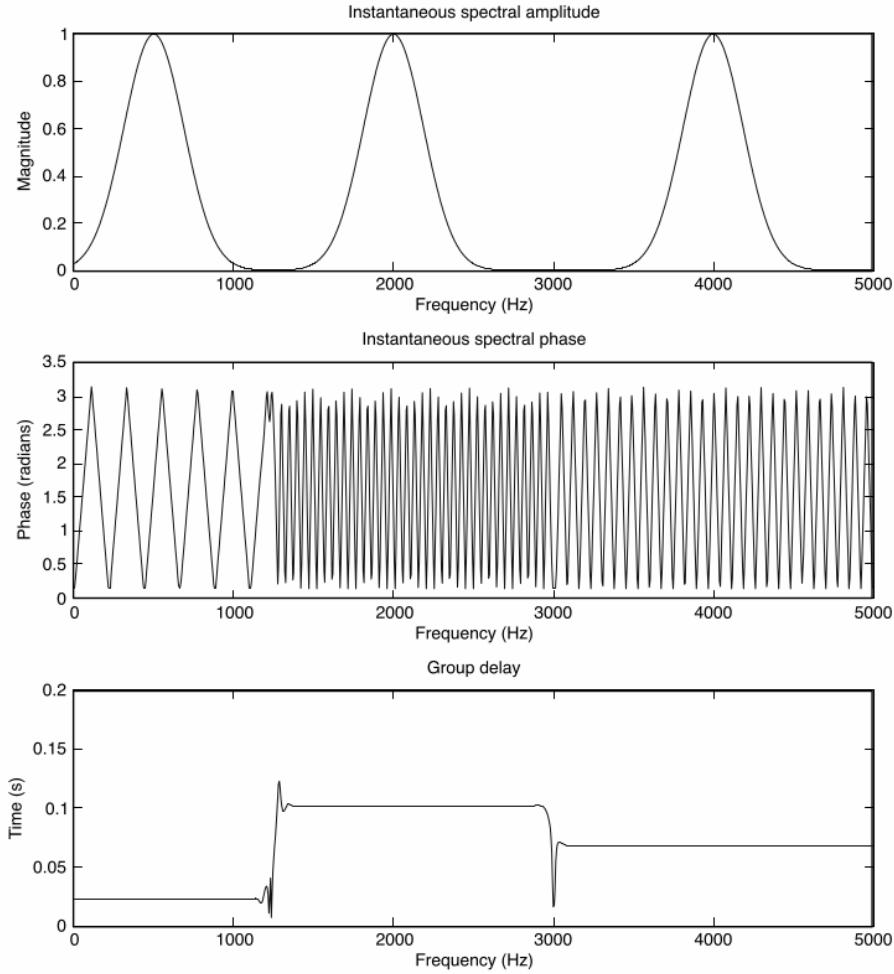


Figure 1.3: Representations of windowed sinusoids in the frequency domain

The group delay is, as stated previously, related to the spectral phase. If the spectral phase of a function f is denoted by $\theta_f(\xi) = \arg \hat{f}(\xi)$, then the group delay Θ_f of f is given by

$$\Theta_f(\xi) := -\frac{1}{2\pi} \frac{d}{d\xi} \theta_f(\xi). \quad (1.7)$$

The group delay of a function f corresponds for each frequency $\xi \in \mathbb{R}$ to one, and only one, value $\Theta_f(\xi)$. The group delay, as defined in (1.7), provides a value of time as a function of frequency. This value gives an indication of the time of appearance of a certain frequency. Since the group delay corresponds to only one value for each frequency, it only denotes the time of appearance of a frequency component that exists at one time instant. In the case that a function consists of a certain frequency component at several time instants, then the group delay denotes a value which doesn't have to be a time instant where the frequency component appears. The concept of a group delay is therefore not as useful for functions whose frequency components appear at several time instants than it is for functions whose frequency components appear only at one moment in time.

Aside the definition of the group delay as in (1.7), there is also a definition of the group delay based on a so-called *time centroid* (Guillemain & Kronland-Martinet, 1996). In this case, the

group delay is given by

$$\Theta_f(\xi) = \Re \left\{ \frac{\int_{\mathbb{R}} t \cdot f(t) e^{i2\pi\xi t} dt}{\int_{\mathbb{R}} f(t) e^{i2\pi\xi t} dt} \right\},$$

where the symbol \Re denotes the real part. Just as the group delay can be expressed in terms of a time centroid, the instantaneous frequency can be expressed in terms of a frequency centroid. In this case, the instantaneous frequency $\Phi_f(t)$ is given by

$$\Phi_f(t) = \Re \left\{ \frac{\int_{\mathbb{R}} \xi \cdot \hat{f}(\xi) e^{-i2\pi\xi t} d\xi}{\int_{\mathbb{R}} \hat{f}(\xi) e^{-i2\pi\xi t} d\xi} \right\}.$$

The temporal properties of a function can in general be obtained by representing this function in the time domain. The Fourier transform \hat{f} of a function f can be mapped into the time domain through the inverse Fourier transform. The inverse Fourier transform \mathcal{F}^{-1} of a Fourier transform $\hat{f} \in L^2(\mathbb{R})$ is given by

$$f(t) = (\mathcal{F}^{-1}\hat{f})(t) = \int_{\mathbb{R}} \hat{f}(\xi) e^{i2\pi\xi t} d\xi.$$

The inverse Fourier transform of \hat{f} is again the function f represented in the time domain. From the inverse Fourier transform it can be deduced that each value $f(t)$ of the function f can be represented by a continuous superposition of complex exponentials $e^{i2\pi\xi t}$.

1.3 Operators associated with Fourier transforms

There are several important operators that have a special relation with the Fourier transforms. These operators are important in the sense that they are useful in applications, but also because they are related to other important operators that are treated in this thesis. In fact, the operators that are described in this section can be seen as “building blocks” for other, more advanced, operators that are treated in the rest of this thesis. For this reason, it is advantageous to have a good understanding of the operators that are treated in this section.

The translation operator \mathcal{T}_τ , that *translates* a function $f \in L^2(\mathbb{R})$ by $\tau \in \mathbb{R}$, is given by

$$(\mathcal{T}_\tau f)(t) = f(t - \tau). \quad (1.8)$$

In the case that the variable t of the function f in (1.8) denotes time, then a translation by τ can be interpreted as a *time shift* of f by τ — see the left subfigures in figure 5.1 and 5.2. Similarly, a translation of a function whose independent variable denotes frequency can be interpreted as a *frequency shift* of this function.

The modulation operator \mathcal{E}_ξ , that *modulates* a function $f \in L^2(\mathbb{R})$ by $\xi \in \mathbb{R}$, is given by

$$(\mathcal{E}_\xi f)(t) = f(t) e^{i2\pi\xi t}. \quad (1.9)$$

A modulation of a function in the time domain equals a multiplication by a spectral component — see the top subfigures in figure 5.1. Such a multiplication results in a shift of the Fourier transform \hat{f} of the function f by ξ . In the same way, it can be shown that a modulation in the frequency domain corresponds with a time shift of the function in the time domain. This results in the next important relation between translation, modulation and the Fourier transform of a function

$$(\mathcal{F}\mathcal{T}_\tau f)(\xi) = (\mathcal{E}_\tau \hat{f})(\xi).$$

From this relation it can be deduced that the translation of a function in the time domain corresponds to the modulation of the Fourier transform of this function. Similarly, the modulation of a function in the time domain corresponds to the translation of the Fourier transform of this function.

The dilation operator \mathcal{D}_v , that *dilates* a function $f \in L^2(\mathbb{R})$ by $v \neq 0$, is given by

$$(\mathcal{D}_\rho)(t) = \frac{1}{\sqrt{|\rho|}} f\left(\frac{t}{\rho}\right). \quad (1.10)$$

The dilation of a function in the time domain could be interpreted as “scaling” this function — see the top subfigures in figure 5.2. Dilating a function in the time domain corresponds to an inverse dilation of the Fourier transform of this function, that is,

$$(\mathcal{F}\mathcal{D}_\rho f)(\xi) = (\mathcal{D}_{\frac{1}{\rho}} \hat{f})(\xi). \quad (1.11)$$

The involution of a function $f \in L^2(\mathbb{R})$ is the function f^\diamond defined by

$$f^\diamond = f^*(-t), \quad (1.12)$$

where f^* denotes the complex conjugated version of f . If the variable t in (1.12) denotes time, then the involution of a function can be interpreted as the “time-reversed” version of its conjugation.

The convolution operator associated with a function $f \in L^2(\mathbb{R})$ in the time domain is denoted as

$$(f * g)(t) = \int_{\mathbb{R}} f(\tau)g(t - \tau)d\tau.$$

The function g that is associated with the convolution operator is in general called the *convolution kernel*. The convolution kernel can be several types of functions, but it is often a so-called *filter*. Note that a filter is also called an *impulse response* and *frequency response* when it is presented in the time domain and frequency domain, respectively. If the convolution kernel is a filter, then convolution is often called *filtering*.

The convolution of a function f with the convolution kernel g can be written in terms of an inner product as

$$(f * g)(t) = \langle f(\tau), g^*(t - \tau) \rangle. \quad (1.13)$$

From writing the convolution in this form it can be deduced that the convolution determines the “similarity” between $f(\tau)$ and a by t translated version of the involution of g , g^\diamond . Using Parseval’s equality (1.6) and the fact that translation in the time domain corresponds to a modulation in the frequency domain, it can be shown that the convolution of the functions f and g in the time domain corresponds to the product of their Fourier transforms \hat{f} and \hat{g} in the frequency domain, that is,

$$\mathcal{F}(f * g) = \hat{f} \cdot \hat{g}.$$

Similarly, it can be shown that the product of two functions in the time domain corresponds to the convolution of their Fourier transforms in the frequency domain,

$$f \cdot g = \mathcal{F}^{-1}(\hat{f} * \hat{g}).$$

Since the convolution operator is related to such a fundamental operation as multiplication, it can be related to almost all operations on functions. Since the operation of filtering is in general well-understood, the convolution leads to valuable interpretations of other, less familiar operators, when these are interpreted as a convolution.

Chapter 2

Fourier uncertainty principles

An uncertainty principle is in general an inequality that involves a function and its spectrum. The uncertainty principle associated with Fourier analysis is therefore an inequality that involves a function and its Fourier transform. In fact, the uncertainty principles associated with Fourier analysis are inequalities that involve the intervals on which the values of a function and its Fourier transform are non-zero. These inequalities describe thus a relation between the intervals on which the values of a function and its Fourier transform are non-zero. It is such a relation that is in general called a *Fourier uncertainty principle*. Every representation that involves a function and its Fourier transform comes with its own version of such an uncertainty principle (Gröchenig, 2003). In fact, the uncertainty principles form an obstruction for all so-called *time-frequency representations* — representations in which both the temporal and spectral aspects of a function can be represented. Since time-frequency representations form an essential part of the rest of this thesis, the associated uncertainty principles cannot be avoided. As the title of this chapter already indicates, only the Fourier uncertainty principles will be treated in this chapter. Some peculiar versions of the uncertainty principle, the ones associated with Gabor and wavelet analysis, are treated in §3.2 and §4.2, respectively.

2.1 The Weyl-Heisenberg inequality

The interval on which a function in the time domain is non-zero can be interpreted as the “duration” of this function. The measure for the duration of a function is related to the moment in time where the function is concentrated. This moment in time can be called the *time centre* of the function since it is centred around this point in the time domain. Since a function can be considered as a density in time (Cohen, 1995), the time centre of a function can be defined as the mean value or average of this function. The time centre \bar{t} of a function $f \in L^2(\mathbb{R})$ is therefore explicitly defined as

$$\bar{t} := \frac{1}{\|f\|_2^2} \int_{\mathbb{R}} t \cdot |f(t)|^2 dt. \quad (2.1)$$

The time centre gives an indication around which time instant the energy of the function is centred. From this time centre the spread of the energy in the time domain can be measured. This time spread is equal to the standard deviation of this function in the time domain with respect to the time centre. For a function $f \in L^2(\mathbb{R})$, the time spread is defined as

$$\Delta_t := \left(\frac{1}{\|f\|_2^2} \int_{\mathbb{R}} (t - \bar{t})^2 |f(t)|^2 dt \right)^{1/2}. \quad (2.2)$$

The time spread gives an indication on how the energy of the function is concentrated in time. For this reason, the time spread of a function can be considered as a measure for its “duration”. If the value of Δ_t is small, then most of the energy of the function is concentrated around its time centre \bar{t} . Similarly, if the value of Δ_t is large, then most of the energy of the function is spread in

time. Note that the time spread only gives an indication of the size of the time interval on which the function is non-zero. In order to determine the actual time interval on which a function is non-zero, the time centre and time spread of this function should be combined, which leads to the time interval

$$[\bar{t} - \Delta_t, \bar{t} + \Delta_t]. \quad (2.3)$$

The “bandwidth” of a function in the time domain is the interval on which its Fourier transform is non-zero. The measure for the bandwidth of a function is related to the frequency around which the function is concentrated. This frequency can be called the *frequency centre* of the function since the energy of its Fourier transform is centred around this frequency. Similarly as the time centre is related to the mean of the function in the time domain, the frequency centre is related to the mean of this function in the frequency domain. For a function $f \in L^2(\mathbb{R})$, the frequency centre is defined as

$$\bar{\xi} := \frac{1}{\|\hat{f}\|_2^2} \int_{\mathbb{R}} \xi \cdot |\hat{f}(\xi)|^2 d\xi, \quad (2.4)$$

where \hat{f} denotes the Fourier transform of the function f . The frequency centre gives an indication of the frequency around which the function f is centred. From this frequency centre the spread of the energy of this function over frequency can be measured. The frequency spread is equal to the standard deviation of the function in the frequency domain with respect to the frequency centre. For a function $f \in L^2(\mathbb{R})$, the frequency spread is defined as

$$\Delta_\xi := \left(\frac{1}{\|\hat{f}\|_2^2} \int_{\mathbb{R}} (\xi - \bar{\xi})^2 |\hat{f}(\xi)|^2 dt \right)^{1/2}. \quad (2.5)$$

The frequency spread gives an indication on how the energy of the function is concentrated in frequency. For this reason, the frequency spread of a function can be considered as a measure for its bandwidth. If the value of Δ_ξ is small, then most of the energy of the function is concentrated around its frequency centre $\bar{\xi}$. Similarly, if the value of Δ_ξ is large, then most of the energy of the function is spread in frequency. Note that the frequency spread only determines the size of the frequency interval on which the function is non-zero. In order to determine the actual frequency interval on which a function is non-zero, the frequency centre and frequency spread of this function should be combined, which leads to the frequency interval

$$[\bar{\xi} - \Delta_\xi, \bar{\xi} + \Delta_\xi]. \quad (2.6)$$

The time and frequency spreads as defined in (2.2) respectively (2.5) can be related to each other through the so-called *Weyl-Heisenberg inequality*, which is considered as the most common quantitative description of an uncertainty principle (Folland & Sitaram, 1997). For a function $f \in L^2(\mathbb{R})$, the Weyl-Heisenberg inequality is given by

$$\left(\int_{\mathbb{R}} (t - a)^2 |f(t)|^2 dt \right)^{1/2} \left(\int_{\mathbb{R}} (\xi - b)^2 |\hat{f}(\xi)|^2 dt \right)^{1/2} \geq \frac{\|\hat{f}\|_2^2}{4\pi},$$

for any $a, b \in \mathbb{R}$. Now, by setting $a = \Delta_t$ and $b = \Delta_\xi$, a relation between the time spread and frequency spread of a function follows as a consequence from the Weyl-Heisenberg inequality,

$$\Delta_t \Delta_\xi \geq \frac{1}{4\pi}. \quad (2.7)$$

This inequality, relating the time and frequency spread of a function, allows several interpretation. Two of such interpretations will now be enumerated:

1. The inequality given in (2.7) shows that the product of the time spread and the frequency spread is bounded from below by the constant $(4\pi)^{-1}$. Thus, the time spread and frequency

spread of a function cannot be both arbitrary small since their product is bounded from below. In other words, it isn't possible that the values of a function are non-zero on an arbitrary small interval in both the time domain and in the frequency domain. The relation given in (2.7) describes therefore the duality that is connected to Fourier analysis, namely that the concentration of a function in one domain causes an inverse concentration in the conjugate domain, and vice versa. Thus, if a function has a well-concentrated representation in the time domain, then it will have a non-concentrated representation in the frequency domain, and vice versa.

2. The inequality given in (2.7) allows aside interpretations of the properties of a function in the time and frequency domain, also an interpretation of the properties of a function in the so-called *time-frequency domain*, a domain in which time and frequency are represented by two independent axes. As the name already suggests, a time-frequency domain could be seen as a compound of the time domain and the frequency domain. Since the representation of a function in the time domain corresponds for each value of time to one function value, and since the representation of a function in the frequency domain corresponds for each value of frequency to one value of its Fourier transform, the representation of a function in the time-frequency domain corresponds for each value of time and frequency to one value of the function. A representation of a function in the time-frequency domain has thus two independent variables, namely a variable that denotes time and a variable that denotes frequency. A representation of a function in the time-frequency domain is in general called a *time-frequency representation* of this function, and is a function of two independent variables. The plane on which this function is represented is called the *time-frequency plane*.

The time and frequency centre of a function, and the time and frequency spread of a function can all be represented on a time-frequency plane. In order to do so, the time centre \bar{t} and frequency centre $\bar{\xi}$ should be considered as the pair $(\bar{t}, \bar{\xi})$, and the time spread Δ_t and Δ_ξ should be considered as the pair (Δ_t, Δ_ξ) . Just as the time centre \bar{t} and frequency centre $\bar{\xi}$ of a function denote the position in time respectively frequency where the function is concentrated, so denotes the pair $(\bar{t}, \bar{\xi})$ the position in time and frequency where the function is concentrated. Similarly, just as the values of the time spread Δ_t and frequency spread Δ_ξ denote the size of the intervals $[\bar{t} - \Delta_t, \bar{t} + \Delta_t]$ and $[\bar{\xi} - \Delta_\xi, \bar{\xi} + \Delta_\xi]$ on which a function in the time domain respectively the frequency domain is concentrated, so denotes the pair (Δ_t, Δ_ξ) the size of the area

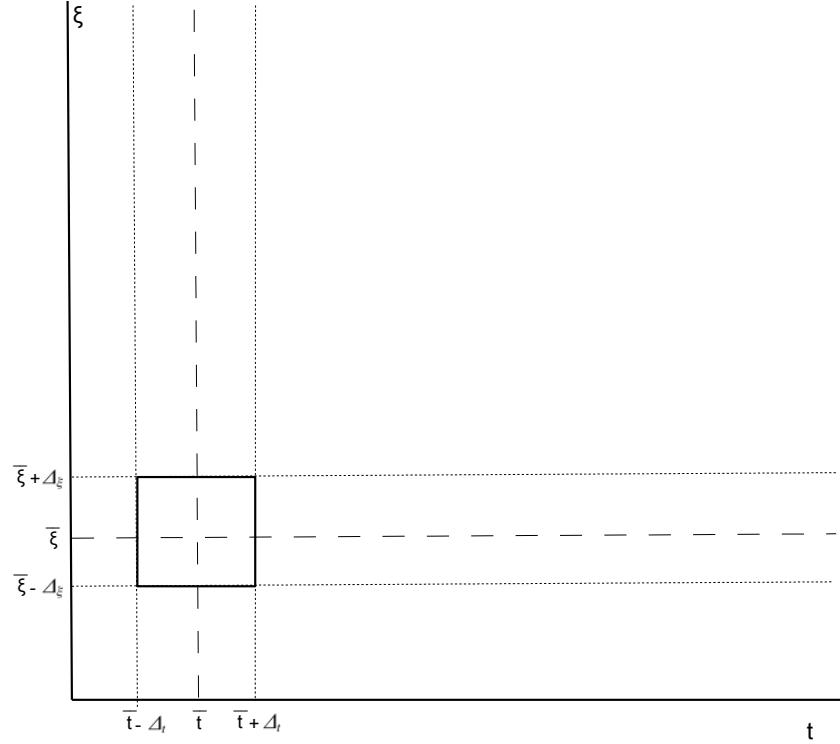
$$[\bar{t} - \Delta_t, \bar{t} + \Delta_t] \times [\bar{\xi} - \Delta_\xi, \bar{\xi} + \Delta_\xi], \quad (2.8)$$

on which a function in the time-frequency domain is concentrated. This area forms a rectangle on the time-frequency plane, and is in general called a *Heisenberg box* — see figure 2.1. The position of the Heisenberg box on the time-frequency plane is thus determined by the coordinates $(\bar{t}, \bar{\xi})$, and the size of the Heisenberg box is determined by the pair (Δ_t, Δ_ξ) .

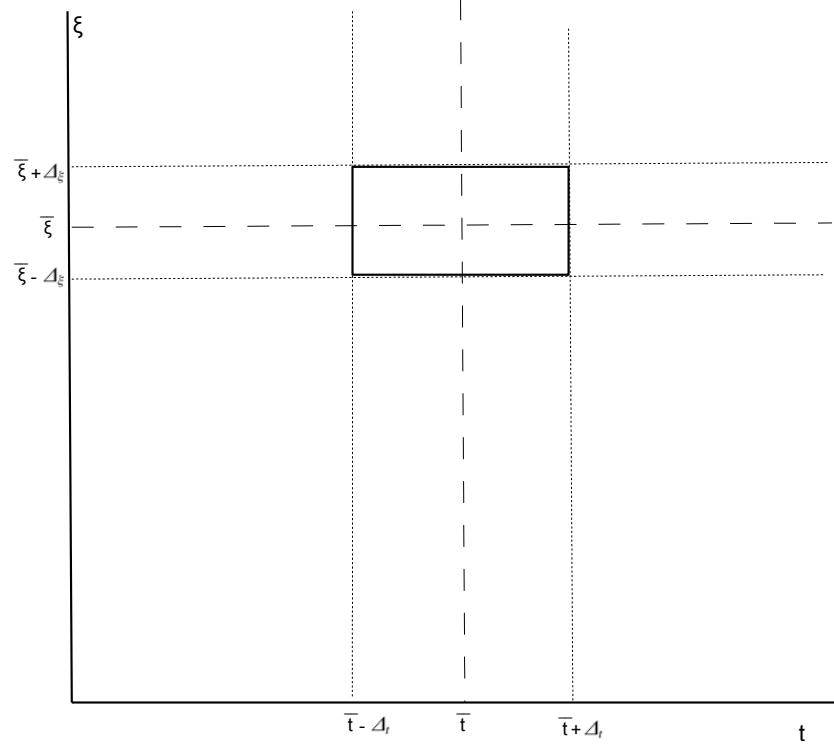
Now the concepts of a time-frequency domain and a Heisenberg box are introduced, their relation with the uncertainty principle given in (2.7) is obvious. Since the product $\Delta_t \Delta_\xi$ determines an area on the time-frequency plane, it follows from (2.7) that this area can't be arbitrary small since the area $\Delta_t \Delta_\xi$ is bounded from below.

2.2 Slepian-Pollack inequality

In the previous section it was stated that the duration of a function in the time domain is the interval on which the value of this function is non-zero. If the value $f(t)$ of a function f is zero for almost all t outside a certain time interval $[-T, T]$, then the function is said to be *time-limited* to the interval $[-T, T]$. In this case, the duration of f is limited to the time interval $[-T, T]$, and the duration of this function is $2T$. Note that it was stated in the previous section that a function that has a time centre \bar{t} and a time spread Δ_t is concentrated on the interval given in (2.3). Such



(a) An example of a Heisenberg box where the values of Δ_t and Δ_ξ are equal.



(b) An example of a Heisenberg box where the value of Δ_t is higher than the value of Δ_ξ .

Figure 2.1: Examples of Heisenberg boxes on the time-frequency plane

a function has thus a duration $2\Delta_t$. The set of all functions $f \in L^2(\mathbb{R})$ that are time-limited to the interval $[-T, T]$ will be denoted by $L_{[-T, T]}^2(\mathbb{R})$. A function which is an element of $L_{[-T, T]}^2(\mathbb{R})$ can thus be interpreted as a function that has a finite energy and a finite duration. The energy of a function on a time interval $[-T, T]$ is given by

$$\int_{-T}^T |f(t)|^2 dt. \quad (2.9)$$

For every $f \in L_{[-T, T]}^2(\mathbb{R})$, this is equal to $\|f\|_2^2$ since the energy of these functions is concentrated on the interval $[-T, T]$.

In the previous section, the bandwidth of a function was defined as the interval on which the Fourier transform of this function is non-zero. Due to this, a function is said to be *band-limited* to an interval $[-\Omega, \Omega]$ if the value of its Fourier transform is zero for almost every ξ outside this interval. In this case, the bandwidth of this function is limited to the interval $[-\Omega, \Omega]$, and the bandwidth of this function is 2Ω . The set of all functions $f \in L^2(\mathbb{R})$ that are band-limited to $[-\Omega, \Omega]$ are elements of the so-called *Paley-Wiener space* $\mathcal{PW}_\Omega(\mathbb{R})$. A function that is an element of this space can thus be interpreted as a function that has a finite energy and a finite bandwidth. The energy of a function on a frequency interval $[-T, T]$ is given by

$$\int_{-\Omega}^{\Omega} |\hat{f}(\xi)|^2 d\xi. \quad (2.10)$$

For every $f \in \mathcal{PW}_\Omega(\mathbb{R})$, this is equal to $\|f\|_2^2$ since the energy of these functions is concentrated on the interval $[-\Omega, \Omega]$.

For a function with a finite energy, a so-called *time concentration* and *frequency concentration* can be defined. The time and frequency concentrations of a function give an indication of the amount of energy that is concentrated on a certain time respectively frequency interval, relative to the total energy of the function. The *time concentration* of a function is therefore defined as the energy of this function on a certain time interval, divided by its total energy. The time concentration α_T of a function $f \in L^2(\mathbb{R})$ is explicitly given by

$$\alpha_T^2 = \frac{\int_{-T}^T |f(t)|^2 dt}{\int_{\mathbb{R}} |f(t)|^2 dt}. \quad (2.11)$$

From this it can be deduced that the time concentration of a function $f \in L_{[-T, T]}^2(\mathbb{R})$ is 1 since for such a function the energy on the interval $[-T, T]$ equals its total energy. The *frequency concentration* of a function is next defined as the energy of the Fourier transform of this function on a certain frequency interval, divided by its total energy. The frequency concentration β_Ω for a function $f \in L^2(\mathbb{R})$ is given by

$$\beta_\Omega^2 = \frac{\int_{-\Omega}^{\Omega} |\hat{f}(\xi)|^2 d\xi}{\int_{\mathbb{R}} |\hat{f}(\xi)|^2 d\xi}. \quad (2.12)$$

For the same reason as the time concentration of a function that is time-limited to the interval $[-T, T]$ is 1, the frequency concentration of a function $f \in \mathcal{PW}_\Omega(\mathbb{R})$ is 1.

The time and frequency concentrations that are defined in (2.11) respectively (2.12) are related to each other through the so-called *Slepian-Pollack inequality*, which can also be considered as a Fourier uncertainty principle (Hogan & Lakey, 2007). The Slepian-Pollack inequality is given by

$$T\Omega \geq q(\alpha, \beta),$$

where $q(\alpha, \beta)$ is a function that depends on α_T and β_Ω and whose value is non-negative. This inequality shows that it is impossible to concentrate the energy of a function to an arbitrary finite interval $[-T, T]$ in the time domain and an arbitrary finite interval $[-\Omega, \Omega]$ in the frequency domain. Another interpretation: the energy of a function can't be represented by an arbitrary small area on the time-frequency plane.

Chapter 3

Gabor theory

Gabor theory aims to represent a function as a sum of so-called *window functions*. Such a representation of a function corresponds to a representation of a function in the time-frequency domain. The mapping that maps a function from the time domain to the time-frequency domain is, in the case of Gabor theory, called the *Gabor transform*. Its inverse, the *inverse Gabor transform*, maps a function from the time-frequency domain to the time domain. The Gabor transform and its inverse are therefore methods that allow one to analyse respectively synthesize functions whose spectral characteristics evolve over time or whose temporal characteristics evolve over frequency. This is in contrast to the Fourier transform and its inverse, which are methods that allow one to analyse respectively synthesize functions with certain spectral characteristics or temporal characteristics, but these characteristics doesn't evolve over time or frequency.

The Gabor transform possesses in general two versions, namely a *continuous* and a *discrete* version. In this chapter, only the continuous version will be treated. The discrete version will be covered in Chapter 5, in the context of Gabor frames. For a more in-depth discussion of both the continuous and discrete Gabor transform, the reader is referred to Gröchenig's book (Gröchenig, 2001).

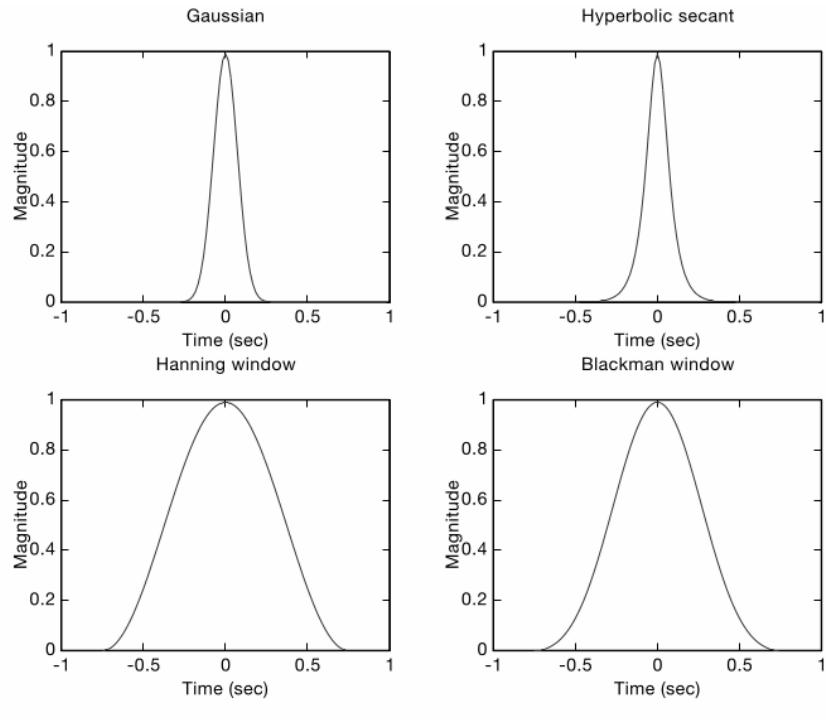
3.1 Short-time Fourier transform

The *short-time Fourier transform* or *continuous Gabor transform* is an example of a method for time-frequency analysis. The short-time Fourier transform can be expressed in several forms and has therefore different interpretations. In this section, several of such interpretations will be treated since they all possess certain characteristics that are interesting for the analysis and synthesis of acoustical signals. The main idea behind the short-time Fourier transform is to obtain a “local frequency domain”, that is, to obtain the spectral content of a function on a finite, time interval. Such a local frequency domain denotes the spectral content of a function on this time interval. A sequence of these local frequency domains next denotes how the spectral content of a function evolves over time.

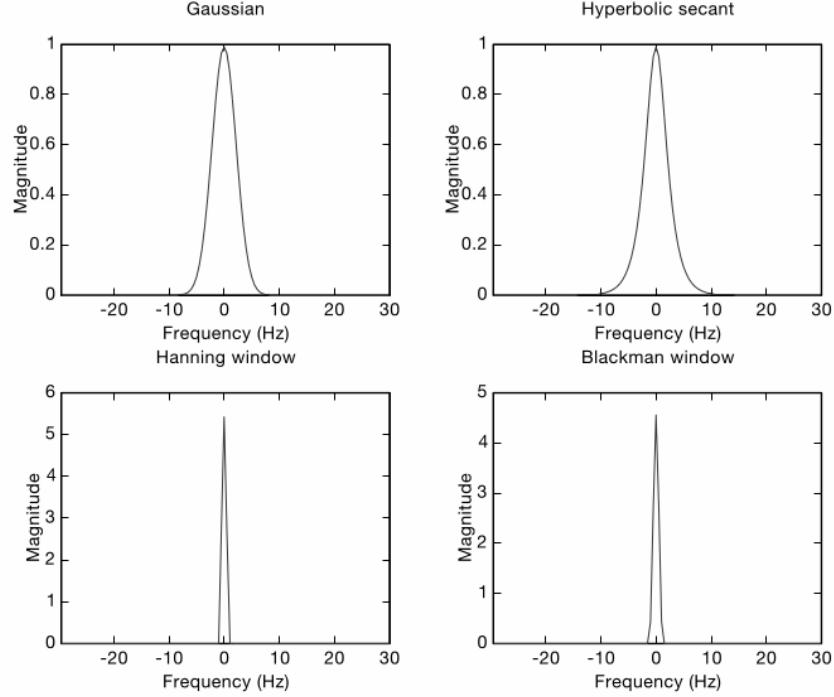
The main form in which the short-time Fourier transform, with respect to the window function $g \in L^2(\mathbb{R})$, is in general expressed for a function $f \in L^2(\mathbb{R})$ is

$$(\mathcal{V}_g f)(\tau, \nu) = \int_{\mathbb{R}} f(t) g^*(t - \tau) e^{-i2\pi\nu t} dt, \quad \nu, \tau \in \mathbb{R}. \quad (3.1)$$

Here, the window function g is a function whose value is zero outside a finite interval, and is in general concentrated around the origin. A function in the time domain whose values are non-zero only on a certain interval was in Chapter 2 called *time-limited*. Furthermore, the time instant around which the energy of a function is centred was in Chapter 2 called the *time centre* — see formula (2.1). A window function could therefore be interpreted as a time-limited function whose time centre equals zero. Examples of window functions are the Gaussian function, the hyperbolic Secant, the Hanning window and the Blackman window. Plots of these examples of window functions are shown in figure 3.1.



(a) Window functions in the time domain



(b) Fourier transform of window functions

Figure 3.1: Examples of window functions in the time and frequency domain

The short-time Fourier transform expressed in the form of (3.1) could be interpreted as the Fourier transform of the product of the function f and a translated version of the window function g , that is, as the Fourier transform of

$$\begin{aligned} f_\tau(t) &= f(t)g^*(t - \tau) \\ &= f(t) \cdot (\mathcal{T}_\tau g^*)(t), \end{aligned}$$

where \mathcal{T}_τ denotes the translation operator as defined in (1.8). Since the window function g with which the function f is multiplied is, by definition, zero outside a certain interval, the function f_τ is a version of f localized around the time instant τ . Since the function f_τ is localized, its Fourier transform equals the short-time Fourier transform, that is,

$$(\mathcal{V}_g f)(\tau, \nu) = (\mathcal{F}f_\tau)(\tau, \nu), \quad \nu, \tau \in \mathbb{R}, \quad (3.2)$$

which gives an indication of the spectral content of a function on the time interval determined by the window function. Note, however, that the short-time Fourier transform does not give an indication of the spectral content of the function f around the time position τ , but that it gives an indication of the spectral content of the product function $f \cdot g^*$, centred around the time instant τ . The characteristics of the window function are therefore of fundamental importance for the characteristics of the short-time Fourier transform.

Instead of expressing the short-time Fourier transform, as defined in (3.1), in terms of the translation operator, it can also be defined in terms of the modulation operator. In this case, the short-time Fourier transform is expressed as the convolution between a function and modulated versions of the window function. The short-time Fourier transform of a function f , with respect to a window function g , is in this case explicitly given by

$$(\mathcal{V}_g f)(\tau, \nu) = f(t) * g_\nu^\diamond(\tau),$$

where g_ν^\diamond denotes

$$\begin{aligned} g_\nu^\diamond(t) &= g^*(-t)e^{i2\pi\nu t} \\ &= \mathcal{E}_\nu g^\diamond, \end{aligned}$$

and where \mathcal{E}_ν is the modulation operator as defined in (1.9). Since convolution can be considered as filtering, the short-time Fourier transform expressed in this form leads to the interpretation of the short-time Fourier transform as the filtering of a function f through the filters g_ν^\diamond , which are time-reversed, modulated versions of a window function. The short-time Fourier transform can therefore be considered as a *filter bank*.

In chapter 1, it was stated that the convolution of two functions in the time domain corresponds to a multiplication of the Fourier transforms of these functions in the frequency domain. For this reason, the short-time Fourier transform could also be expressed in terms of the Fourier transform of a function. This definition of the short-time Fourier transform could be seen as a definition of the short-time Fourier transform in the frequency domain. The definition of the short-time Fourier transform in the frequency domain is given by

$$(\mathcal{V}_g f)(\tau, \nu) = e^{-i2\pi\nu\tau} \int_{\mathbb{R}} \hat{f}(\xi) \hat{g}^*(\xi - \nu) e^{i2\pi\xi\tau} d\xi, \quad \tau, \xi \in \mathbb{R}, \quad (3.3)$$

where \hat{f} denotes the Fourier transform of the analysed function f , and where \hat{g} denotes the Fourier transform of the window function g . The short-time Fourier transform expressed in this form can be interpreted as the inverse Fourier transform of the function $\hat{f}(\xi) \hat{g}^*(\xi - \nu) e^{-i2\pi\nu\tau}$, that is,

$$(\mathcal{V}_g(\tau, \nu)) = (\mathcal{F}^{-1}(\hat{f} \cdot \mathcal{E}_{-\tau} \mathcal{T}_\nu \hat{g}^*))(\tau, \nu).$$

Here, the translated version of \hat{g} , that is, $\hat{g}(\xi - \nu)$ is the window function in the frequency domain localized around the frequency ν . Since this window function is in general concentrated on a finite

frequency interval, the function whose inverse Fourier transform equals the short-time Fourier transform is a modulated version of \hat{f} localized around the frequency ν .

The definition of the short-time Fourier transform in the frequency domain was based on the fact that a convolution in the time domain corresponds to a multiplication in the frequency domain. Yet another form of the short-time Fourier transform can be obtained due to the fact that the convolution operator can be rewritten as an inner product — see formula (1.13). The short-time Fourier transform can thus also be written as an inner product. This leads to the expression of the short-time Fourier transform as the inner product between a function and translated and modulation versions of a window function. If a translated and modulation version of a window function $g \in L^2(\mathbb{R})$ is given by

$$\begin{aligned} g_{\nu,\tau} &= g(t - \tau) e^{i2\pi\nu t} \\ &= \mathcal{E}_\nu \mathcal{T}_\tau g, \end{aligned}$$

then the short-time Fourier transform of a function f can be expressed as

$$(\mathcal{V}_g f)(\tau, \nu) = \langle f, g_{\nu,\tau} \rangle \quad (3.4)$$

$$= \langle f, \mathcal{E}_\nu \mathcal{T}_\tau g \rangle, \quad (3.5)$$

where $\nu, \tau \in \mathbb{R}$. Since an inner product between two functions can be interpreted as the “similarity” or “overlap” between these two functions, the short-time Fourier transform expressed as an inner product can be interpreted as the “similarity” or “overlap” between the function f and the window functions $g_{\nu,\tau}$. Since the functions $g_{\nu,\tau}$ are both concentrated in time and frequency, the overlap between these functions and the function f gives an indication of how f is concentrated in time and frequency. The short-time Fourier transform expressed in this form will be dominant in the rest of this thesis since it leads to the construction of so-called *Gabor frames* — see chapter 5. It is also this form of the short-time Fourier transform that gives a valuable interpretation of the inverse short-time Fourier transform.

The inverse short-time Fourier transform \mathcal{V}_h^{-1} , with respect to a window function $h \in L^2(\mathbb{R})$, of a function $c \in L^2(\mathbb{R}^2)$ is given by

$$(\mathcal{V}_h^{-1} c)(t) = \iint_{\mathbb{R}} c(\tau, \nu) h_{\tau,\nu} d\tau d\nu \quad (3.6)$$

$$= \iint_{\mathbb{R}} c(\tau, \nu) \mathcal{E}_\nu \mathcal{T}_\tau h d\tau d\nu. \quad (3.7)$$

The inverse short-time Fourier transform constructs thus a function in the time domain by representing it as a continuous superposition of translated and modulated versions of a window function. Since all these versions of a window function are concentrated in time and frequency, the inverse short-time Fourier transform is a method to construct functions whose temporal characteristics evolve over frequency or whose spectral characteristics evolve over time.

3.2 Time-frequency resolution of the short-time Fourier transform

In the previous section it was shown that the short-time Fourier transform can be defined both in the time domain and in the frequency domain — see formula (3.1) and (3.3), respectively. These definitions of the short-time Fourier transform are both with respect to a window function. For this reason depends the result of the short-time Fourier transform on this window function. The window function associated with the short-time Fourier transform in the time domain determines the time spread of the function that is mapped into the frequency domain. The window function associated with the short-time Fourier transform in the frequency domain determines the frequency spread of the function that is mapped into the time domain. The window function associated with the short-time Fourier transform determines therefore how the analysed function is presented on the time-frequency plane by the short-time Fourier transform. The shape of the window function

is therefore a fundamental factor that defines the spectral and temporal characteristics that the short-time Fourier transform possesses.

In the case that the window function associated with the short-time Fourier transform in the time domain is centred around the time centre \bar{t} and has a time spread Δ_t , then a by τ translated version of this window function is concentrated on the interval given by

$$[\tau + \bar{t} - \Delta_t, \tau + \bar{t} + \Delta_t].$$

This interval is thus the time interval on which a translated version of a window function is concentrated. Note that this interval is equal to the time interval given in (2.3) except for the factor τ , which is caused by the translation of the window function by τ . In order to ease notation, the interval $[\tau + \bar{t} - \Delta_t, \tau + \bar{t} + \Delta_t]$ will from now on be denoted by $\mathbf{I}_g(\tau)$, where g is the window function that has a time centre \bar{t} and a time spread Δ_t , and where τ is the translation parameter by which g is translated. Since, by definition, a window function is zero-valued outside a certain interval, the value of g is trivially zero outside the interval $\mathbf{I}_g(\tau)$. By the definition of a window function it was also stated that a window function is in general concentrated around the origin, that is, its time centre is 0. If this is indeed the case, then the interval $\mathbf{I}_g(\tau)$ reduces to $[\tau - \Delta_t, \tau + \Delta_t]$.

Just as one can obtain the time interval on which a translated window function is concentrated, one can also obtain the frequency interval on which its Fourier transform is concentrated. In order to do so, recall that a translated window function in the time domain corresponds to a modulated version of the Fourier transform of the window function. In the case that the window function \hat{g} , associated with the short-time Fourier transform in the frequency domain, is centred around the frequency centre $\bar{\xi}$, and has a frequency spread Δ_ξ , then a by ν modulated version of \hat{g} is concentrated on the interval given by

$$[\nu + \bar{\xi} - \Delta_\xi, \nu + \bar{\xi} + \Delta_\xi].$$

This interval is the frequency interval on which the window function is concentrated. Note that this interval is equal to the frequency interval given in (2.6) except for the factor ν , which is caused by the modulation of the window function by ν . This frequency interval associated with the window function g will from now on shortly be denoted by $\mathbf{I}_{\hat{g}}(\nu) := [\nu + \bar{\xi} - \Delta_\xi, \nu + \bar{\xi} + \Delta_\xi]$.

From the intervals $\mathbf{I}_g(\tau)$ and $\mathbf{I}_{\hat{g}}(\nu)$, it can next be deduced that the window function g is concentrated on the area in the time-frequency plane given by

$$[\tau + \bar{t} - \Delta_t, \tau + \bar{t} + \Delta_t] \times [\nu + \bar{\xi} - \Delta_\xi, \nu + \bar{\xi} + \Delta_\xi]. \quad (3.8)$$

The window function is thus concentrated on this area in the time-frequency domain. This area has a rectangular shape and is the Heisenberg box of the by τ translated and by ν modulated window function g . In order to get an impression on how this Heisenberg box looks on the time-frequency plane, note that in figure 2.1 examples of a by 0 translated and a by 0 modulated version of a window function are shown. To ease the notation, the area will from now on be denoted by

$$\begin{aligned} \mathbf{A}_g(\tau, \nu) &:= [\tau + \bar{t} - \Delta_t, \tau + \bar{t} + \Delta_t] \times [\nu + \bar{\xi} - \Delta_\xi, \nu + \bar{\xi} + \Delta_\xi] \\ &:= \mathbf{I}_g(\tau) \times \mathbf{I}_{\hat{g}}(\nu). \end{aligned}$$

Note that this area is equal to the area defined in (2.8) except for the factors τ and ν , which are caused by the translation respectively the modulation of g . The position of $\mathbf{A}_g(\tau, \nu)$ on the time-frequency plane is primarily determined by the variables τ and ν , but also by the time centre \bar{t} and the frequency centre $\bar{\xi}$ of g , if they are non-zero. The size of $\mathbf{A}_g(\tau, \nu)$ is entirely determined by the time spread Δ_t and the frequency spread Δ_ξ of the window function g . Due to the uncertainty principle expressed in (2.7), the time spread Δ_t and the frequency spread Δ_ξ of the window function can't be both arbitrary small. Therefore, the area $\mathbf{A}_g(\tau, \nu)$ on which it is concentrated in the time-frequency domain can't be arbitrary small. The form of this area is, as its size, determined by the time spread Δ_t and the frequency spread Δ_ξ of g . Since both Δ_t and Δ_ξ are constant for all positions τ and ν , and for all \bar{t} and $\bar{\xi}$, both the size of the area $\mathbf{A}_g(\tau, \nu)$ and the shape of $\mathbf{A}_g(\tau, \nu)$ are constant for all points (τ, ν) on the time-frequency plane. Thus, the time-frequency resolution that the short-time Fourier transform possesses is equal at each position in the time-frequency plane.

3.2.1 Examples

In the next paragraphs several representations of translated and modulated versions of Gaussian window function will be given. The representations that are given are the graphs of the Gaussian window functions in the time domain, the frequency domain and the time-frequency domain. The main reason that there is chosen for a Gaussian window function is because it possesses a version whose time and frequency spread are equal, a version whose time spread is greater than its frequency spread and a version whose time spread is less than its frequency spread.

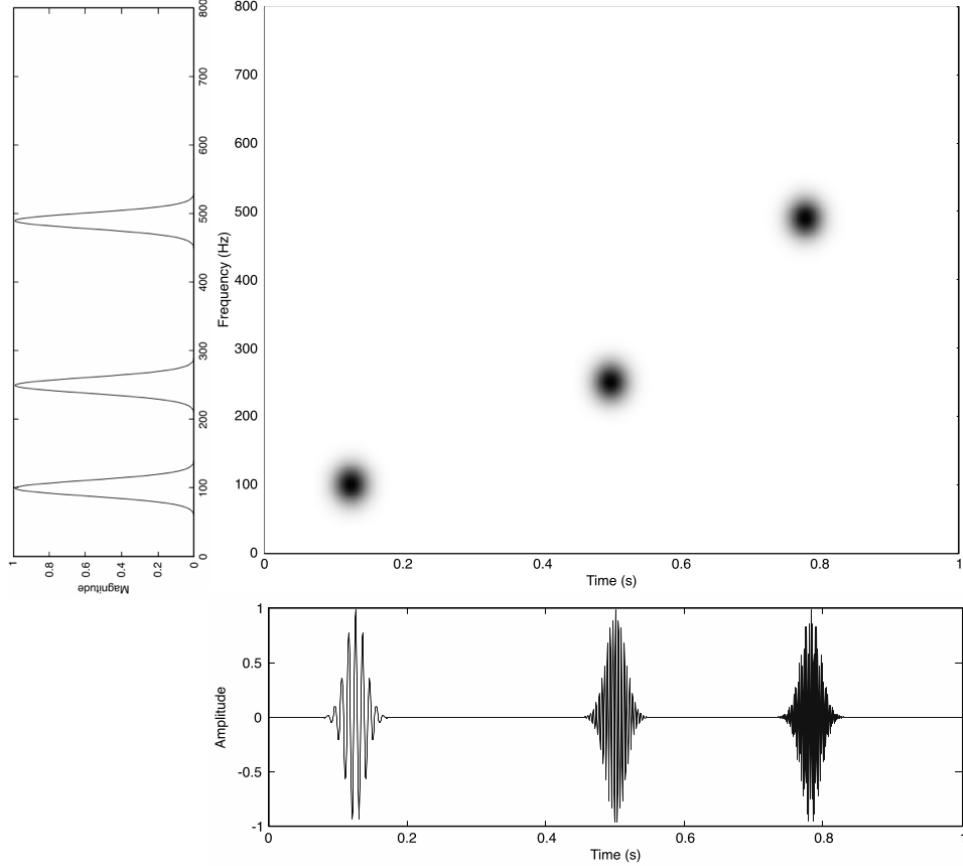


Figure 3.2: Representations of a translated and modulated Gaussian window function whose time and frequency spread are equal.

Example 3.1 In figure 3.2 several representations of translated and modulated versions of a Gaussian window function, of which the time and the frequency spread are equal, are shown. The bottom representation shows the translated and modulated versions of this Gaussian window function in the time domain; the left representation shows them in the frequency domain; and the top-right representation shows them on the time-frequency plane. Since the time and frequency spread of this Gaussian window function are equal, the size of the intervals $I_g(\tau)$ and $I_{\hat{g}}(\nu)$ are equal, and $A_g(\tau, \nu)$ corresponds to a square on the time-frequency plane centred around the point (τ, ν) . Since the size of the Gaussian is invariant under translation and modulation, the different translated and modulated versions of the Gaussian window function all correspond to the same area $A_g(\tau, \nu)$, but at different points (τ, ν) on the time-frequency plane.

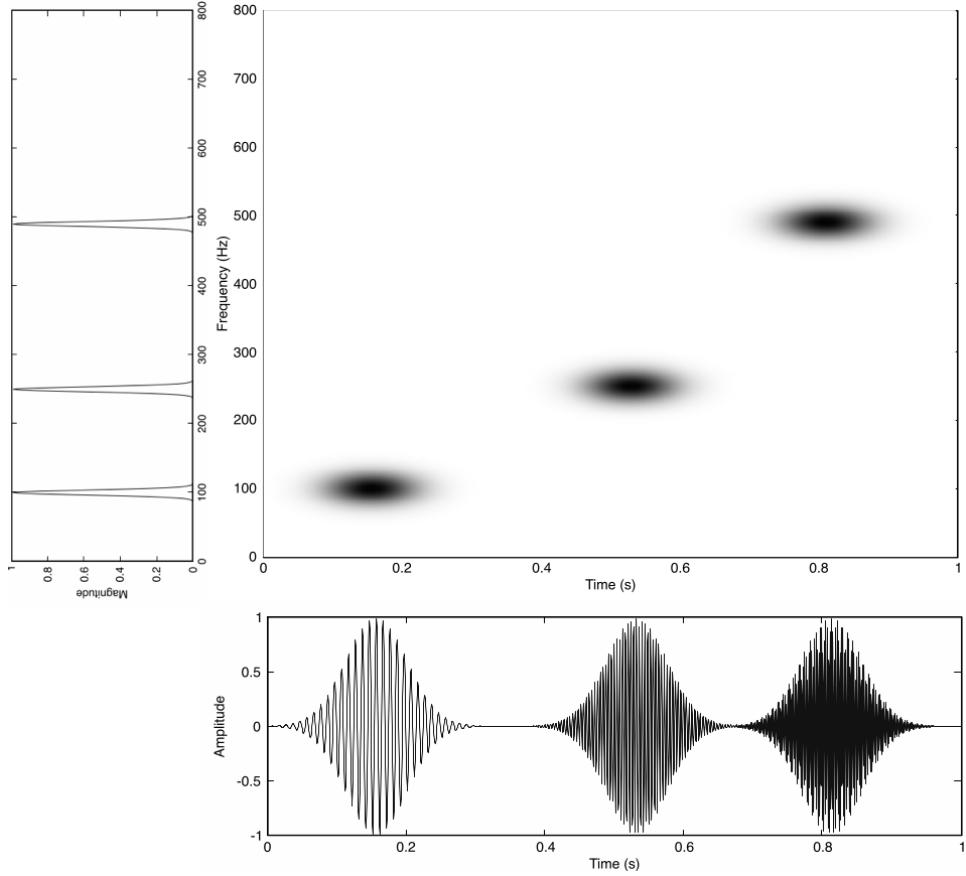


Figure 3.3: Representations of a translated and modulated Gaussian window function whose time spread is greater than its frequency spread.

Example 3.2 In figure 3.3 several representations of translated and modulated versions of a Gaussian window function, of which the time spread is greater than its frequency spread, are shown. The bottom representation shows the translated and modulated versions of this Gaussian window function in the time domain; the left representation shows them in the frequency domain; and the top-right representation shows them on the time-frequency plane. Since the time spread and the frequency spread of this Gaussian window function aren't equal, the size of the intervals $\mathbf{I}_g(\tau)$ and $\mathbf{I}_{\hat{g}}(\nu)$ aren't equal, and $\mathbf{A}_g(\tau, \nu)$ corresponds to a rectangular shape other than a square on the time-frequency plane. Since the size of the Gaussian is invariant under translation and modulation, the different translated and modulated versions of the Gaussian window function all correspond to the same area $\mathbf{A}_g(\tau, \nu)$, but at different points (τ, ν) on the time-frequency plane.

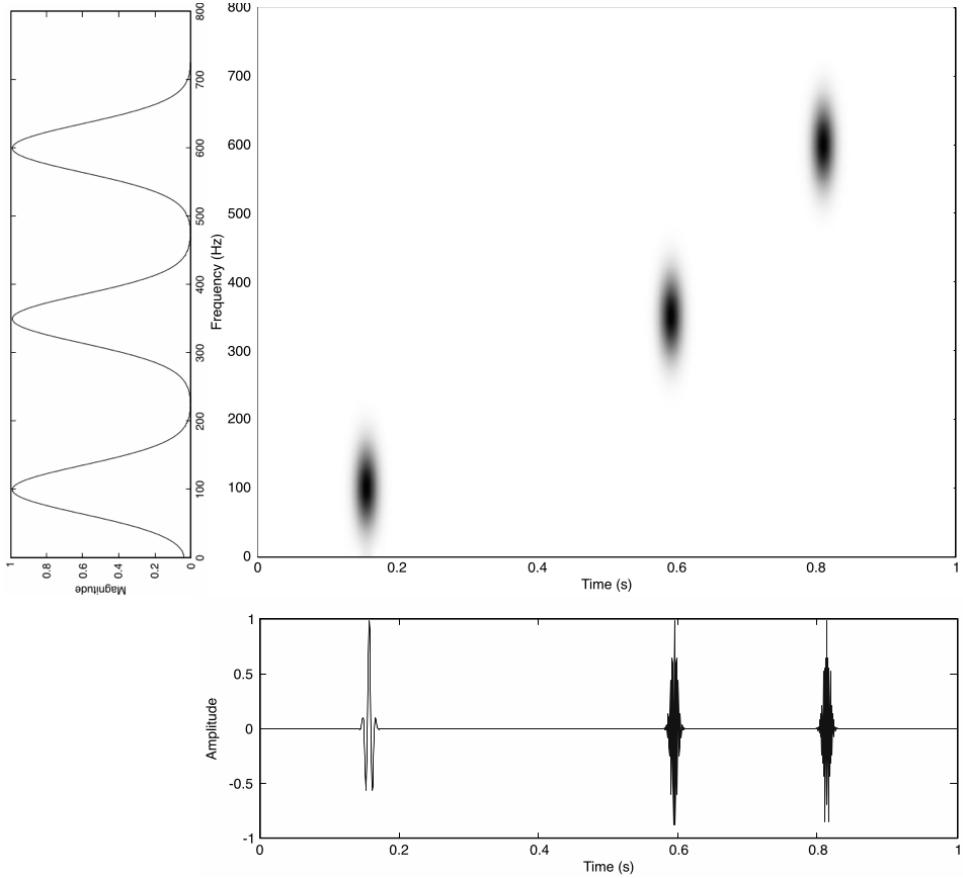


Figure 3.4: Representations of a translated and modulated Gaussian window function whose time spread is less than its frequency spread.

Example 3.3 In figure 3.4 several representations of translated and modulated versions of a Gaussian window function, of which the time spread is less than its frequency spread, are shown. The bottom representation shows the translated and modulated versions of this Gaussian window function in the time domain; the left representation shows them in the frequency domain; and the top-right representation shows them on the time-frequency plane. Since the time and frequency spread of this Gaussian window function aren't equal, the size of the intervals $\mathbf{I}_g(\tau)$ and $\mathbf{I}_{\hat{g}}(\nu)$ aren't equal, and $\mathbf{A}_g(\tau, \nu)$ corresponds to a rectangular shape other than a square on the time-frequency plane. Since the size of the Gaussian is invariant under translation and modulation, the different translated and modulated versions of the Gaussian window function all correspond to the same area $\mathbf{A}_g(\tau, \nu)$, but at different points (τ, ν) on the time-frequency plane.

3.3 Time-frequency representations based on the short-time Fourier transform

In the previous sections it was shown that the short-time Fourier transform of a function provides a representation of this function on the time-frequency plane. The short-time Fourier transform of a function f , with respect to a window function g , provides at each point (τ, ν) of the time-frequency plane one value, namely $(\mathcal{V}_g f)(\tau, \nu)$. This value of the short-time Fourier transform is complex-valued, and it therefore possesses different forms in which it can be represented. All these different representations are valuable for analysing the temporal and spectral characteristics of functions. Due to the non-invertibility of these representations, these representations aren't directly applicable for the synthesis or modification of functions. However, very recently some

attempts were made to use these representations for the synthesis and modification of functions (Holighaus et al., 2015). In this section, several of such representations based on the short-time Fourier transform in polar form will be treated. In this treatment, the main focus will be on the application of these methods for analysis purposes only.

If $\mathcal{V}_g f$ denotes the short-time Fourier transform, with respect to a window function g , of a function f , then the value $|(\mathcal{V}_g f)(\tau, \nu)|$ denotes the magnitude of the short-time Fourier transform of f at (τ, ν) . Since the short-time Fourier transform gives rise to several interpretations, also the magnitude of the short-time Fourier transform has several interpretations. For example, in the case that the short-time Fourier transform is interpreted as the Fourier transform of a function f localized at τ (as in formula (3.2)), then $|(\mathcal{V}_g f)(\tau, \nu)|$ could be interpreted as the magnitude of the spectral component $e^{i2\pi t\nu}$ localized at the time instant τ . Another example, in the case that the short-time Fourier transform is interpreted as the inner product between a function f and translated and modulated versions of a window function g (as in formula (3.4)), then the value $|(\mathcal{V}_g f)(\tau, \nu)|$ may be interpreted as the “similarity” between the magnitude of f and the window function g . Whatever the interpretation of the short-time Fourier transform is, the value $|(\mathcal{V}_g f)(\tau, \nu)|$ corresponds always to a magnitude at the time instant τ and the frequency ν . The value of this magnitude is a non-negative number, and has an obvious interpretation due to its relation to energy. Based on this relation between amplitude and energy, the function $|\mathcal{V}_g f|^2$ denotes an energy distribution in the time-frequency plane, and is in general called the *spectrogram*. Each value of the spectrogram can be interpreted as an energy at time τ and frequency ν .

The complex-valued short-time Fourier transform consists aside a magnitude also of an argument or phase. If $\mathcal{V}_g f$ denotes the short-time Fourier transform, with respect to a window function g , of a function f , then the value $\arg(\mathcal{V}_g f)(\tau, \nu)$ denotes its phase at the point (τ, ν) . Since the short-time Fourier transform gives rise to several interpretations, also the argument of the short-time Fourier transform has several interpretations. For example, in the case that the short-time Fourier transform is interpreted as the Fourier transform of a function f localized at τ (as in formula (3.2)), then $\arg(\mathcal{V}_g f)(\tau, \nu)$ could be interpreted as the phase of the spectral component $e^{i2\pi t\nu}$ that is localized at time τ . Another example, in the case that the short-time Fourier transform is interpreted as the inner product between a function f and translated and modulated versions of a window function g (as in formula (3.4)), then the value $\arg(\mathcal{V}_g f)(\tau, \nu)$ may be interpreted as the “similarity” between the phase of the function f and the window function g . In both interpretations, the value $\arg(\mathcal{V}_g f)(\tau, \nu)$ corresponds to a phase at time τ and frequency ν . This phase is a value in the interval $]-\pi, \pi]$, and the interpretation of this value is in general not obvious. However, the information that the phase possesses is important for the modification of the short-time Fourier transform or the reconstruction of the analysed function from the short-time Fourier transform. An obvious representation of the information that the phase possesses is therefore valuable. Such a representation is the derivative of the phase over time or over frequency. These derivatives of the phase of the short-time Fourier transform with respect to time and frequency can be interpreted as a *local instantaneous frequency* and a *local group delay*, respectively. Both properties are called *local* since they are not the instantaneous frequency and the group delay of a function in general, but of an in time respectively in frequency localized version of this function.

If $\phi(\tau, \nu) := \arg(\mathcal{V}_g f)(\tau, \nu)$ where $(\mathcal{V}_g f)(\tau, \nu)$ denotes the short-time Fourier transform of f , with respect to the window function g , then the local instantaneous frequency Φ_f of f is defined as

$$\Phi_f(\tau, \nu) := \nu + \frac{1}{2\pi} \frac{\delta}{\delta \tau} \phi(\tau, \nu). \quad (3.9)$$

For each value τ and ν , the local instantaneous frequency Φ_f corresponds to one value $\Phi_f(\tau, \nu)$. This value is the centroid in frequency of the short-time Fourier transform localized around (τ, ν) , and it denotes the frequency $\xi \in \mathbf{I}_g(\nu)$ around which the value $(\mathcal{V}_g f)(\tau, \nu)$ is concentrated. Note that the values of $\Phi_f(\tau, \nu)$ and the values $|(\mathcal{V}_g f)(\tau, \nu)|$ are the values that together form the values of the so-called *phase vocoder* (Flanagan & Golden, 1966).

Just as the standard instantaneous frequency of a function can be computed through a derivative and a centroid, also the local instantaneous frequency associated with the short-time Fourier

transform, as defined in formula (3.9), can be computed as a centroid (Auger & Flandrin, 1995). In this case, the localized centroid in frequency is given by

$$\Phi_f(\tau, \nu) = \nu + \Im \left\{ \frac{(\mathcal{V}_{g'} f)(\tau, \nu) \cdot (\mathcal{V}_g f)^*(\tau, \nu)}{|(\mathcal{V}_g f)(\tau, \nu)|^2} \right\}, \quad (3.10)$$

which could next be reduced to

$$\Phi_f(\tau, \nu) = \nu + \Im \left\{ \frac{(\mathcal{V}_{g'} f)(\tau, \nu)}{(\mathcal{V}_g f)(\tau, \nu)} \right\}, \quad (3.11)$$

where the window function g' in both (3.10) and (3.11) denotes the derivative of g , that is, $g'(t) := \frac{d}{dt}g(t)$. The local centroid in frequency in formula (3.10) is thus defined in terms of three short-time Fourier transforms, namely one with respect to the window function g' , one conjugated version with respect to the window function g , and the spectrogram with respect to the window function g . The local centroid in frequency in formula (3.11) is defined with respect to two short-time Fourier transforms, namely one with respect to the window function g' , and one with respect to the window function g .

If $\theta(\tau, \nu) := \arg(\mathcal{V}_g f)(\tau, \nu)$ where $\mathcal{V}_g f$ denotes the short-time Fourier transform of f , with respect to the window function g , then the local group delay of f is defined as

$$\Theta_f(\tau, \nu) := -\frac{1}{2\pi} \frac{\delta}{\delta \nu} \theta(\tau, \nu).$$

For each value τ and ν , the local group delay Θ_f corresponds to one value $\Theta_f(\tau, \nu)$. This value is the centroid in time of the short-time Fourier transform localized around (τ, ν) , and it denotes the time instant $t \in \mathbf{I}_g(\tau)$ around which the value $(\mathcal{V}_g f)(\tau, \nu)$ is concentrated. Just as the centroid in frequency could be computed from spectrograms, also the centroid in time can be computed from spectrograms (Auger & Flandrin, 1995). In this case, the centroid in time is given by

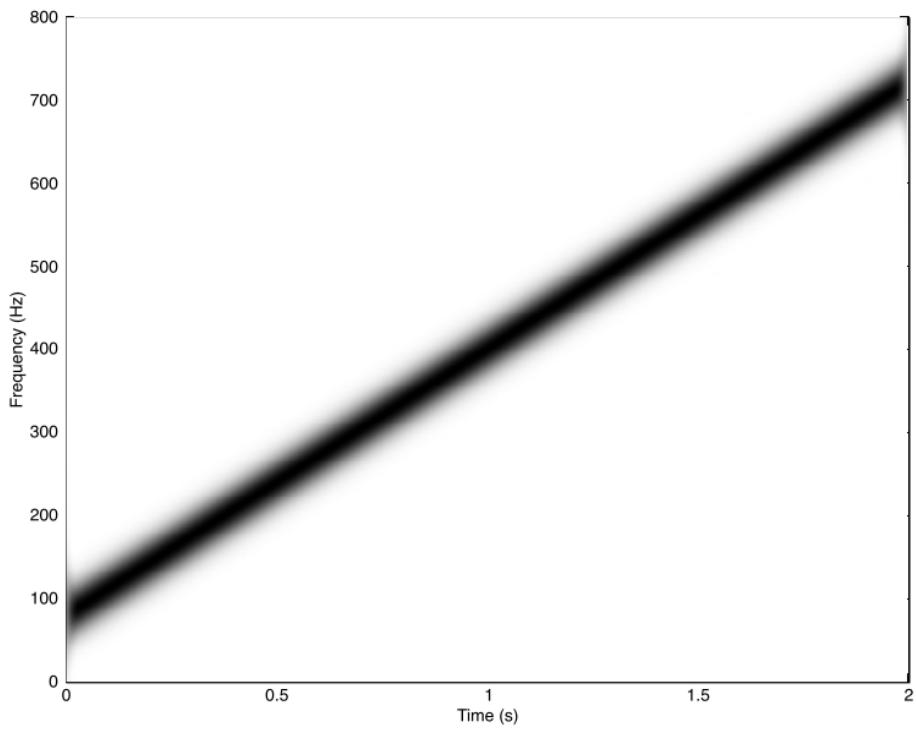
$$\Theta_f(\tau, \nu) = \tau - \Re \left\{ \frac{(\mathcal{V}_{T_g} f)(\tau, \nu) \cdot (\mathcal{V}_g f)^*(\tau, \nu)}{|(\mathcal{V}_g f)(\tau, \nu)|^2} \right\}, \quad (3.12)$$

which could next be reduced to

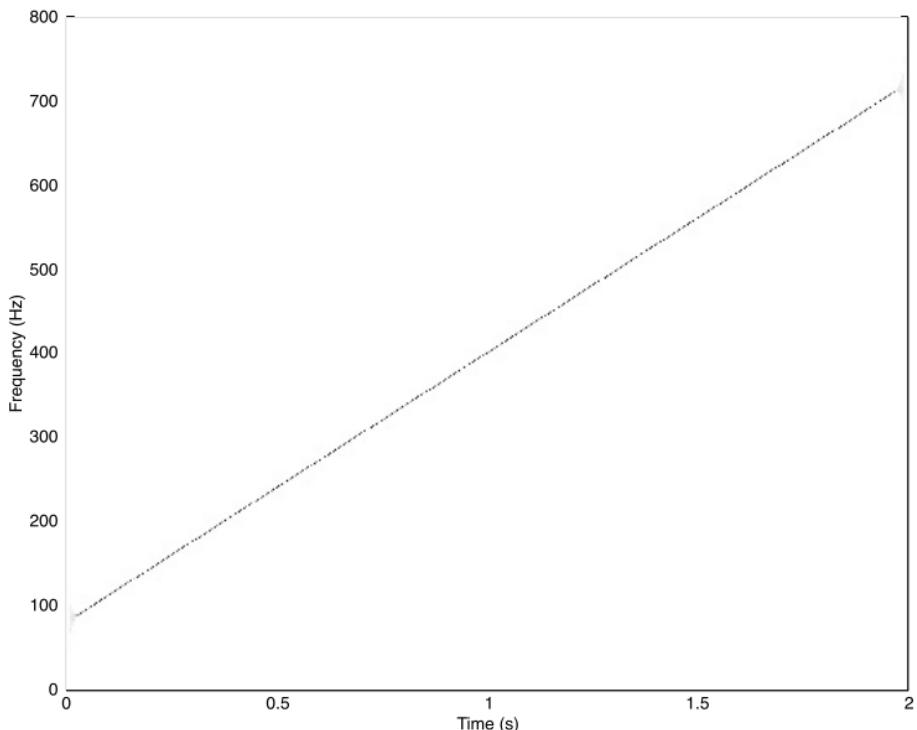
$$\Theta_f(\tau, \nu) = \tau - \Re \left\{ \frac{(\mathcal{V}_{T_g} f)(\tau, \nu)}{(\mathcal{V}_g f)(\tau, \nu)} \right\}, \quad (3.13)$$

where the window function T_g in both (3.12) and (3.13) is the function defined by $T_g(t) := t \cdot g(t)$. The local centroid in time in formula (3.12) is therefore defined in terms of three short-time Fourier transforms, namely one with respect to the window function T_g , one conjugated version with respect to the window function g , and the spectrogram with respect to the window function g . The local centroid in time in formula (3.13) is defined with respect to two short-time Fourier transforms, namely one with respect to the window function T_g , and one with respect to the window function g .

The definitions of the local centroids in time and frequency lead to the so-called *reassigned spectrogram*. The main difference between an ordinary spectrogram and the reassigned spectrogram is that the value $|(\mathcal{V}_g f)(\tau, \nu)|^2$ is assigned to the points $(\Phi_f(\tau, \nu), \Theta_f(\tau, \nu))$ instead of (τ, ν) , the point where the ordinary spectrogram was computed. This reassignment has the advantage that the values of the reassigned spectrogram better-reflect the actual energy of the analysed function than the ordinary spectrogram does — see figure 3.5.



(a) Spectrogram of a linear frequency modulated function



(b) Reassigned spectrogram of a linear frequency modulated function

Figure 3.5: Spectrograms of a linear frequency modulated function

Chapter 4

Wavelet theory

Wavelet theory aims to represent a function as a sum of translated and dilated versions of so-called *wavelet functions*. Such a representation of a function corresponds to a representation of this function in the so-called *time-scale domain*. The mapping that maps a function from the time domain to the time-scale domain is, in the case of wavelet theory, called the *wavelet transform*. Its inverse, the *inverse wavelet transform*, maps a function from the time-scale domain to the time domain. Using an identification that relates the dilation of a function to a value of frequency, the wavelet transform can also be considered as a time-frequency transform. Therefore, the wavelet transform admits aside a time-scale representation also a time-frequency representation. The wavelet transform and its inverse are therefore methods that can be used for the analysis respectively synthesis of functions whose spectral characteristics evolve over time or whose temporal characteristics evolve over frequency. The wavelet transform is for this purpose thus an alternative for the Gabor transform.

The wavelet transform possesses in general two versions, namely a *continuous* and a *discrete* version. In this chapter, only the continuous version will be treated. A discrete version of the wavelet transform will be covered in Chapter 5, in the context of wavelet frames. For a more in-depth discussion of both the continuous and discrete wavelet transform, the reader is referred to Daubechies' book (Daubechies, 1992). For a more extensive discussion of the applications of the wavelet transforms for the analysis and synthesis of acoustical signals, the reader is referred to the classical paper by Kronland-Martinet, Morlet & Grossmann (Kronland-Martinet et al., 1987).

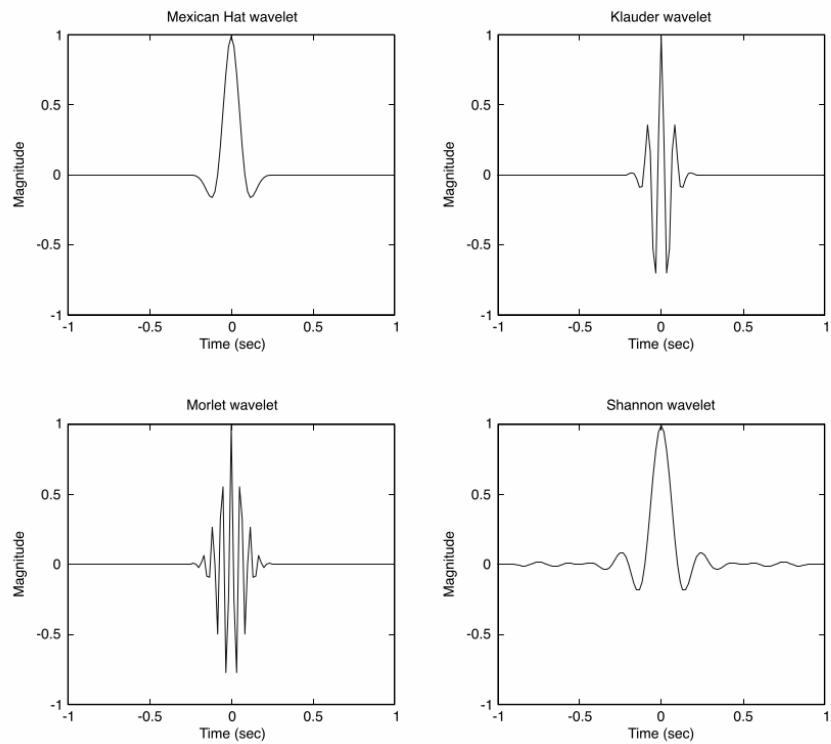
4.1 Wavelet transform

The *continuous wavelet transform* is a method that can be used for both time-frequency and time-scale analysis. Just as the Gabor transform can be represented in several forms, also the wavelet transform can be represented in several forms. All these different forms lead to different interpretations, which all offer a different viewpoint that might be of interest for the analysis and synthesis of acoustical signals. For this reason, several of such different forms of the wavelet transform and their different interpretations will be treated in this section.

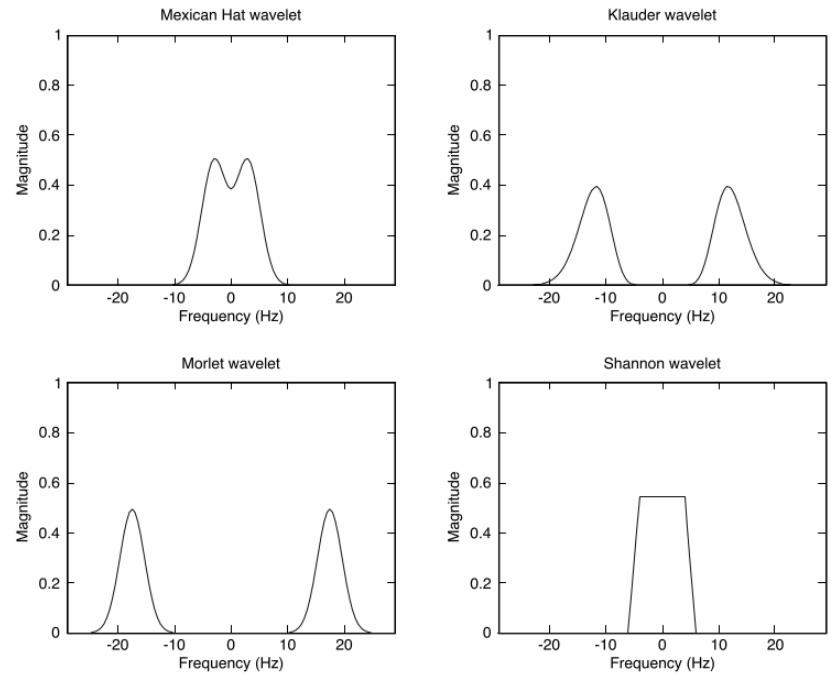
The main form in which the wavelet transform, with respect to a wavelet function $h \in L^2(\mathbb{R})$, of a function $f \in L^2(\mathbb{R})$ is in general expressed is

$$(\mathcal{W}_h f)(\tau, \rho) = \frac{1}{\sqrt{|\rho|}} \int_{\mathbb{R}} f(t) h^* \left(\frac{t - \tau}{\rho} \right) dt. \quad (4.1)$$

Here the wavelet function h is, by definition, a function that is zero-valued outside a certain time interval and whose mean is zero. This leads to the interpretation of such a function as a “little wave”, which is the literal meaning of the word “wavelet”. A function in the time domain whose values are non-zero only on a certain time interval was in chapter 2 called *time-limited*. A wavelet could therefore be seen as a time-limited function with a zero mean. Examples of such functions are the Mexican Hat wavelet, Klauder wavelet, Morlet wavelet, and Shannon wavelet — see Figure 4.1.



(a) Wavelet functions in the time domain



(b) Fourier transform of wavelet functions

Figure 4.1: Examples of wavelet functions

The wavelet transform, as expressed in (4.1), could also be expressed in terms of a convolution. In this case, the wavelet transform should be expressed as the convolution between a function f and dilated versions of the wavelet function h , in which case it is explicitly given by

$$(\mathcal{W}_h f)(\tau, \rho) = f(\tau) * h_\rho^\diamond(\tau),$$

where h_ρ^\diamond denotes

$$\begin{aligned} h_\rho^\diamond(t) &= \frac{1}{\sqrt{|\rho|}} h^* \left(\frac{-t}{\rho} \right) \\ &= (\mathcal{D}_\rho h^\diamond)(t), \end{aligned}$$

and where \mathcal{D}_ρ is the dilation operator as defined in (1.10). Since convolution can be considered as filtering, the wavelet transform expressed in this form leads to the interpretation of the wavelet transform as the filtering of a function f through the filters h_ρ^\diamond , which are time-reversed, dilated versions of a wavelet function. The wavelet transform can therefore be considered as a *filter bank*.

The short-time Fourier transform, which was introduced in the previous chapter, can be expressed in both the time domain and the frequency domain since it could be expressed as a convolution in the time domain. Since the wavelet transform can also be expressed as a convolution in the time domain, it possesses also a definition in the frequency domain. The definition of the wavelet transform in the frequency domain is given by

$$(\mathcal{W}_h f)(\tau, \rho) = \sqrt{|\rho|} \int_{\mathbb{R}} \hat{f}(\xi) \hat{h}^*(\rho\xi) e^{i2\pi\xi\tau} d\xi$$

where \hat{f} denotes the Fourier transform of the function f , and where \hat{h} denotes the Fourier transform of the wavelet function h . The wavelet transform expressed in this form can be interpreted as the inverse Fourier transform of the function $\hat{f}(\xi) \sqrt{|\rho|} h(\rho\xi)$, that is,

$$(\mathcal{W}_h f)(\tau, \rho) = (\mathcal{F}^{-1}(\hat{f} \cdot \mathcal{D}_{\frac{1}{\rho}} \hat{h}^*))(\tau, \rho) \quad (4.2)$$

Here, the dilated version of \hat{h} , that is, $\sqrt{|\rho|} h(\rho\xi)$, is the wavelet function h in the frequency domain dilated by ρ^{-1} . In general, the value ρ , by which a function is dilated, doesn't give a direct indication of the frequency ν around which \hat{h} is localized. However, the dilation parameter ρ and modulation parameter ν can be related through a formal identification that is based on the frequency centre $\bar{\xi}$ of the wavelet function \hat{h} (Flandrin, 1998). This formal identification is defined as

$$\nu = \frac{\bar{\xi}}{\rho}. \quad (4.3)$$

The position of a dilated version of a wavelet function in the frequency domain is thus completely determined by the frequency centre of this wavelet function and the dilation parameter. Therefore, the position of a translated and dilated version of wavelet function on the time-frequency plane is completely determined by its frequency centre, the dilation parameter and the translation parameter. Using the identification given in (4.3), the wavelet transform admits aside its time-scale representation also a time-frequency interpretation.

Just as the short-time Fourier transform can also be expressed in terms of an inner product since it could be written in terms of a convolution, so can the wavelet transform also be expressed as an inner product since it can be written in terms of a convolution. This leads to the expression of the wavelet transform of a function as the inner product between this and translated and dilated versions of the wavelet function. If a translated and dilated version of a window function h is given by

$$\begin{aligned} h_{\tau, \rho}(t) &= \frac{1}{\sqrt{|\rho|}} h \left(\frac{t - \tau}{\rho} \right) \\ &= (\mathcal{T}_\tau \mathcal{D}_\rho h)(t), \end{aligned}$$

then the wavelet transform of a function f can be expressed as

$$\begin{aligned} (\mathcal{W}_h f)(\tau, \rho) &= \langle f, h_{\tau, \rho} \rangle \\ &= \langle f, \mathcal{T}_\tau \mathcal{D}_\rho h \rangle, \end{aligned}$$

where $\rho, \tau \in \mathbb{R}$, $\rho \neq 0$. The wavelet transform expressed in this form can be interpreted as the “similarity” or “overlap” between f and the functions $h_{\tau, \rho}$. The wavelet transform expressed in this form will be dominant in the rest of this thesis since the translated and dilated wavelet functions lead to the construction of so-called *wavelet frames* — see chapter 5. It is also this form that gives a valuable interpretation to the inverse wavelet transform.

The inverse wavelet transform \mathcal{W}_h^{-1} , associated with the wavelet function h , of a function $c \in L^2(\mathbb{R}^2)$ is given by

$$(\mathcal{W}_h^{-1} c)(t) = \iint_{\mathbb{R}} \mathcal{W}_h(\tau, \rho) h_{\tau, \rho} \frac{d\tau d\rho}{\rho^2} \quad (4.4)$$

$$= \iint_{\mathbb{R}} \mathcal{W}_h(\tau, \rho) \mathcal{T}_\tau \mathcal{D}_\rho h \frac{d\tau d\rho}{\rho^2}. \quad (4.5)$$

The inverse wavelet transform constructs thus a function in the time domain by representing it as a continuous superposition of translated and dilated versions of a wavelet function. Since all these versions of a wavelet function are concentrated in time and frequency, the inverse wavelet transform is a method to construct functions whose temporal characteristics evolve over frequency or whose spectral characteristics evolve over time.

4.2 Time-frequency resolution of the wavelet transform

The wavelet transform admits, just as the short-time Fourier transform, a definition in the time domain and the frequency domain — see formula (4.1) and (4.2), respectively. These definitions of the wavelet transform are both with respect to a wavelet function. For this reason, the result of the wavelet transform is completely determined by this wavelet function. The wavelet function associated with the wavelet transform in the time domain determines the time spread of the function that is mapped to the frequency domain. The wavelet function associated with the wavelet transform in the frequency domain determines the frequency spread of the function that is mapped into the time domain. The wavelet function associated with the wavelet transform determines therefore how the analysed function is represented on the time-frequency plane by the wavelet transform. For this reason, the shape of the wavelet function is a fundamental factor that defines the spectral and temporal characteristics that the wavelet transform possesses.

In the case that the wavelet function h is centred around the time centre \bar{t} and has a time spread Δ_t , as defined in (2.1) respectively (2.2), then a by ρ dilated version of h positioned at τ is concentrated on the interval given by

$$[\tau + \rho\bar{t} - \rho\Delta_t, \tau + \rho\bar{t} + \rho\Delta_t].$$

This interval is the time interval on which the wavelet function is concentrated. Since a wavelet function is zero-valued outside a certain interval, the value of h is zero outside the interval on which it is concentrated. The interval $[\tau + \rho\bar{t} - \rho\Delta_t, \tau + \rho\bar{t} + \rho\Delta_t]$ will from now on be denoted by $\mathbf{I}_{h_\rho}(\tau)$, where h_ρ denotes the by ρ dilated wavelet function that has a time centre \bar{t} , a time spread Δ_t , and where τ is the translation parameter by which h is centred. This interval $\mathbf{I}_{h_\rho}(\tau)$ differs from the interval $\mathbf{I}_g(\tau)$ of a translated and modulated window function g by that $\mathbf{I}_{h_\rho}(\tau)$ depends on the dilation parameter ρ , whereas $\mathbf{I}_g(\tau)$ is constant. The higher the value of ρ , the larger the interval on which h_ρ is concentrated. On the contrary, the lower the value of ρ , the smaller the interval on which h_ρ is concentrated.

Just as one can obtain the time interval on which a translated and dilated wavelet function is concentrated, one can also obtain the frequency interval on which its Fourier transform is concentrated. In the case that the wavelet function \hat{h} associated with the wavelet transform in the

frequency domain is centred around the frequency centre $\bar{\xi}$ and has frequency spread Δ_ξ , then a by ρ dilated version of \hat{h} is concentrated on the interval given by

$$[\rho^{-1}\bar{\xi} - \rho^{-1}\Delta_\xi, \rho^{-1}\bar{\xi} + \rho^{-1}\Delta_\xi] = [\nu - \rho^{-1}\Delta_\xi, \nu + \rho^{-1}\Delta_\xi]$$

This interval is the frequency interval on which a by ρ dilated wavelet function is concentrated, and will be abbreviated as $\mathbf{I}_{\hat{h}_\rho}(\nu) := [\nu - \rho^{-1}\Delta_\xi, \nu + \rho^{-1}\Delta_\xi]$. Just as the size of the time interval $\mathbf{I}_{h_\rho}(\tau)$ depends on the dilation parameter ρ , also the interval $\mathbf{I}_{\hat{h}_\rho}(\nu)$ depends on ρ . From the relation given in formula (1.11), it follows that a by ρ dilated wavelet function in the time domain corresponds to a by $\frac{1}{\rho}$ dilated wavelet function in the frequency domain. Therefore, the higher the value of ρ , the smaller the interval on which \hat{h}_ρ is concentrated,. On the contrary, the smaller the value of ρ , the larger the interval on which \hat{h}_ρ is concentrated.

From the time interval $\mathbf{I}_{h_\rho}(\tau)$ and frequency interval $\mathbf{I}_{\hat{h}_\rho}(\nu)$ it can be deduced that a by ρ dilated wavelet function is concentrated on the area in the time-frequency plane given by

$$[\tau + \rho\bar{t} - \rho\Delta_t, \tau + \rho\bar{t} + \rho\Delta_t] \times [\nu - \rho^{-1}\Delta_\xi, \nu + \rho^{-1}\Delta_\xi],$$

which will from now on be denoted by

$$\mathbf{A}_{h_\rho}(\tau, \nu) := \mathbf{I}_{h_\rho}(\tau) \times \mathbf{I}_{\hat{h}_\rho}(\nu).$$

The position of this area on the time-frequency plane is determined by the parameters τ and $\nu = \bar{\xi}/\rho$. The size of the area is determined by $\rho\Delta_t$ and Δ_ξ/ρ of the by ρ dilated wavelet function h . Since the time spread Δ_t and frequency spread Δ_ξ of the wavelet function h are constant, the size of the area $\mathbf{A}_{h_\rho}(\tau, \nu)$ is also constant for all τ and ν . However, since ρ varies for different ν , the shape of $\mathbf{A}_{h_\rho}(\tau, \nu)$ differs also for different ρ . Thus, the time-frequency resolution that the wavelet transform possesses isn't constant for all points on the time-frequency plane, but differs per position on the time-frequency plane.

4.2.1 Examples

In the next two paragraphs several representations of a translated and dilated versions of Morlet wavelet functions will be given. The representations that are given are translated and dilated versions of a Morlet wavelet function in the time domain, the frequency domain and the time-frequency domain. Since the Morlet wavelet function can be interpreted as a modulated Gaussian it possesses, as the Gaussian window function, a version whose time and frequency spread are equal, a version whose time spread is greater than its frequency spread and a version whose time spread is less than its frequency spread.

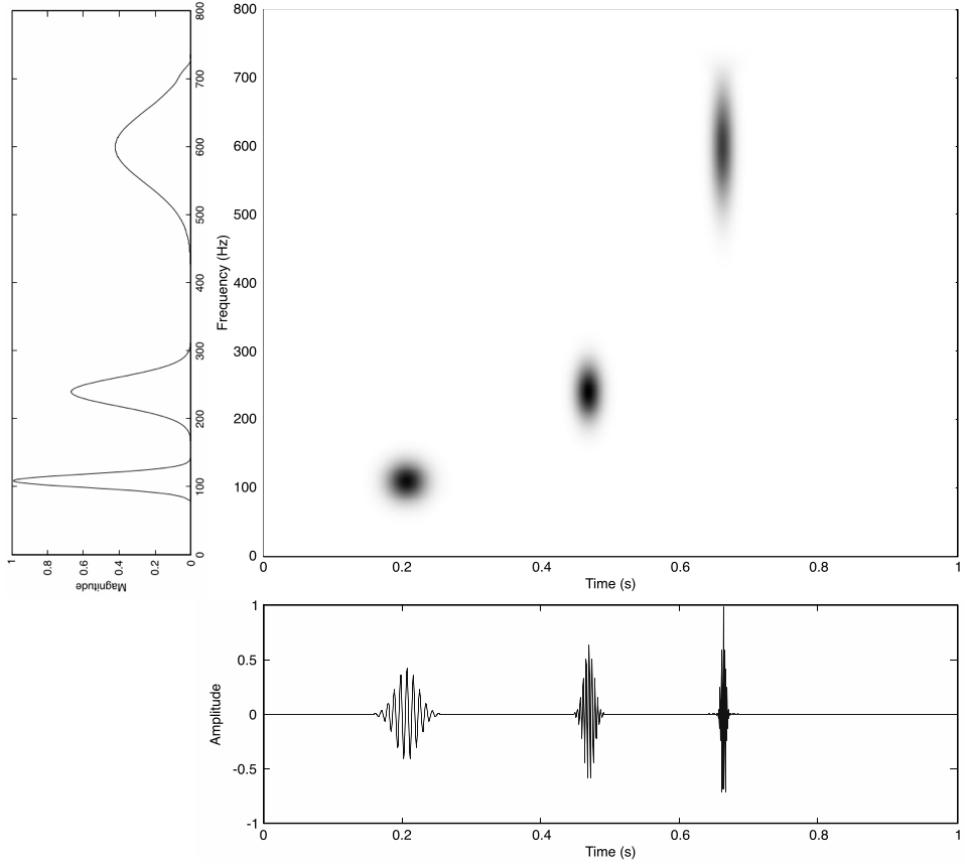


Figure 4.2: Representations of translated and dilated Morlet window functions

Example 4.1 In figure 4.2 several representations of translated and dilated versions of a Morlet wavelets function, of which the time spread and the frequency spread are equal, are shown. The bottom representation shows the translated and modulated versions of a Morlet wavelet function in the time domain; the left representation shows them in the frequency domain; and the top-right representation shows them on the time-frequency plane. Since the time spread and the frequency spread of this Morlet wavelet function are equal, the intervals $\mathbf{I}_{h_1}(\tau)$ and $\mathbf{I}_{\hat{h}_1}(\nu)$ are equal, and $\mathbf{A}_{h_1}(\tau, \nu)$ corresponds to a square on the time-frequency plane. However, since the size of the Morlet wavelet function is variant under dilation, that is, since its differs for different scaled versions, the sizes of the intervals $\mathbf{I}_{h_\rho}(\tau)$ and $\mathbf{I}_{\hat{h}_\rho}(\nu)$ are in general not equal, and $\mathbf{A}_{h_\rho}(\tau, \nu)$ corresponds not to a square for all values of ρ . This means that for different values of ρ the shape of $\mathbf{A}_{h_\rho}(\tau, \nu)$ is also different.

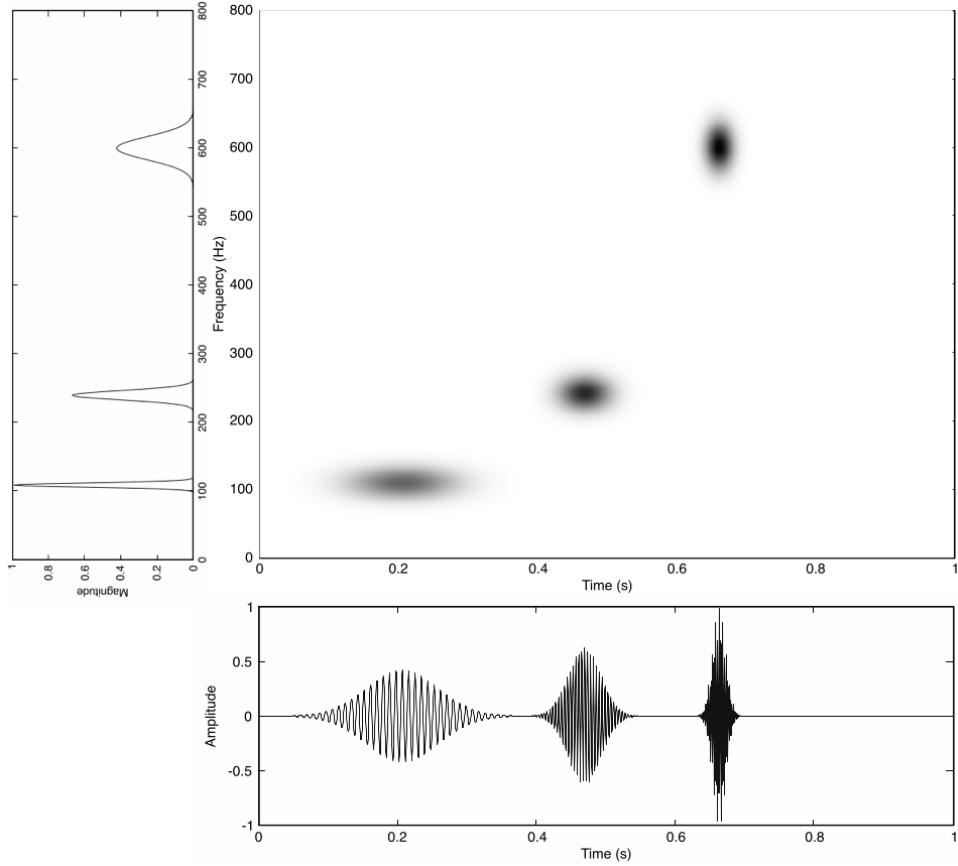


Figure 4.3: Representations of translated and dilated Morlet wavelet functions

Example 4.2 In figure 4.3 several representations of translated and dilated versions of a Morlet wavelet function, of which the time spread is greater than its frequency spread, are shown. The bottom representation shows the translated and dilated versions of this Morlet wavelet function in the time domain; the left representation shows them in the frequency domain; and the top-right representation shows them on the time-frequency plane. Since the time spread and frequency spread of this Morlet wavelet function aren't equal, the intervals $\mathbf{I}_{h_1}(\tau)$ and $\mathbf{I}_{\hat{h}_1}(\nu)$ aren't equal, and $\mathbf{A}_{h_1}(\tau, \nu)$ corresponds to a rectangular shape other than a square on the time-frequency plane. However, since the size of the wavelet function differs for different scaled versions, the sizes of the intervals $\mathbf{I}_{h_\rho}(\tau)$ and $\mathbf{I}_{\hat{h}_\rho}(\nu)$ are in general not equal, and the shape of $\mathbf{A}_{h_\rho}(\tau, \nu)$ differs for different values of ρ (which is clearly visible in Figure 4.3).

4.3 Time-frequency representations based on the wavelet transform

In the previous sections it was shown that the wavelet transform provides a representation of a function on the time-scale plane. Using the identification $\rho = \frac{\bar{\xi}}{v}$ where $\bar{\xi}$ denotes the frequency centre of a wavelet function, the time-scale representation can be linked to a time-frequency representation. The wavelet transform of a function f , with respect to a window function h , provides at each point (τ, ρ) of the time-scale and time-frequency plane one value, namely $(\mathcal{W}_h f)(\tau, \rho)$ respectively $(\mathcal{W}_h f)(\tau, \frac{\bar{\xi}}{v})$. This value of the wavelet transform is complex-valued, and it can therefore be represented in several forms. In this section several of such representations based on the wavelet transform in polar form will be presented.

If $\mathcal{W}_h f$ denotes the wavelet transform, with respect to a wavelet function h , of a function f , then the value $|\mathcal{W}_h f(\tau, \rho)|$ denotes the magnitude of the wavelet transform of f at (τ, ρ) . Since the wavelet transform gives rise to several interpretations, also the magnitude of the wavelet transform has different interpretations. For example, if the wavelet transform is interpreted as the inner product between the function f and translated and dilated versions of a wavelet function h , then the value $|\mathcal{W}_h f(\tau, \rho)|$ can be interpreted as the “similarity” of the magnitude of f and the wavelet function h . The value of this magnitude is non-negative and has an obvious interpretation due to its relation to energy. Based on this relation between magnitude and energy, the function $|\mathcal{W}_h f|^2$ denotes an energy distribution in the time-scale plane and is in general called the *scalogram*. Each value of the scalogram can be interpreted as the energy at time τ and scale ρ .

The complex-valued wavelet transform consists aside a magnitude also of an argument or phase. If $\mathcal{W}_h f$ denotes the wavelet transform, with respect to a wavelet function h , of a function f , then the value $\arg \mathcal{W}_h f(\tau, \rho)$ denotes its phase at time (τ, ρ) . Since the wavelet transform gives rise to several interpretations, also the argument of the wavelet transform has different interpretations. For example, if the wavelet transform is interpreted as the inner product of the analysed function and translated and dilated versions of a wavelet function, then the value $\arg \mathcal{W}_h f(\tau, \rho)$ can be interpreted as “similarity” of the phase of f and the wavelet function h . The value of this phase is a value in the interval $]-\pi, \pi]$, and the interpretation of this value is in general not obvious. However, the information that the phase possesses is fundamental for modifications of the wavelet transform or for the reconstruction of the analysed function from the wavelet transform. An obvious representation of the information that the phase possesses is therefore valuable. Such representations are, just as in Gabor theory, the local centroids. In the case of the wavelet transform, these local centroids are local centroids in time and scale, respectively. Since scale can be related to frequency through the identification given in formula (4.3), also a local centroid in frequency associated with the wavelet transform can be derived.

If $\mathcal{W}_h f$ denotes the wavelet transform, with respect to the wavelet function h , of a function f , then the local centroid in scale Ψ_f of f is given by

$$\Psi_f(\tau, \rho) := -\frac{\rho \bar{\xi}}{\Im \left\{ \frac{(\mathcal{W}_{h'} f)(\tau, \rho) \cdot (\mathcal{W}_h f)^*(\tau, \rho)}{|(\mathcal{W}_h f)(\tau, \rho)|^2} \right\}}, \quad (4.6)$$

which could next be reduced to

$$\Psi_f(\tau, \rho) = -\frac{\rho \bar{\xi}}{\Im \left\{ \frac{(\mathcal{W}_{h'} f)(\tau, \rho)}{(\mathcal{W}_h f)(\tau, \rho)} \right\}}, \quad (4.7)$$

where the wavelet function h' in (4.6) and (4.7) denotes the derivative of h , that is, $h'(t) := \frac{d}{dt} h(t)$. For each value τ and ρ , the local centroid in scale Ψ_f corresponds to one value $\Psi_f(\tau, \rho)$. This value denotes the scale $\rho \in \mathbf{I}_{h_\rho}(\bar{\xi}/\rho)$ around which the value $(\mathcal{W}_h f)(\tau, \rho)$ is concentrated. This value of scale could, in analogy with the term local instantaneous frequency, be called the *local instantaneous scale*. This local instantaneous scale could be converted into a local instantaneous frequency

(Chassande-Mottin & Flandrin, 2001). This can be done by using the following identification

$$\Phi_f(\tau, \rho) := \frac{\bar{\xi}}{\Psi_f(\tau, \rho)}. \quad (4.8)$$

Note that this local instantaneous frequency has still τ and ρ as independent variables. However, due to the relation between the dilation parameter ρ and a modulation parameter ν , a local instantaneous frequency with τ and ν as independent variables can be derived (Auger & Flandrin, 1995). The local instantaneous frequency is then explicitly given by

$$\Phi_f(\tau, \nu) = \nu + \Im \left\{ \frac{(\mathcal{W}_{h'} f)(\tau, \rho) \cdot (\mathcal{W}_h f)^*(\tau, \rho)}{\rho |(\mathcal{W}_h f)(\tau, \rho)|^2} \right\} \quad (4.9)$$

$$= \nu + \Im \left\{ \frac{(\mathcal{W}_{h'} f)(\tau, \rho)}{\rho (\mathcal{W}_h f)(\tau, \rho)} \right\}. \quad (4.10)$$

The definitions of the local centroid in scale as given in the formulae (4.6), (4.7), and the local centroids in frequency as given in (4.8), (4.9) and (4.10), are all expressed in terms of several wavelet transforms. The local centroid in scale as in (4.6) consists of three wavelet transforms, namely one with respect to the window function h' , one conjugated version with respect to h and a scalogram with respect to h . The same wavelet transforms are used in the computation of (4.9). For the computation of (4.7) and (4.10) only two wavelet transforms are used, namely one with respect to h' and one with respect to h .

If $\mathcal{W}_h f$ denotes the wavelet transform, with respect to the wavelet function h , of a function f , then the local centroid in time Θ_f of f is given by

$$\Theta_f(\tau, \rho) := \tau + \Re \left\{ \frac{\rho (\mathcal{W}_{T_h} f)(\tau, \rho) \cdot (\mathcal{W}_h f)^*(\tau, \rho)}{|(\mathcal{W}_h f)(\tau, \rho)|^2} \right\}, \quad (4.11)$$

which could next be reduced to

$$\Theta_f(\tau, \rho) = \tau + \Re \left\{ \frac{\rho (\mathcal{W}_{T_h} f)(\tau, \rho)}{(\mathcal{W}_h f)(\tau, \rho)} \right\}, \quad (4.12)$$

where the wavelet function T_h in both (4.11) and (4.12) is the function defined by $T_h(t) = t \cdot h(t)$. For each value τ and ρ , the local centroid in time corresponds to one value $\Theta_f(\tau, \rho)$. This value is the centroid in time of the wavelet transform localized around (τ, ρ) , and it denotes the time instant $t \in \mathbf{I}_{h_\rho}(\tau)$, where the value $(\mathcal{W}_h f)(\tau, \nu)$ is concentrated. Since the dilation parameter ρ can be related to a frequency parameter ν , the time centroid expressed in (4.11) and (4.12) can be rewritten as a function with time τ and frequency ν as independent variables (Auger & Flandrin, 1995). In this case, formulae (4.11) and (4.12) become

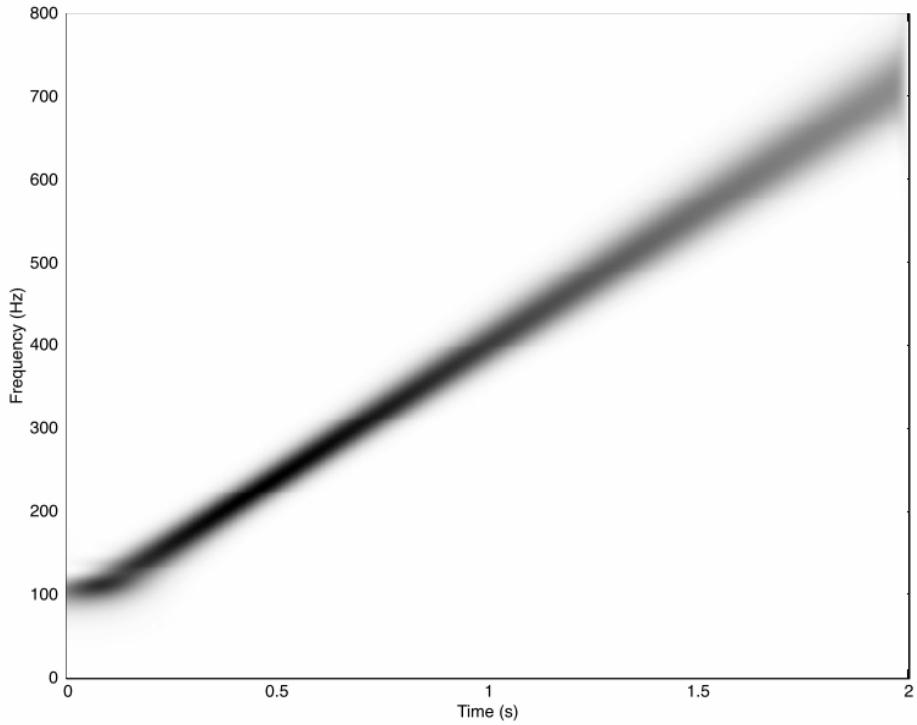
$$\Theta_f(\tau, \nu) = \tau - \Re \left\{ \frac{\rho (\mathcal{W}_{T_h} f)(\tau, \rho) \cdot (\mathcal{W}_h f)^*(\tau, \rho)}{|(\mathcal{W}_h f)(\tau, \rho)|^2} \right\} \quad (4.13)$$

respectively

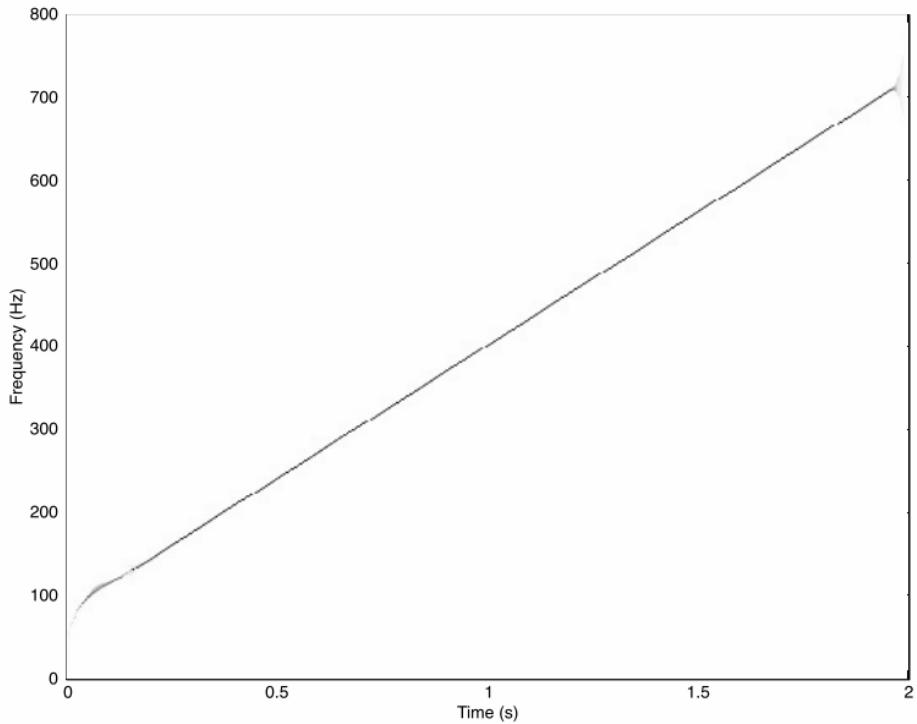
$$\Theta_f(\tau, \nu) = \tau - \Re \left\{ \frac{\rho (\mathcal{W}_{T_h} f)(\tau, \rho)}{(\mathcal{W}_h f)(\tau, \rho)} \right\}. \quad (4.14)$$

The forms of the local centroids in time as given in (4.11), (4.12), (4.13) and (4.14) consist all of several wavelet transforms. The local centroid in time, as defined in the formulae (4.6) and (4.13), consist of three wavelet transforms, namely one with respect to the window function T_h , one conjugated version with respect to h and a scalogram with respect to h . The local centroids in frequency, as given in (4.7) and (4.14), only use two wavelet transforms, namely one with respect to T_h and one with respect to h .

Just as the local centroids in frequency and time lead to the reassigned spectrogram, so does the local centroids in scale and time lead to the *reassigned scalogram*. The reassignment of the value $|(\mathcal{W}_h f)(\tau, \nu)|^2$ to the points $(\Phi_f(\tau, \rho), \Theta_f(\tau, \rho))$ instead of (τ, ρ) has the advantage that the values of the reassigned scalogram better-reflect the actual energy of the analysed function on the time-scale plane than the ordinary scalogram does — see figure 4.4.



(a) Scalogram of a linear frequency modulated function



(b) Reassigned scalogram of a linear frequency modulated function

Figure 4.4: Scalograms of a linear frequency modulated function

Chapter 5

Frame theory

Frame theory allows one to represent a function as a superposition of so-called *elementary functions*. These elementary functions are functions that are in general derived from a single, fixed function by elementary operations. This single, fixed function, from which the elementary functions are derived, is in general called the *atom* or *generator* of these elementary functions. A collection of such elementary functions is in general called a *system*. Such a system is called a *frame* if, and only if, it satisfies the so-called *frame condition*. If a collection of elementary functions forms a frame, then any function can be represented as a superposition of these elementary functions. Each elementary function in such a superposition has a corresponding coefficient, which determines the relative weight of the elementary function in the superposition. These corresponding coefficients can be obtained by an application of an analysis operator on the function which should be represented. Next, by an application of a synthesis operator on these obtained coefficients, the analysed function can be perfectly reconstructed, and is represented by a superposition of elementary functions. The fact that a function can be perfectly reconstructed from a collection of coefficients when the elementary functions form a frame is the main characteristic that distinguishes frames from collections of functions that doesn't form a frame. Thus, if one wants to represent a function as a superposition of elementary functions, then the elementary functions should form a frame. On the contrary, if a collection of elementary functions doesn't form a frame, then a function can't be represented as a superposition of these elementary functions.

In the previous chapters, several examples of specific methods that allow representations of functions as a superposition of elementary functions were given, namely Fourier theory, Gabor theory and wavelet theory. In Fourier theory, Gabor theory and wavelet theory, the elementary functions were complex exponentials, translated and modulated versions of a window function and translated and dilated versions of a wavelet function, respectively. These methods allowed representations of functions as a *continuous superposition* of elementary functions. The collections of elementary functions that were associated with these methods are examples of so-called *continuous frames*. In this thesis, the topic of continuous frames will not be treated further since the technical details that are needed for such a treatment are far beyond the scope of this thesis. However, so-called *discrete frames* form the topic of this chapter. In fact, they form the core of this thesis. In the same manner as it is possible to represent a function as a continuous superposition of elements of a continuous frame, it is possible to represent a function as a *discrete superposition* of elements of a discrete frame. Discrete frames allow thus a representation of a function f as

$$f = \sum_{n \in \mathbb{I}} c_n q_n, \quad (5.1)$$

where $\{c_n\}_{n \in \mathbb{I}}$ is a collection of coefficients, and where $\{q_n\}_{n \in \mathbb{I}}$ is the collection of elementary functions. A representation of a function in the form expressed in (5.1) is in general called an *atomic decomposition*. In order to represent all functions of a certain set of functions as an atomic decomposition, the collection of elementary functions should form a frame for this set of functions. Thus, in order to represent all functions of $L^2(\mathbb{R})$ as an atomic decomposition, the collection of elementary functions should form a frame for $L^2(\mathbb{R})$.

5.1 General frame theory

A collection of functions forms a frame for $L^2(\mathbb{R})$ if it satisfies the so-called *frame condition* or *frame inequality* associated with $L^2(\mathbb{R})$. This frame condition states that a collection of functions $\{f_n\}_{n \in \mathbb{I}}$ forms a frame for $L^2(\mathbb{R})$ if there exist two positive constants A, B such that

$$A\|f\|_2^2 \leq \sum_{n \in \mathbb{I}} |\langle f, f_n \rangle|^2 \leq B\|f\|_2^2, \quad (5.2)$$

for all $f \in L^2(\mathbb{R})$. The constants A and B are called the *lower* and the *upper frame bound*, respectively. These frame constants are intimately related to the representation of a function in the form of an atomic decomposition. The interested reader is referred to appendix A for a detailed discussion on frame theory and atomic decompositions.

The frame condition, as given in (5.2), allows several interpretations. In this thesis, the interpretation of the frame condition related to the energy of functions will be dominant. The term $A\|f\|_2^2$, on the left-hand side of the left inequality, can be interpreted as the energy of a function f , multiplied by the positive constant A . Since this function f is, by assumption, an element of $f \in L^2(\mathbb{R})$, it has a finite energy, and $A\|f\|_2^2$ is therefore a non-negative constant. The same reasoning shows that the term $B\|f\|_2^2$, on the right-hand side of the right inequality, is a non-negative constant. Thus, if the collection $\{f_n\}_{n \in \mathbb{I}}$ forms a frame, then the term $\sum_{n \in \mathbb{I}} |\langle f, f_n \rangle|^2$ should be bounded from below by the non-negative constant $A\|f\|_2^2$ and bounded from above by the non-negative constant $B\|f\|_2^2$. An equivalent expression of the term $\sum_{n \in \mathbb{I}} |\langle f, f_n \rangle|^2$ is $\|\langle f, f_n \rangle\|_2^2$, which could be interpreted as the total energy of the scalars $\langle f, f_n \rangle$. These scalars can next be interpreted as the “similarity” or the “overlap” between the function f and the elementary function f_n . Thus, a collection $\{f_n\}_{n \in \mathbb{I}}$ forms a frame for $L^2(\mathbb{R})$ if the total energy of the sequence of scalars $\langle f, f_n \rangle$ is non-negative and finite for all $f \in L^2(\mathbb{R})$. In Appendix A, it is proved that this condition is a guarantee that atomic decompositions of functions $f \in L^2(\mathbb{R})$ are possible. Such representations are allowed by the properties of several important operators associated with frames.

Operators associated with frames

There are several important operators associated with frames including the analysis operator, the synthesis operator and the frame operator.

If the collection of functions $\{f_n\}_{n \in \mathbb{I}}$ forms a frame, then the *analysis operator* \mathcal{C} , associated with the frame $\{f_n\}_{n \in \mathbb{I}}$, is given by

$$\mathcal{C} : L^2(\mathbb{R}) \rightarrow \ell^2(\mathbb{I}), \quad f \mapsto \{\langle f, f_n \rangle\}_{n \in \mathbb{I}}.$$

The analysis operator maps a function $f \in L^2(\mathbb{R})$ into a collection of scalars $\{\langle f, f_n \rangle\}_{n \in \mathbb{I}} \in \ell^2(\mathbb{I})$. These scalars are called the *frame coefficients* associated with the function f and the frame $\{f_n\}_{n \in \mathbb{I}}$. Just as any inner product between two functions can be interpreted as the “similarity” or the “overlap” between these two functions, so can a frame coefficient be interpreted as the “similarity” or the “overlap” between the analysed function f and a frame element $f_n \in \{f_n\}_{n \in \mathbb{I}}$. The frame operator is therefore a method to extract certain properties of a function. The kind of properties that are extracted from the function through the analysis operator depends trivially on the frame that is associated with this operator.

The adjoint of the analysis operator is the synthesis operator. If the collection $\{f_n\}_{n \in \mathbb{I}}$ forms a frame, then the *synthesis operator* \mathcal{R} , associated with the frame $\{f_n\}_{n \in \mathbb{I}}$, is given by

$$\mathcal{R} : \ell^2(\mathbb{I}) \rightarrow L^2(\mathbb{R}), \quad \{c_n\}_{n \in \mathbb{I}} \mapsto \sum_{n \in \mathbb{I}} c_n f_n.$$

The synthesis operator maps a collection of scalars $\{c_n\}_{n \in \mathbb{I}} \in \ell^2(\mathbb{I})$ into a function $\sum_{n \in \mathbb{I}} c_n f_n \in L^2(\mathbb{R})$. This function is constructed through a discrete superposition of elements of the frame $\{f_n\}_{n \in \mathbb{I}}$, with respect to the scalars $\{c_n\}_{n \in \mathbb{I}}$. These scalars give an indication of the relative weight of the associated frame element f_n in the discrete superposition.

The last operator associated with frames that will be introduced is the so-called *frame operator*. If the collection $\{f_n\}_{n \in \mathbb{I}}$ forms a frame, then the *frame operator* \mathcal{S} , associated with the frame $\{f_n\}_{n \in \mathbb{I}}$, is given by

$$\mathcal{S} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}), \quad f \mapsto \sum_{n \in \mathbb{I}} \langle f, f_n \rangle f_n.$$

The frame operator maps a function $f \in L^2(\mathbb{R})$ into another function $\sum_{n \in \mathbb{I}} \langle f, f_n \rangle f_n \in L^2(\mathbb{R})$. The way that the frame operator constructs the function can clearly be divided in two parts. First, the function f is mapped into a collection of scalars $\{\langle f, f_n \rangle\}_{n \in \mathbb{I}}$, which is exactly the operation performed by the analysis operator. Then, these scalars $\{\langle f, f_n \rangle\}_{n \in \mathbb{I}}$, which are obtained through the analysis operator, are mapped into the function $\sum_{n \in \mathbb{I}} c_n f_n \in L^2(\mathbb{R})$, which is exactly the operation performed by the synthesis operator. Thus, the frame operator can be seen as the concatenation of the analysis and the synthesis operator associated with the $\{f_n\}_{n \in \mathbb{I}}$. In fact, the frame operator could be defined in terms of the analysis and synthesis operator associated with the frame $\{f_n\}_{n \in \mathbb{I}}$, namely as $\mathcal{S}f := \mathcal{RC}f$.

Due to the invertibility of the frame operator \mathcal{S} — see Appendix A —, a frame of the form $\{\mathcal{S}^{-1}f_n\}_{n \in \mathbb{I}}$ could be constructed. This frame is called the *canonical dual frame* of the frame $\{f_n\}_{n \in \mathbb{I}}$. It is this frame that guarantees that an atomic decompositions is possible if the collection of functions $\{f_n\}_{n \in \mathbb{I}}$ forms a frame. By using the canonical dual frame $\{\mathcal{S}^{-1}f_n\}_{n \in \mathbb{I}}$ for the analysis, and using the frame $\{f_n\}_{n \in \mathbb{I}}$ for the synthesis, the analysed function can be represented as

$$f = \sum_{n \in \mathbb{I}} \langle f, \mathcal{S}^{-1}f_n \rangle f_n \tag{5.3}$$

$$= \mathcal{S}\mathcal{S}^{-1}f. \tag{5.4}$$

On the contrary, by using the frame $\{f_n\}_{n \in \mathbb{I}}$ for the analysis, and using its canonical dual frame $\{\mathcal{S}^{-1}f_n\}_{n \in \mathbb{I}}$ for the synthesis, the analysed function f can be represented as

$$f = \sum_{n \in \mathbb{I}} \langle f, f_n \rangle \mathcal{S}^{-1}f_n \tag{5.5}$$

$$= \mathcal{S}^{-1}\mathcal{S}f. \tag{5.6}$$

In general, a representation of a function as a superposition of elements of a frame is called a *frame decomposition*. If a collection of functions forms a frame, then such a frame decomposition is *always* possible. A frame decomposition is always possible due to the fact that the canonical dual frame can be defined. The frame decomposition given in (5.3) represents a function as a discrete superposition of elements of the frame $\{f_n\}_{n \in \mathbb{I}}$, where the coefficients are given by $\{\langle f, \mathcal{S}^{-1}f_n \rangle\}_{n \in \mathbb{I}}$. Similarly, the frame decomposition given in (5.5) represents a function as a discrete superposition of elements of the frame $\{\mathcal{S}^{-1}f_n\}_{n \in \mathbb{I}}$, where the coefficients are given by $\{\langle f, f_n \rangle\}_{n \in \mathbb{I}}$. Note that both frame decompositions are also atomic decompositions since they have the form as expressed in (5.1). Thus, if a collection of functions forms a frame, then an atomic decomposition is *always* possible due to the fact that a canonical dual frame can be defined. The theory associated with frames offers therefore a framework in which it is guaranteed that atomic decompositions are possible. This guarantee is the main motivation for the use of frames in this thesis.

The canonical dual frame, which guarantees a decomposition of a function, is just one concrete example of a frame that leads to a decomposition of a function. In fact, there are sometimes also other frames associated with $\{f_n\}_{n \in \mathbb{I}}$ than $\{\mathcal{S}^{-1}f_n\}_{n \in \mathbb{I}}$ which guarantee such a decomposition. Every so-called *dual frames* guarantee such a decomposition. In general, any two frames $\{f_n\}_{n \in \mathbb{I}}$, $\{h_n\}_{n \in \mathbb{I}}$ are called *dual frames* if they that satisfy

$$f = \sum_{n \in \mathbb{I}} \langle f, h_n \rangle f_n \tag{5.7}$$

So, if two frames form a dual frame of each other, then an atomic decomposition is possible. Questions as whether there exist other dual frames than the canonical dual frame are covered in Appendix A.

5.2 Gabor frames

In chapter 3, it was stated that Gabor theory aims to represent functions as a sum of translated and modulated versions of a window function. The representations that were presented in that chapter were all representations of functions as a continuous superposition of elements of the collection

$$\{\mathcal{E}_\nu \mathcal{T}_\tau g\} = \{g(t - \tau) e^{i2\pi\nu t}\}, \quad \nu, \tau \in \mathbb{R}. \quad (5.8)$$

It can be shown that this collection of elementary functions forms a *continuous frame*. In order to represent functions as a discrete superposition of translated and modulated versions of a window function, the index of the collection of elementary functions given in (5.8) should be discretized. A natural discretization is to set $\tau = n\alpha$ and $\nu = m\beta$, where α and β are two positive constants that range over the integers (Daubechies, 1992). Due to this discretization, the collection of translated and modulated versions of a window function becomes

$$\{\mathcal{E}_{m\beta} \mathcal{T}_{n\alpha} g\} = \{g(t - n\alpha) e^{i2\pi m\beta t}\}, \quad m, n \in \mathbb{Z}. \quad (5.9)$$

An element of this collection is called a *Gabor atom* — see figure 5.1. The collection of Gabor atoms is called a *Gabor system*, and will from now on be denoted by

$$G(g, \alpha, \beta) = \{g_{m,n}\}_{m,n \in \mathbb{Z}}, \quad g_{m,n} := \mathcal{E}_{m\beta} \mathcal{T}_{n\alpha} g. \quad (5.10)$$

If the Gabor system satisfies the frame condition given in (5.2), that is, if it forms a frame for $L^2(\mathbb{R})$, then it will be called a *Gabor frame* for $L^2(\mathbb{R})$. If this is the case, then it is possible to represent any function $f \in L^2(\mathbb{R})$ as a discrete superposition of Gabor atoms. In appendix A, several conditions are given under which a Gabor system forms a frame, but in the main part of this thesis it will simply be assumed that a Gabor system is a Gabor frame since it is then possible to focus entirely on the applications of frames for the analysis and synthesis of signals.

Operator associated with Gabor frames

Since the Gabor system $G(g, \alpha, \beta)$ forms, by assumption, a frame, there are several operators associated with it. Examples of these operators are the analysis operator, the synthesis operator and the frame operator. In this paragraph, these examples of operators will be treated explicitly since this gives a clear overview of the possibilities of these operators.

The analysis operator \mathcal{C}_G , associated with a Gabor frame $G(g, \alpha, \beta)$, is given by

$$\begin{aligned} \mathcal{C}_G f &= \langle f, g_{m,n} \rangle \\ &= \int_{\mathbb{R}} f(t) g_{m,n}^*(t) dt, \end{aligned}$$

where $m, n \in \mathbb{Z}$. The analysis operator associated with a Gabor frame will simply be called the *Gabor transform*. The Gabor transform maps a function f into a collection of scalars $\{\langle f, g_{m,n} \rangle\}_{m,n \in \mathbb{Z}}$. These scalars are in general called *Gabor coefficients* and they can, as any inner product, be interpreted as the “similarity” or the “overlap” between the analysed function and the frame elements, which are in this case Gabor atoms. Thus, the Gabor coefficients give an indication of both the spectral and the temporal characteristics of the analysed function. It does so by representing this function as a discrete set of points on the time-frequency plane. A part of the points of this set are shown in figure 5.3a. Each point in this figure corresponds to the location of a Gabor coefficient on the time-frequency plane. In order to see the relation between the Gabor coefficients and the short-time Fourier transform — both represent a function on the time-frequency plane — it is advantageous to write the Gabor transform in its full main form, which is

$$(\mathcal{C}_G f)(n, m) = \int_{\mathbb{R}} f(t) g(t - n\alpha) e^{-i2\pi m\beta t} dt, \quad (5.11)$$

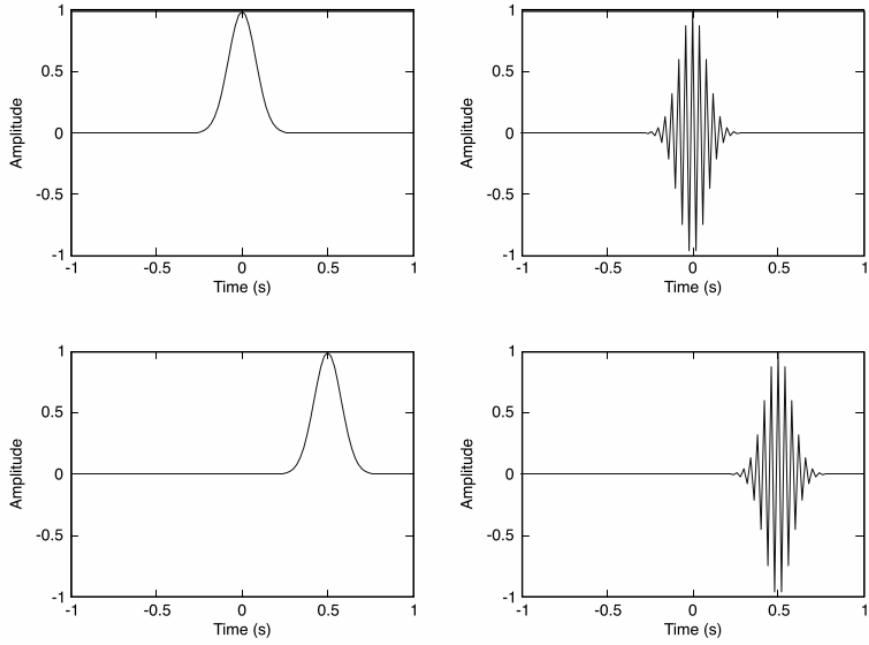


Figure 5.1: Examples of Gabor atoms. The top left subfigure is a Gaussian window function, the top right subfigure a modulated Gaussian window, the bottom left subfigure a translated Gaussian window and the bottom right subfigure a translated, modulated Gaussian window.

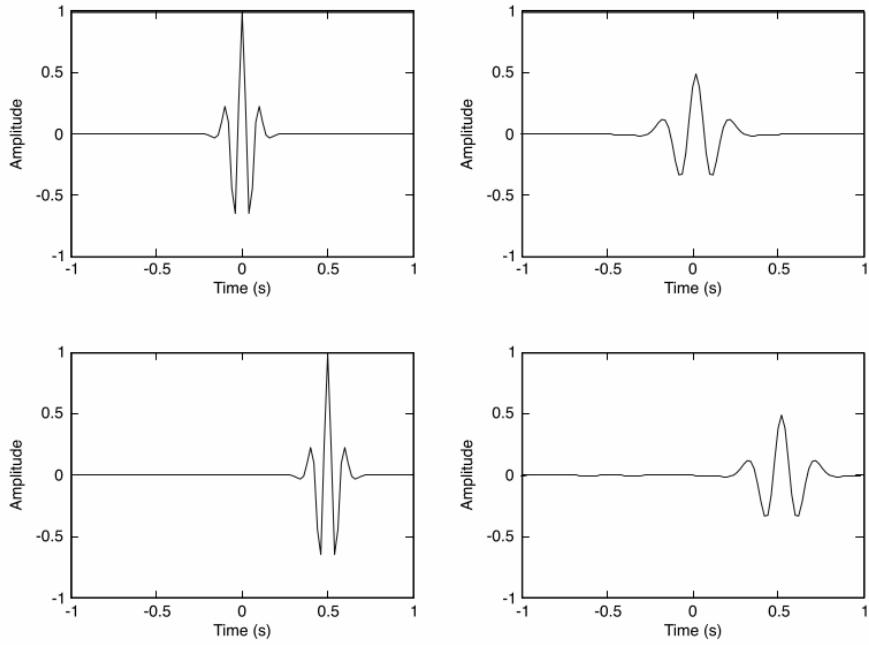


Figure 5.2: Examples of wavelet atoms. The top left subfigure is a Klauder wavelet function, the top right subfigure a dilated Klauder wavelet, the bottom left subfigure a translated Klauder wavelet and the bottom right subfigure a translated, dilated Klauder wavelet.

where $m, n \in \mathbb{Z}$. The relation between the Gabor transform as expressed in (5.11) and the short-time Fourier transform as expressed in (3.1) is now quite obvious. The Gabor transform is namely a subsampled version of the short-time Fourier transform. In fact, the Gabor coefficients $(\mathcal{C}_G f)(n, m)$ correspond to the points $(n\alpha, m\beta)$ of the short-time Fourier transform, that is,

$$(\mathcal{C}_G f)(n, m) = (\mathcal{V}_g f)(n\alpha, m\beta),$$

where $m, n \in \mathbb{Z}$. Note that this is exactly the discretization of the collection of functions given in (5.8), which yielded the collection given in (5.9).

The synthesis operator \mathcal{R}_G , associated with a Gabor frame $G(g, \alpha, \beta)$, is given by

$$\mathcal{R}_G c = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n} g_{m,n}.$$

This synthesis operator maps a collection of scalars $\{c_{m,n}\}_{m,n \in \mathbb{Z}}$ in the time-frequency domain into a function $\mathcal{R}_G c$ in the time domain. In order to get a valuable interpretation of the functions that can be constructed through this synthesis operator, it is advantageous to express the synthesis operator in its full form. Expressing this synthesis operator explicitly in terms of translated and modulation versions of a window function yields

$$(\mathcal{R}_G c)(t) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n} g(t - n\alpha) e^{i2\pi m \beta t}, \quad t \in \mathbb{R}.$$

The synthesis operator expressed in this form leads to the interpretation of the Gabor synthesis operator as a discrete superposition of Gabor atoms. Since these Gabor atoms possess both temporal and spectral characteristics, it is with this synthesis operator possible to synthesize functions whose spectral characteristics evolve over time or whose temporal characteristics evolve over frequency. The scalars from which such functions are synthesized determine the relative weight of the Gabor atom in the superposition. Aside this interpretation, the synthesis operator expressed in its full form also clearly shows the relation between the Gabor synthesis operator and the inverse short-time Fourier transform, as expressed in (3.7). The inverse short-time Fourier transform is namely a continuous superposition of translated and modulated versions of a window function, whereas the Gabor synthesis operator is a discrete superposition of discrete Gabor atoms.

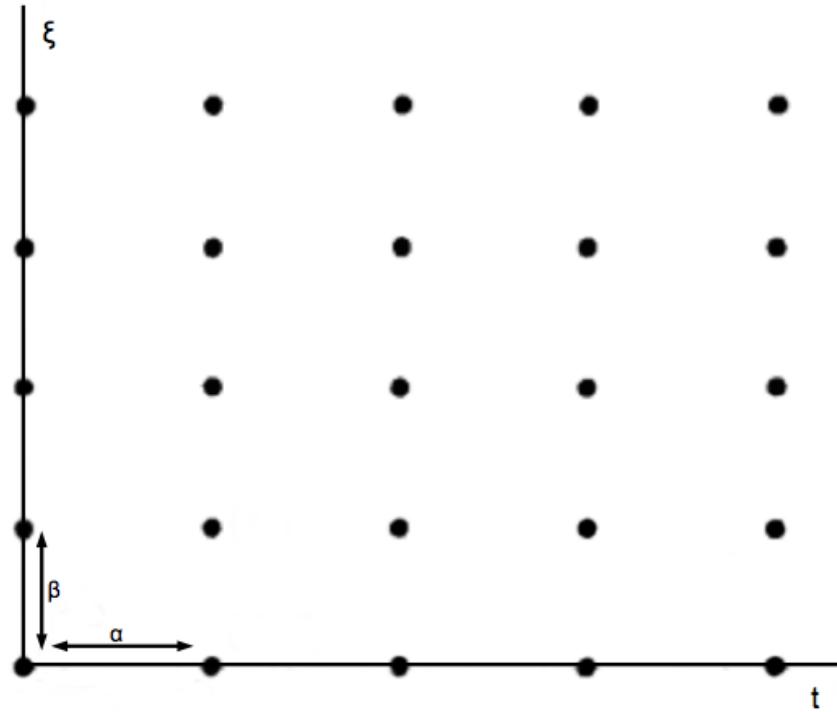
The frame operator \mathcal{S}_G , associated with a Gabor frame $G(g, \alpha, \beta)$, is given by

$$\mathcal{S}_G f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, g_{m,n} \rangle g_{m,n}.$$

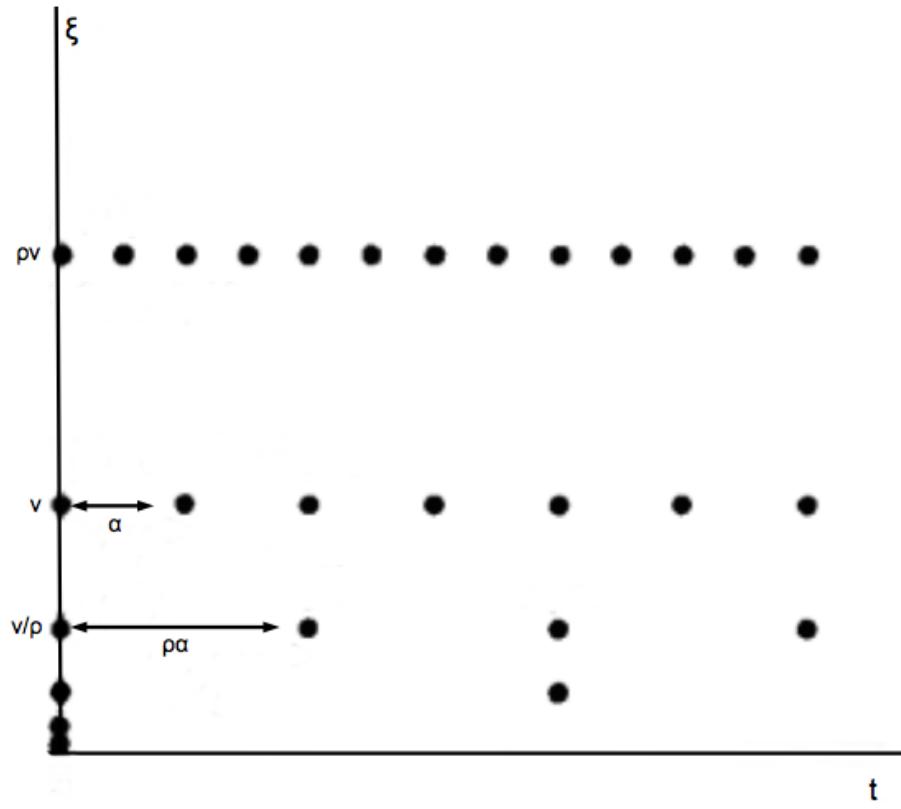
This frame operator maps a function f into another function, namely into the function $\mathcal{S}_G f$. This frame operator can be considered as the concatenation of the Gabor transform and the Gabor synthesis operator. The operation that this operator performs can therefore be divided into two parts. First, the Gabor coefficients $\{\langle f, g_{m,n} \rangle\}_{m,n \in \mathbb{Z}}$ are obtained, which can be obtained by an application of the Gabor transform, associated with the frame $\{g_{m,n}\}_{m,n \in \mathbb{Z}}$, on the function f . Second, the obtained Gabor coefficients $\{\langle f, g_{m,n} \rangle\}_{m,n \in \mathbb{Z}}$ are mapped into the function $\mathcal{S}_G f$, which can be obtained by applying the Gabor synthesis operator, associated with the frame $\{g_{m,n}\}_{m,n \in \mathbb{Z}}$, on the Gabor coefficients. The Gabor frame operator can thus be seen as the concatenation of the analysis and synthesis operator associated with one Gabor frame, namely the Gabor $\{g_{m,n}\}_{m,n \in \mathbb{Z}}$. However, if instead the concatenation of the analysis and synthesis operator associated with two dual Gabor frames is considered, then this leads to the representation

$$f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, h_{m,n} \rangle g_{m,n}, \tag{5.12}$$

where $\{g_{m,n}\}_{m,n \in \mathbb{Z}}$ and $\{h_{m,n}\}_{m,n \in \mathbb{Z}}$ are two dual Gabor frames. The representation of a function as expressed in (5.12) is in general called the *Gabor series* of this function. The Gabor series shows that any function can be represented by a discrete superposition of translated and modulated versions of a window function.



(a) Time-frequency localization of the Gabor coefficients.



(b) Time-frequency localization of wavelet coefficients

Figure 5.3: Examples of localization of frame coefficients on the time-frequency plane. The figures are strongly modified versions of figures that originally occurred in (Daubechies, 1992)

5.3 Wavelet frames

In chapter 4, it was stated that wavelet theory aims to represent a function as a sum of translated and dilated versions of a wavelet function. The representations that were presented in that chapter were all representations of functions as a continuous superposition of elements of the collection

$$\{\mathcal{T}_\tau \mathcal{D}_\rho\} = \left\{ \frac{1}{\sqrt{|\rho|}} h\left(\frac{t-\tau}{\rho}\right) \right\}, \quad \tau, \rho \in \mathbb{R}. \quad (5.13)$$

It can be shown that this collection of elementary functions forms a *continuous frame*. In order to represent functions as a discrete superposition of translated and dilated versions of a wavelet function, the index of the collection of elementary functions in (5.13) should be discretized. A natural discretization is to set $\rho = \rho^j$ and $\tau = n\alpha\rho^j$, where $\rho > 1$ and $\alpha > 0$ are constants and j, n range over the integers (Daubechies, 1992). Due to this discretization, the collection of translated and dilated versions of a wavelet function becomes

$$\{\mathcal{T}_{n\alpha\rho^j} \mathcal{D}_{\rho^j} h\} = \left\{ \frac{1}{\sqrt{|\rho^j|}} h\left(\frac{t-n\alpha\rho^j}{\rho^j}\right) \right\} \quad (5.14)$$

$$= \{\rho^{-j/2} h(\rho^{-j} t - n\alpha)\}, \quad (5.15)$$

where $j, n \in \mathbb{Z}$. An element of this collection is called a *wavelet atom* — see figure 5.2. The collection of wavelet atoms is called a *wavelet system*, and will from now on be denoted by

$$W(h, \rho, \alpha) = \{h_{j,n}\}_{j,n \in \mathbb{Z}}, \quad h_{j,n} := \mathcal{T}_{n\alpha\rho^j} \mathcal{D}_{\rho^j} h. \quad (5.16)$$

If the wavelet system satisfies the frame condition given in (5.2), that is, if it forms a frame for $L^2(\mathbb{R})$, then it will be called a *wavelet frame* for $L^2(\mathbb{R})$. In this case, it is possible to represent any function $f \in L^2(\mathbb{R})$ as a discrete superposition of wavelet atoms. In appendix A, several conditions are given under which a wavelet system forms a frame, but in the rest of the main part of this thesis it will simply be assumed that a wavelet system forms a frame since it is then possible to focus entirely on the applications of such frames for the analysis and synthesis of signals.

Operator associated with wavelet frames

Since the wavelet system $W(h, \rho, \alpha)$ forms, by assumption, a frame, there are several operators associated with it. Examples of these are the analysis operator, the synthesis operator and the frame operator. In this paragraph, these examples of operators will be treated explicitly since this gives a clear overview of the possibilities and characteristics of these operators, which is useful information for the applications.

The analysis operator \mathcal{C}_W , associated with a wavelet frame $W(h, \rho, \alpha)$, is given by

$$\begin{aligned} \mathcal{C}_W f &= \langle f, h_{j,n} \rangle \\ &= \int_{\mathbb{R}} f(t) h_{j,n}^*(t) dt, \end{aligned}$$

where $j, n \in \mathbb{Z}$. The analysis operator associated with a wavelet frame is a discrete version of a wavelet transform. This analysis operator maps a function f into a collection of coefficients $\{\langle f, h_{j,n} \rangle\}_{j,n \in \mathbb{Z}}$. These scalars are in general called *wavelet coefficients* and they can, as any inner product, be interpreted as the “similarity” or the “overlap” between the analysed function and the frame elements, which are in this case wavelet atoms. Since the dilation parameters ρ can be related to a frequency ν through the identification given in (4.3), the wavelet coefficients give an indication of both the spectral and temporal characteristics of the analysed function. It does so by representing the analysed function as a discrete set of points on the time-scale or time-frequency plane. A part of the set of points on the time-frequency plane is shown in figure 5.3b. Each point in this figure corresponds to the location of a wavelet coefficient on the time-frequency plane. In

order to see the relation between these wavelet coefficients and the continuous wavelet transform — both represent a function on the time-scale or time-frequency plane — it is advantageous to write the analysis operator associated with a wavelet frame in its full form. Expressing this analysis operator explicitly in terms of translated and dilated versions of a wavelet function yields

$$(\mathcal{C}_W f)(n, j) = \rho^{-j/2} \int_{\mathbb{R}} f(t) h(\rho^{-j}t - n\alpha) dt, \quad j, n \in \mathbb{Z}. \quad (5.17)$$

The relation between this analysis operator and the continuous wavelet transform as expressed in (4.1) is now quite obvious. The analysis operator, associated with a wavelet frame, is namely a sub-sampled version of the continuous wavelet transform. In fact, the wavelet coefficients $(\mathcal{C}_W f)(n, j)$ correspond to the points $(n\alpha\rho^j, \rho^j)$ of the continuous wavelet transform, that is,

$$(\mathcal{C}_G f)(n, j) = (\mathcal{W}_h f)(n\alpha\rho^j, \rho^j), \quad j, n \in \mathbb{Z}.$$

Note that these discrete points correspond exactly to the discretization of the collection of functions given in (5.13), which yielded the wavelet system given in (5.14).

The synthesis operator \mathcal{R}_W , associated with a wavelet frame $W(h, \rho, \alpha)$, is given by

$$\mathcal{R}_W c = \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{j,n} h_{j,n}$$

This synthesis operator maps a collection of coefficients $\{c_{j,n}\}_{j,n \in \mathbb{Z}}$ in the time-scale or time-frequency plane into a function $\mathcal{R}_W c$ in the time domain. In order to get a valuable interpretation of the functions that can be constructed through this synthesis operator, it is advantageous to express the synthesis operator in its full form. Expressing this synthesis operator explicitly in terms of translated and dilated versions of a wavelet function yields

$$(\mathcal{R}_W c)(t) = \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{j,n} \rho^{-j/2} h(\rho^{-j}t - n\alpha).$$

The synthesis operator expressed in this form leads to the interpretation of the wavelet synthesis operator as a discrete superposition of wavelet atoms. Since these wavelet atoms possess both temporal and spectral characteristics, it is with this synthesis operator possible to synthesize functions whose spectral characteristics evolve over time or whose temporal characteristics evolve over frequency. The scalars from which such functions are synthesized determine the relative weight of the wavelet atom in the superposition. Aside this interpretation, the synthesis operator expressed in its full form also clearly shows the relation between this synthesis operator and the inverse continuous wavelet transform, as expressed in (4.5). The inverse continuous wavelet transform is namely a continuous superposition of translated and dilated versions of a wavelet function, whereas the synthesis operator is a discrete superposition of discrete wavelet atoms.

The frame operator \mathcal{S}_W , associated with a wavelet frame $W(h, \rho, \alpha)$, is given by

$$\mathcal{S}_W f = \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, h_{j,n} \rangle h_{j,n}.$$

This frame operator maps a function f into another function, namely $\mathcal{S}_W f$. This frame operator can be considered as the concatenation of the wavelet analysis and synthesis operator. The operation that this operator performs can therefore be divided into two parts. First, the wavelet coefficients $\{\langle f, h_{j,n} \rangle\}_{j,n \in \mathbb{Z}}$ are obtained, which can be obtained by an application of the analysis operator, associated with the frame $\{h_{j,n}\}_{j,n \in \mathbb{Z}}$, on the function f . Second, the obtained wavelet coefficients $\{\langle f, h_{j,n} \rangle\}_{j,n \in \mathbb{Z}}$ are mapped into the function $\mathcal{S}_W f$, which can be obtained by applying the wavelet synthesis operator, associated with the frame $\{h_{j,n}\}_{j,n \in \mathbb{Z}}$, on the wavelet coefficients. The wavelet frame operator can thus be seen as the concatenation of the analysis and synthesis operator associated with one wavelet frame, namely the wavelet frame $\{h_{j,n}\}_{j,n \in \mathbb{Z}}$. However, if

instead the concatenation of the analysis and synthesis operator associated with two dual wavelet frames is considered, then this leads to the representation

$$f = \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, g_{j,n} \rangle h_{j,n}, \quad (5.18)$$

where $\{h_{j,n}\}_{j,n \in \mathbb{Z}}$ and $\{g_{j,n}\}_{j,n \in \mathbb{Z}}$ are two dual wavelet frames. The representation of a function as expressed in (5.18) is in general called the *wavelet series* of this function. The wavelet series shows that any function can be represented by a discrete superposition of translated and dilated versions of a wavelet function.

5.4 Non-stationary Gabor frames

Both Gabor and wavelet theory aim to represent a function in terms of elementary functions. In the case of Gabor theory these elementary functions are translated and modulated versions of a window function, and in the case of wavelet theory these elementary functions are translated and dilated versions of a wavelet function. In both cases, the elementary functions are derived from one single, fixed generator or atom by elementary operations. The theory associated with non-stationary Gabor theory extends this principle by considering a collection of several generators rather than one single generator. Thus, the elementary functions associated with non-stationary Gabor theory are derived from a collection of generators rather than a single generator. The operation that is applied on these generators, in order to obtain the elementary functions, is modulation. Since the generators can be functions in the time domain or functions in the frequency domain, there are two types of collections of elementary functions associated with non-stationary Gabor theory, namely a collection of elementary functions associated with non-stationary Gabor theory defined in the time domain and the frequency domain. In order to distinguish the non-stationary Gabor theory in the time domain from the non-stationary Gabor theory in the frequency domain, the first type is in general called time-side non-stationary Gabor theory and the second type frequency-side non-stationary Gabor theory.

In this section both the time-side non-stationary Gabor theory and the frequency-side non-stationary Gabor theory will be introduced. Aside the introduction of these theories, it will also be shown that the Gabor and wavelet frames that were introduced in the previous two sections are examples of non-stationary Gabor frames. There are even more well-known signal transforms that can be treated as non-stationary Gabor frames. Examples of these are the constant-Q transform and general filter banks (Balazs et al., 2011). That several well-known signal transforms can be treated as non-stationary Gabor frames is an interesting fact since this is also a guarantee that the transform is invertible. This is, for example, not the case for the constant-Q transform, as introduced by Brown & Puckette (Brown, 1991; Brown & Puckette, 1992). Thus, when these transforms are treated as special cases of non-stationary Gabor theory, they can also be used for the synthesis of signals.

5.4.1 Non-stationary Gabor frames in time domain

The collection of elementary functions associated with non-stationary Gabor theory in the time domain is called a *time-side non-stationary Gabor system*. The definition of such a non-stationary Gabor system is explicitly given by

$$N(g_n, \beta_n)_{n \in \mathbb{Z}} = \{g_{m,n}\}_{m,n \in \mathbb{Z}}, \quad g_{m,n} := \mathcal{E}_{m\beta_n} g_n, \quad (5.19)$$

where $\{g_n\}_{n \in \mathbb{Z}} \subseteq L^2(\mathbb{R})$ is a collection of window functions, and where $\{\beta_n\}_{n \in \mathbb{Z}}$ is a collection of modulation parameters with $\beta_n > 0$. Note that both the window functions and the modulation parameters are indexed by $n \in \mathbb{Z}$. For this reason, there is with each window function g_n one associated modulation parameter β_n . The window functions are assumed to be well-localized in the time domain and centred around positions in time α_n . An element of a non-stationary Gabor

system in the time domain will be called a *time-side non-stationary Gabor atom*, and is explicitly given by

$$g_{m,n}(t) = g_n(t)e^{i2\pi m\beta_n t}.$$

Such a non-stationary Gabor atom is thus a window function $g_n \in L^2(\mathbb{R})$ modulated by $m\beta_n$. As was stated earlier, for each window function g_n there is one associated modulation parameter β_n . Due to this, it might be possible that each window function is modulated by a different frequency. Note that in a Gabor system there is just one window function g and one modulation parameter β . A non-stationary Gabor system $N(g_n, b_n)$ with $g_n = \mathcal{T}_{n\alpha}g$ and $\beta_n = \beta$ for all $n \in \mathbb{Z}$ corresponds therefore to a Gabor system $G(g, \alpha, \beta)$ with $g \in L^2(\mathbb{R})$ and positive constants α, β . Since there is only one window function in a Gabor system, the time-frequency resolution is constant for all positions on the time-frequency plane — see figure 5.3a. However, there are several window functions in a non-stationary Gabor system and it is therefore possible that the time-frequency resolution differs per point on the time-frequency plane. As was stated earlier, each window function g_n of a time-side non-stationary Gabor system is assumed to be centred at position in time α_n . This position in time can be different for each window function. The time-frequency resolution of a time-side non-stationary Gabor system might therefore evolve over time. In figure 5.4a, an example of the time-frequency localization of non-stationary Gabor coefficients is shown. Here it is clearly shown that the time-frequency resolution can differ for each position in time.

Just as any collection of elementary functions, a time-side non-stationary Gabor system forms a frame if it satisfies the frame condition. If a time-side non-stationary Gabor system satisfies the frame condition given in (5.2), then it is called a *time-side non-stationary Gabor frame* for $L^2(\mathbb{R})$. In this case, all functions $f \in L^2(\mathbb{R})$ can be represented as a discrete superposition of time-side non-stationary Gabor atoms. In the rest of this section, it is assumed that the non-stationary Gabor system $N(g_n, \beta_n)_{n \in \mathbb{Z}}$ forms a frame for $L^2(\mathbb{R})$. Now, since it forms a frame, there are several operator associated with it.

Operators associated with non-stationary Gabor frames in the time domain

The analysis operator \mathcal{C}_N , associated with the non-stationary Gabor frame $N(g_n, \beta_n)_{n \in \mathbb{Z}}$, is given by

$$\begin{aligned}\mathcal{C}_N f &= \langle f, g_{m,n} \rangle \\ &= \int_{\mathbb{R}} f(t) g_{m,n}^*(t) dt,\end{aligned}$$

where $m, n \in \mathbb{Z}$. The analysis operator associated with a non-stationary Gabor frame maps a function f into a collection of scalars $\{\langle f, g_{m,n} \rangle\}_{m,n \in \mathbb{Z}}$. These scalars are in general called *time-side non-stationary Gabor coefficients* and they can be interpreted as the “similarity” or the “overlap” between the analysed function and the frame elements, which are in this case the time-side non-stationary Gabor atoms. See figure 5.4a for an example of where these coefficients might be located on the time-frequency plane.

The synthesis operator \mathcal{R}_N , associated with a non-stationary Gabor frame $N(g_n, \beta_n)_{n \in \mathbb{Z}}$, is given by

$$\mathcal{R}_N c = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n} g_{m,n}.$$

This synthesis operator maps a collection of scalars $\{c_{m,n}\}_{m,n \in \mathbb{Z}}$ into a function $\mathcal{R}_N c$. In order to get a valuable interpretation of the synthesis operator it is advantageous to express it in its full form, that is, expressing it explicitly in terms of the modulated versions of several window functions. Expressing the synthesis operator explicitly in terms of modulated versions of several window functions yields

$$(\mathcal{R}_N c)(t) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n} g_n(t) e^{i2\pi m\beta_n t}.$$

The synthesis operator expressed in this form lead to the interpretation of the non-stationary Gabor synthesis operator as a discrete superposition of time-side non-stationary Gabor atoms. The window functions associated with these atoms are, by assumption, well-localized in time and are also well-localized in frequency due to the modulation. Thus, these non-stationary Gabor atoms are well-localized in both time and frequency. Due to this, the synthesis operator associated with these atoms is a method that allows the synthesis of functions whose spectral characteristics evolve over time or whose temporal characteristics evolve over frequency. Furthermore, the time-frequency resolution that is connected to these functions can also evolve over time.

The frame operator \mathcal{S}_N , associated with the non-stationary Gabor frame $N(g_n, \beta_n)_{n \in \mathbb{Z}}$, is given by

$$\mathcal{S}_N f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, g_{m,n} \rangle g_{m,n}.$$

This frame operator maps a function f into another function, namely $\mathcal{S}_N f$. This frame operator can be considered as the concatenation of the non-stationary Gabor analysis and synthesis operator. The operation that this frame operator performs can therefore be divided into two parts. First, the non-stationary coefficients $\{\langle f, g_{m,n} \rangle\}_{m,n \in \mathbb{Z}}$ are obtained, which can be obtained by applying the non-stationary Gabor analysis operator, associated with the frame $\{g_{m,n}\}_{m,n \in \mathbb{Z}}$, on the function f . Second, the obtained non-stationary Gabor coefficients are mapped into the function $\mathcal{S}_N f$, which can be obtained by an application of the non-stationary Gabor synthesis operator, associated with the frame $\{g_{m,n}\}_{m,n \in \mathbb{Z}}$, on the obtained non-stationary Gabor coefficients. The non-stationary Gabor frame operator can thus be seen as the concatenation of the analysis and synthesis operator associated with one time-side non-stationary Gabor frame, namely $\{g_{m,n}\}_{m,n \in \mathbb{Z}}$. However, if instead the concatenation of the analysis and synthesis operator associated with two dual non-stationary Gabor frames is considered, then this leads to the representation.

$$f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, g_{m,n} \rangle h_{m,n}, \quad (5.20)$$

where $\{h_{m,n}\}_{m,n \in \mathbb{Z}}$ and $\{g_{m,n}\}_{m,n \in \mathbb{Z}}$ are two dual non-stationary Gabor frames. The representation of a function as in (5.20) is called a *time-side non-stationary Gabor series* of this function. This time-side non-stationary Gabor series shows that any function $f \in L^2(\mathbb{R})$ can be represented by a discrete superposition of non-stationary Gabor atoms.

5.4.2 Non-stationary Gabor frames in frequency domain

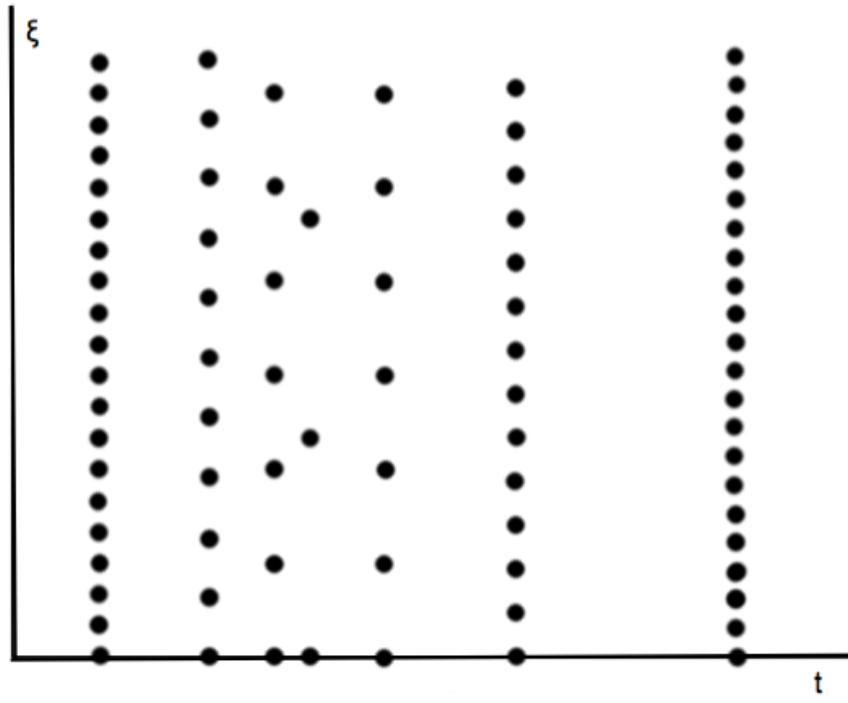
As was stated in the introduction of this section, non-stationary Gabor systems possess aside a definition in the time domain also a definition in the frequency domain. In both cases the non-stationary Gabor systems can be defined through the application of a modulation operator on several window functions. In the case of a non-stationary Gabor system in the frequency domain, these window functions are functions in the frequency domain. Therefore, if $\{\hat{h}_m\}_{m \in \mathbb{Z}} \subseteq L^2(\mathbb{R})$ is a collection of window functions h_m in the frequency domain, then a non-stationary Gabor system in the frequency domain is given by

$$N(\hat{h}_m, \alpha_m)_{m \in \mathbb{Z}} = \{\hat{h}_{m,n}\}_{m,n \in \mathbb{Z}}, \quad \hat{h}_{m,n} := \mathcal{E}_{-n\alpha_m} \hat{h}_m,$$

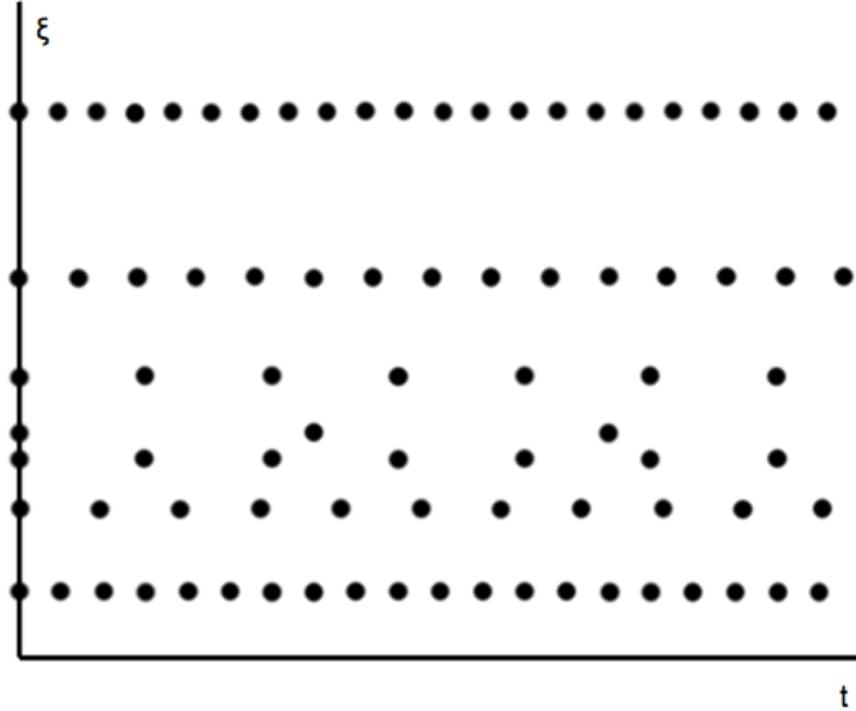
where $\{\alpha_m\}_{m \in \mathbb{Z}}$ is a collection of modulation parameters with $\alpha_m > 0$. Such a non-stationary Gabor system will be called a *frequency-side non-stationary Gabor system*. The window functions are assumed to be well-localized in the frequency domain and centred around positions in frequency β_n . An element of a frequency-side non-stationary Gabor system is explicitly given by

$$\hat{h}_{m,n}(\xi) = \hat{h}_m(\xi) e^{-i2\pi n \alpha_m \xi},$$

and is called a *frequency-side non-stationary Gabor atom*. Such a non-stationary Gabor atom is thus a window function $\hat{h}_m \in L^2(\mathbb{R})$ modulated by $-n\alpha_m$. For each window function g_m there is



(a) Time-frequency localization of non-stationary Gabor coefficients in the time domain.



(b) Time-frequency localization of non-stationary Gabor coefficients in the frequency domain.

Figure 5.4: Examples of localization of non-stationary Gabor frame coefficients on the time-frequency plane. The figures are based on examples that originally occurred in (Balazs et al., 2011).

one associated modulation parameter α_m . Rewriting the non-stationary Gabor system in terms of $h_{m,n}$, that is, in terms of the inverse Fourier transform of the frequency-side non-stationary Gabor atoms yields

$$N(h_m, \alpha_m)_{m \in \mathbb{Z}} = \{h_{m,n}\}_{m,n \in \mathbb{Z}}, \quad h_{m,n} := \mathcal{T}_{n\alpha_m} h_m,$$

where $h_{m,n}$ is explicitly given by

$$h_{m,n}(t) = h_m(t - n\alpha_m).$$

The frequency-side non-stationary Gabor system expressed in the time domain consists thus of by $n\alpha_m$ translated versions of window functions h_m . Again, there is for each window function h_m one translation parameter α_m . Note that in a wavelet system there is just one generator h and one translation parameter α . A non-stationary Gabor system $N(h_m, \alpha_m)_{m \in \mathbb{Z}}$ with $h_m = \mathcal{D}_m$ and $\alpha_m = \alpha$ for all $m \in \mathbb{Z}$ corresponds therefore to a wavelet system $W(h, \alpha, m)$ with $h \in L^2(\mathbb{R})$ and positive constants α, m . The time spread of the wavelet function is changed by the dilation operator, but at each position in time α there is a version of wavelet function centred. In the case of a non-stationary Gabor system, there is one window function at each time position α_m . The time spread of this window function can be chosen freely since at each time position one window function h_m may be chosen. This leads to the possibility of evolving the time-frequency resolution over frequency. In figure 5.4b, an example of the time-frequency localization of non-stationary Gabor coefficients is shown. Here it is clearly shown that the time-frequency resolution can differ for each different frequency.

Since all the other systems of functions in this thesis were also defined in the time domain, the frequency-side non-stationary Gabor system will be denoted by the collection of functions in the time domain $\{h_{m,n}\}_{m,n \in \mathbb{Z}}$. However, to distinguish it from the time-side non-stationary Gabor system $N(g_n, \beta_n)_{n \in \mathbb{Z}}$, it will be denoted by $N(\hat{h}_n, \alpha_n)_{n \in \mathbb{Z}}$. Just as any other system of functions, a frequency-side non-stationary Gabor system forms a frame if it satisfies the frame condition. If the frequency-side non-stationary Gabor system satisfies the frame condition given in (5.2), then it is called a *frequency-side non-stationary Gabor frame* for $L^2(\mathbb{R})$. In the rest of this section it is assumed that the frequency-side non-stationary Gabor system is a frame for $L^2(\mathbb{R})$. Since $N(\hat{h}_n, \alpha_n)_{n \in \mathbb{Z}}$ forms a frame, there are several operator associated with it.

Operators associated with non-stationary Gabor frames in the frequency domain

The analysis operator \mathcal{C}_N , associated with the non-stationary Gabor frame $N(\hat{h}_n, \alpha_n)_{n \in \mathbb{Z}}$, is given by

$$\begin{aligned} \mathcal{C}_N f &= \langle f, h_{m,n} \rangle \\ &= \int_{\mathbb{R}} f(t) h_{m,n}^*(t) dt, \end{aligned}$$

where $m, n \in \mathbb{Z}$. The analysis operator associated with a non-stationary Gabor frame maps a function f into a collection of scalars $\{\langle f, h_{m,n} \rangle\}_{m,n \in \mathbb{Z}}$. These scalars are in general called *frequency-side non-stationary Gabor coefficients* and they can be interpreted as the “similarity” or the “overlap” between the analysed function and the frame elements, which are in this case the frequency-side non-stationary Gabor atoms.

The synthesis operator \mathcal{R}_N , associated with a non-stationary Gabor frame $N(\hat{h}_n, \alpha_n)_{n \in \mathbb{Z}}$, is given by

$$\mathcal{R}_N c = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n} h_{m,n}.$$

This synthesis operator maps a collection of scalars $\{c_{m,n}\}_{m,n \in \mathbb{Z}}$ into a function $\mathcal{R}_N c$. In order to get a valuable interpretation of the synthesis operator it is advantageous to express it in its full

form, that is, expressing it explicitly in terms of the translated versions of several window functions. Expressing this synthesis operator explicitly in terms of translated versions of a window function yields

$$(\mathcal{R}_N c)(t) = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n} h_n(t - n\alpha_m)$$

The synthesis operator expressed in this form lead to the interpretation of the non-stationary Gabor synthesis operator as a discrete superposition of frequency-side non-stationary Gabor atoms. The window functions associated with these atoms are, by assumption, well-localized in frequency and are also well-localized in time due to the translation. Thus, these non-stationary Gabor atoms are well-localized in both time and frequency. Due to this, the synthesis operator associated with these atoms is a method that allows the synthesis of functions whose spectral characteristics evolve over time or whose temporal characteristics evolve over frequency. Furthermore, the time-frequency resolution that is connected to these functions can also evolve over frequency.

The frame operator \mathcal{S}_N , associated with the non-stationary Gabor frame $N(\hat{h}_n, \alpha_n)_{n \in \mathbb{Z}}$, is given by

$$\mathcal{S}_N f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, h_{m,n} \rangle h_{m,n}.$$

This frame operator maps a function f into another function, namely the function $\mathcal{S}_N f$. This frame operator can be considered as the concatenation of the non-stationary Gabor analysis and synthesis operator. The operation that this frame operator performs can therefore be divided into two parts. First, the non-stationary coefficients $\{\langle f, h_{m,n} \rangle\}_{m,n \in \mathbb{Z}}$ are obtained, which can be obtained by applying the non-stationary Gabor analysis operator, associated with the frame $\{h_{m,n}\}_{m,n \in \mathbb{Z}}$, on the function f . Second, the obtained coefficients are mapped into the function $\mathcal{S}_N f$, which can be obtained by an application of the non-stationary Gabor synthesis operator, associated with the frame $\{h_{m,n}\}_{m,n \in \mathbb{Z}}$, on the obtained wavelet coefficients. The non-stationary Gabor frame operator can thus be seen as the concatenation of the analysis and synthesis operator associated with one non-stationary Gabor frame, namely $\{h_{m,n}\}_{m,n \in \mathbb{Z}}$. However, if instead the concatenation of the analysis and synthesis operator associated with two dual non-stationary Gabor frames is considered, then this leads to the representation

$$f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, g_{m,n} \rangle h_{m,n}, \quad (5.21)$$

where $\{h_{m,n}\}_{m,n \in \mathbb{Z}}$ and $\{g_{m,n}\}_{m,n \in \mathbb{Z}}$ are two dual non-stationary Gabor frames. The representation of a function as in (5.21) is called a *frequency-side non-stationary Gabor series* of this function. This frequency-side non-stationary Gabor series shows that any function $f \in L^2(\mathbb{R})$ can be represented by a discrete superposition of non-stationary Gabor atoms.

Part II

Analysis, Synthesis and Analysis-Resynthesis

Chapter 6

Analysis

The analysis of a signal is the extraction of information from this signal. The analysis of a signal associated with an atomic decomposition is the extraction of information from this signal by representing this signal as a collection of discrete-indexed scalars. In order to obtain this collection of scalars, the signal should be mapped into these scalars. For this reason can the analysis associated with an atomic decomposition be seen as a mapping that maps a signal into a collection of scalars. The general form of such a mapping is

$$f \longmapsto \langle f, h_k \rangle,$$

where f is the signal to be analysed and where $\{h_k\}_{k \in \mathbb{I}}$ is a collection of so-called *analysing functions*, which are the functions used for the analysis. Thus, the signal is mapped from a function in the time domain to a collection of scalars in a so-called *transform domain*. The scalars that are obtained through the analysis mapping are elements of the transform domain, and for this reason they will be called *transform coefficients*. Since these transform coefficients are inner products between the analysed signal and a collection of analysing functions, they can be interpreted as the “similarity” or “overlap” between the analysed signal and these analysing functions. For this reason, the analysing functions should possess the same kind of information as should be extracted from the signal to be analysed. For example, if the spectral characteristics of a signal is the desired kind of information to extract from the signal, then the analysing functions should possess certain spectral characteristics. Similarly, if both the temporal and the spectral characteristics of a signal to be analysed should be extracted, then the analysing functions should possess both temporal and spectral characteristics.

The transform domain, the domain of which the transform coefficients are elements, is inseparable related to the analysing functions since they are the functions which, together with the analysed signal, determine the transform coefficients. In fact, the analysing functions determine the transform domain completely since they are the functions that determine the kind of information that the transform domain contains. The actual values of the transform coefficients are of course determined by both the analysed signal and the analysing functions. The transform coefficients can therefore be interpreted as a representation of the analysed signal in the transform domain. Thus, the transform coefficients represent just the analysed signal, but in the transform domain instead of the time domain. For this reason, it is natural to suppose that the transform coefficients possess exactly the same information as the analysed signal in the time domain does. Since this is the case with an atomic decomposition, it is possible to map the signal from the transform domain back to the time domain. This adjoint mapping is in general called *synthesis*. The synthesis associated with an atomic decomposition constructs a function as a discrete superposition of so-called *expansion functions*, which are the functions used for the expansion. The synthesis associated with atomic decompositions is primarily discussed in the next chapter, but since both mappings are inseparable related to atomic decompositions, it will also be an important feature in the analysis of signals.

Transform coefficients Since the collection of transform coefficients is, as noted previously, determined by the collection of analysing functions used in the analysis mapping, the collection of coefficients is indexed by elements of the same index set as these analysing functions are indexed. The values of the transform coefficients give, as also noted previously, an indication of the relation between the analysed signal and the analysing functions. These values of the transform coefficients are in general *complex*, that is, the transform coefficients are complex-valued scalars. The values of the transform coefficients can therefore be represented in several forms: they can be presented in their Cartesian form or in their polar form. The values of the transform coefficients are in both forms of course exactly the same; only the representation of the values is different in both forms. For this reason, both forms lead to different interpretations.

In the Cartesian form, the transform coefficients consists of a real part and an imaginary part. For a complex-valued number z in the Cartesian form, $z = a + bi$, the real part a and imaginary part b are defined as

$$a = \Re(z), \quad b = \Im(z),$$

where $a, b \in \mathbb{R}$. The value $\Re(z)$ denotes thus the value of the real part of the complex number z , and the value $\Im(z)$ denotes the value of the imaginary part of z .

In the polar form, the transform coefficients consists of a magnitude and a phase. For a complex number z in polar form, $z = re^{i\varphi}$, the magnitude r and phase φ are defined as the modulus respectively the argument of z , that is,

$$r = |z|, \quad \varphi = \arg z, \tag{6.1}$$

where r is a non-negative real number, and where $\varphi \in \mathbb{R}$ is in general the principal value in the open-closed interval $] -\pi, \pi]$.

In this thesis, the transform coefficients will only be represented in their polar form since this leads to the most valuable interpretations. The interpretation of the magnitude of the transform coefficient is reasonable obvious due to its relation to energy, which is defined as $r^2 = |z|^2$. The magnitude of the transform coefficients in the transform domain leads therefore to a representation of the energy of the analysed signal in the transform domain. However, the interpretation of the phase of the transform coefficients is in general not obvious. The presentation of the phase of the analysed signal in the transform domain is therefore also not obvious. Although the phase values do not lead to an obvious interpretation, the information that these values possess is important. For example, the phase values are of fundamental importance for the perfect reconstruction of a signal from its transform coefficients, and are therefore also necessary for the construction of a signal from modified transform coefficients. A representation of the phase that leads to an obvious interpretation is therefore valuable. Such representations exist and they are, just as in the case of the short-time Fourier transform and the continuous wavelet transform, based on the local centroids associated with these transforms. The local centroids associated with discrete frames are, for the sake to avoid a lot of repetition, not discussed in this thesis. The interested reader is therefore referred to a recent paper that treats these topics (Holighaus et al., 2015).

6.1 Linear analysis

The analysis mapping that maps a signal into a collection of inner products is a *linear* mapping from the time domain to the transform domain. This kind of analysis will therefore be called *linear analysis*. The transform coefficients that are constructed through such a linear analysis can be two types of transform coefficients. The first type of transform coefficients are the inner products between the analysed signal and a predetermined collection of analysing functions. The second type of transform coefficients are the inner products between the analysed signal and a non-predetermined collection of analysing function, but from which the signal can be reconstructed through a predetermined collection of expansion functions. Thus, the first type of transform coefficients constitutes a representation of the analysed signal in the transform domain associated with a collection of predetermined analysing functions. The second type of transform coefficients constitutes a representation of the analysed signal in a transform domain that is not predetermined, but which guarantees that the transform coefficients are *the* transform coefficients which are needed in order to perfectly reconstruct a signal as a discrete superposition of predetermined expansion functions. Note that both types of transform coefficients can lead to a perfect reconstruction of the signal, but that the kind of analysing and expansion functions, to which the obtained transform coefficients correspond, are different.

Since both types of transform coefficients that can be constructed through a linear analysis are different, also the procedures that derive both kind of transform coefficients are different. The first type of transform coefficients can be derived through a direct computation with the predetermined analysing functions. The second type can be derived through an iterative procedure that is associated with the predetermined expansion functions.

6.1.1 Linear analysis with predetermined analysing functions

The transform coefficients associated with a predetermined collection of analysing function can, as stated previously, directly be computed. This type of transform coefficients corresponds to the inner product between the analysed signal and the analysing functions. This kind of linear analysis can therefore be seen as the operator \mathcal{C} given by

$$\mathcal{C} : L^2(\mathbb{R}) \rightarrow \ell^2(\mathbb{I}), \quad \mathcal{C}f = \{\langle f, f_k \rangle\}_{k \in \mathbb{I}} \quad (6.2)$$

where $f \in L^2(\mathbb{R})$ is the signal to be analysed and where $\{f_k\}_{k \in \mathbb{I}} \subseteq L^2(\mathbb{R})$ is a collection of analysing functions. If this collection of analysing functions forms a frame, which is assumed, then the operator in (6.2) is the frame analysis operator associated with this frame. The collection of transform coefficients is thus the collection of frame coefficients associated with the analysing functions. These frame coefficients represent the analysed signal in the transform domain. Applying the frame synthesis operator, associated with a dual frame of the used analysing functions $\{f_k\}_{k \in \mathbb{I}}$, on these frame coefficients, yields the analysed signal back in the time domain, that is,

$$f = \sum_{k \in \mathbb{I}} \langle f, f_k \rangle g_k,$$

where $\{g_k\}_{k \in \mathbb{I}}$ denotes a dual frame of the frame $\{f_k\}_{k \in \mathbb{I}}$. Thus, a signal can be perfectly reconstructed after a linear analysis associated with a predetermined collection of analysing functions.

6.1.2 Linear analysis with predetermined expansion functions

The linear mapping that constructs transform coefficients associated with a predetermined collection of expansion functions is based on an iterative procedure rather than a direct computation — as is the case by the construction of transform coefficients associated with a predetermined collection of analysing functions. This iterative procedure consists of approximating a collection of transform coefficients from which the analysed signal can be constructed through a discrete superposition of the predetermined expansion functions. Thus, this type of linear analysis is inseparable related to a synthesis. The iterative procedure associated with this type of linear analysis

allows the construction of the transform coefficients associated with a predetermined collection of expansion functions if, and only if, the collection of expansion functions form a frame. Since it is assumed that the collection of expansion functions forms a frame, this is applicable to the expansion functions studied in this chapter. There are several iterative algorithms that can perform the iterative procedure including the, what will be called, *frame analysis algorithm* and the *conjugate gradient algorithm*. In this thesis only the so-called frame analysis algorithm will be treated.

Frame analysis algorithm The frame analysis algorithm performs an iterative procedure that allows the construction of a collection of frame coefficients associated with a predetermined collection of expansion functions. This iterative procedure is intimately related to the frame bounds of this collection of expansion functions, and depends only on this collection of expansion functions and the signal to be analysed. So, the frame analysis algorithm constructs a collection of frame coefficients — a collection of inner products between the analysed signal and a collection of analysing functions — without any a priori knowledge on the analysing functions. The frame coefficients are constructed by successive approximations that are completely based on the signal to be analysed and the expansion functions. Applying the frame synthesis operator associated with these expansion functions on the constructed frame coefficients results again in the analysed signal in the time domain.

The main principle behind the frame analysis algorithm is that it approximates a collection of frame coefficients based on the difference between the analysed signal and an estimation of this signal. Such an estimation of the analysed signal is constituted by applying the frame operator, associated with the predetermined collection of expansion functions, on a previous estimation of the analysed signal, which is obtained in a previous iteration.

The actual frame analysis algorithm is presented as pseudocode in algorithm 1. Here, \mathcal{S} and \mathcal{R} denote the frame operator respectively frame synthesis operator. Both operators are the operators associated with the predetermined collection of expansion functions $\{h_k\}_{k \in \mathbb{I}}$.

Algorithm 1 Frame algorithm

```

1:  $r^{(0)} \leftarrow f$ 
2: repeat for  $n = 1, 2, \dots$ 
3:  $r^{(n)} \leftarrow r^{(n-1)} - (\frac{2}{A+B})\mathcal{S}r^{(n-1)}$ 
4:  $c_k^{(n)} \leftarrow \frac{2}{A+B}\langle r^{(0)} + r^{(1)} + \dots + r^{(n-1)}, h_k \rangle$ 
5:  $f^{(n)} \leftarrow \mathcal{R}c_k^{(n)}, \forall k \in \mathbb{I}$ 
6: until  $\|f - f^{(n)}\|_2 = 0$ 
```

In words, the estimation of the analysed signal f after the zeroth iteration, $r^{(0)}$, is this signal itself, that is, $r^{(0)} = f$. The estimation of the analysed signal after the first iteration, $r^{(1)}$, is then constructed by applying the frame operator on the estimation of the analysed signal after the zeroth iteration, $r^{(0)}$, multiplying the obtained function by $2/(A+B)$, and subtracting it from $r^{(0)}$. By iterating this process, a sequence of estimations of the analysed signal can be constructed. Next, by taking the inner product of the sum of these estimations up to the n 'th iteration and an expansion function, and multiplying this inner product by the constant $2/(A+B)$, results in the frame coefficient of the associated expansion function after the n 'th iteration. By applying the frame synthesis operator on these constructed frame coefficients, results in an approximation of the signal to be analysed after n 'th iteration, $f^{(n)}$. By iterating this iterative procedure until a certain stopping criterion is met, the difference between the analysed signal and its estimation decreases with each iteration. It can be shown that the difference between the analysed signal and its estimation after the n 'th iteration is less than or equal to the constant $(\frac{B-A}{B+A})^n \|x\|_2$, where n denotes the number of iterations (Duffin & Schaeffer, 1952). Note that in algorithm 1, the stopping criterion is a difference of zero between the analysed signal and its estimation. If this stopping criterion is met, then the constructed frame coefficients are the ones that correspond to the predetermined collection of expansion functions.

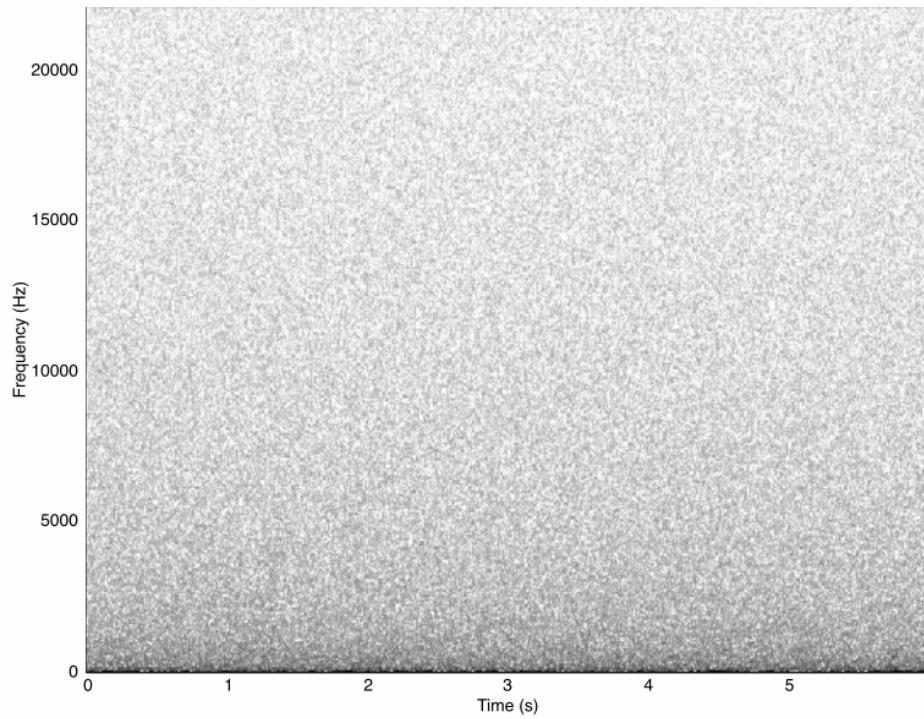
6.2 Examples

In this section several representations of analysed signals will be given. These representations are all obtained through the application of an analysis mapping on a signal in the time domain. The analysed signals all possess properties that characterize sound signals. From these representations it will be deduced how the different analysis methods describe the properties that are characteristic for sound signals.

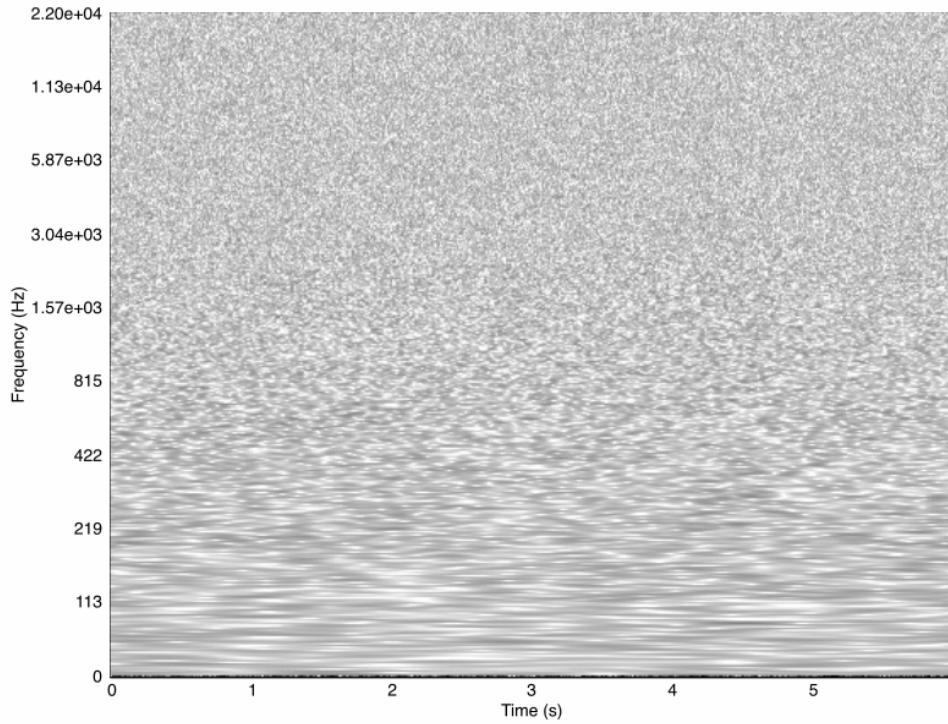
Example 1 In this example, it will be compared how the analysis operator associated with a Gabor frame and the analysis operator associated with a wavelet frame describe a pink noise signal. In figure 6.1 two time-frequency representations of a pink noise signal are given. In this figure, the subfigure 6.1a presents the magnitude of the Gabor coefficients of the pink noise signal, and subfigure 6.1b presents the magnitude of its wavelet coefficients. It can immediately be deduced from these figures that the time-frequency localization of the Gabor and wavelet coefficients are different — see figure 5.4 for the actual time-frequency localization of both types of coefficients. The density of the coefficients indicate that in the case of the Gabor coefficients the time-frequency resolution is equal for all positions on the time-frequency plane, whereas in the case of the wavelet coefficients the time-frequency resolution differs per position. The constant time-frequency resolution in the case of the Gabor coefficients results in a linearly spaced frequency scale, whereas the variable time-frequency resolution in the case of the wavelet coefficients results in a frequency scale that is spaced logarithmically. Note that not every transform with a variable time-frequency resolution results in a frequency scale that is spaced logarithmically, but that this is one of the characteristics of the wavelet transform. This characteristic is of interest for the analysis of acoustical signals since the logarithmic frequency scale corresponds better to the way sound is perceived than a linear frequency scale does.

Example 2 The examples in the second part of this thesis are in general related to discrete frames rather than continuous frames. However, this example will be an exception on this rule since the reassigned versions of transforms associated with discrete frames are, as stated in the introduction of this chapter, not discussed in this thesis. Again, the interested reader is referred to the recent paper (Holighaus et al., 2015).

In this example, the ordinary spectrogram and scalogram will be compared with their reassigned versions. In figure 6.2, the spectrogram and scalogram of a synthesized signal are shown. In figure 6.3, the reassigned versions of the same spectrogram and scalogram are shown. The analysed signal is a signal that is synthesized using methods that are treated in chapter 7. The differences between the spectrogram in figure 6.2a and the scalogram in figure 6.2b clearly show the differences between a Gabor and wavelet analysis. The main difference between both methods that can be deduced from these figures is that the time-frequency resolution of a Gabor analysis is constant for all positions in the time-frequency plane, whereas the time-frequency resolution of a wavelet analysis differs per position. The time-frequency resolution of both methods is approximately equal at the event centred around 8000 Hz at 1 second. The same differences can be deduced through a comparison of figure 6.3a and 6.3b. However, the differences between the spectrogram and scalogram are less obvious when their reassigned versions are compared. This is due to the fact that the reassigned versions of the spectrogram and scalogram reflect the energy of the analysed signal much better than the ordinary versions do. A reassigned version of a transform allows therefore to extract certain properties of an analysed signal that are less accessible for an ordinary version of a transform. For example, the sound event that is centred around 6000 Hz is shown in figure 6.2 as an event that consists of a band of frequencies, but in figure 6.3 it is shown as almost a single frequency component. Detailed information on special characteristics of a signal can thus much better be deduced from a reassigned version of a transform than from its ordinary version.

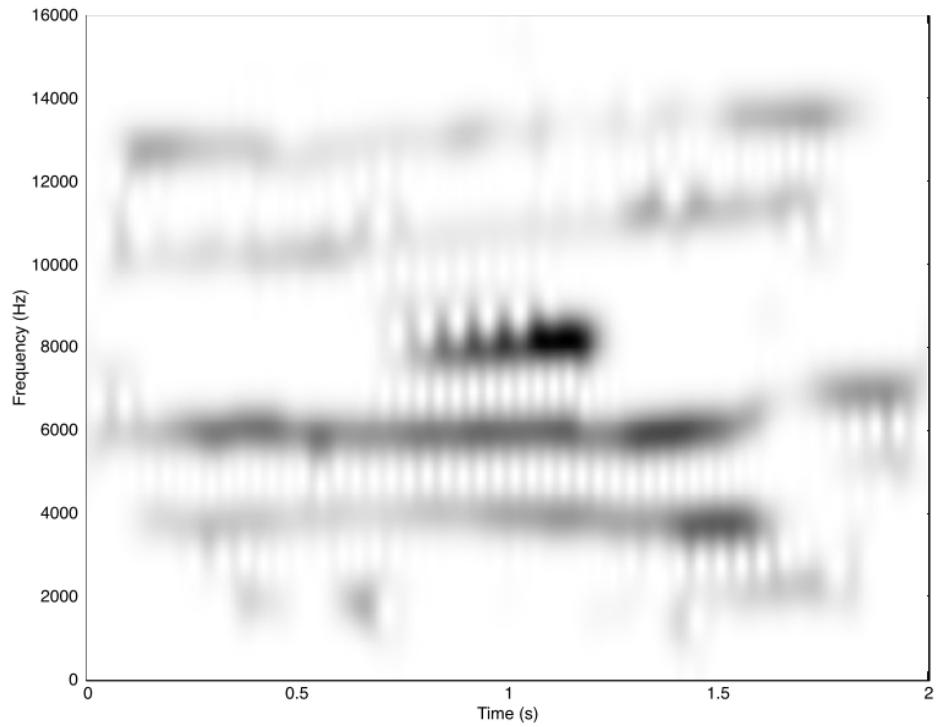


(a) Magnitude of Gabor coefficients

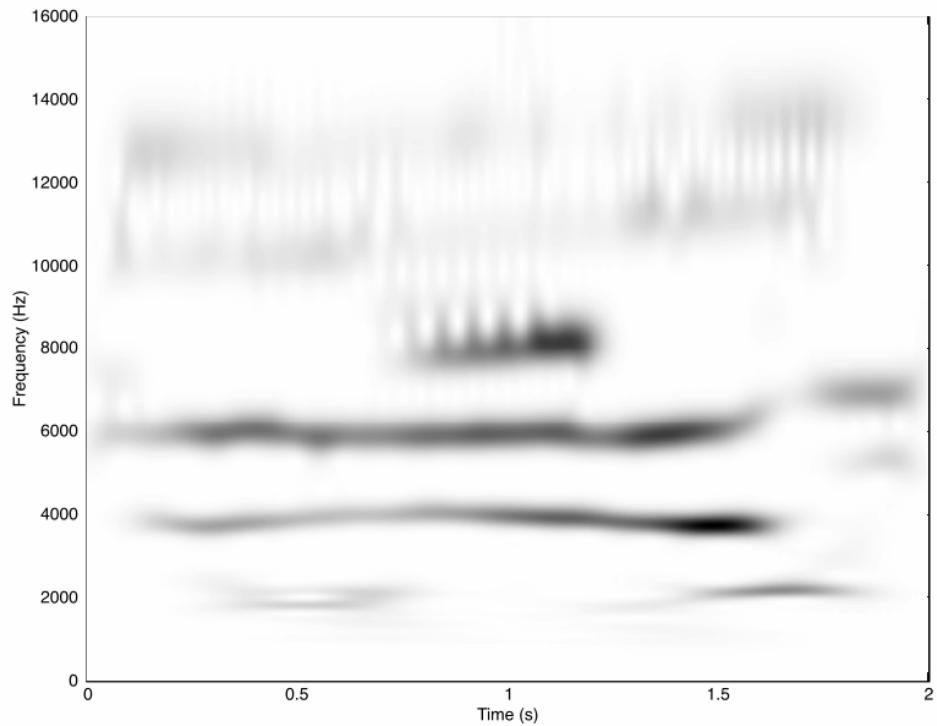


(b) Magnitude of wavelet coefficients

Figure 6.1: Time-frequency representations of a pink noise signal

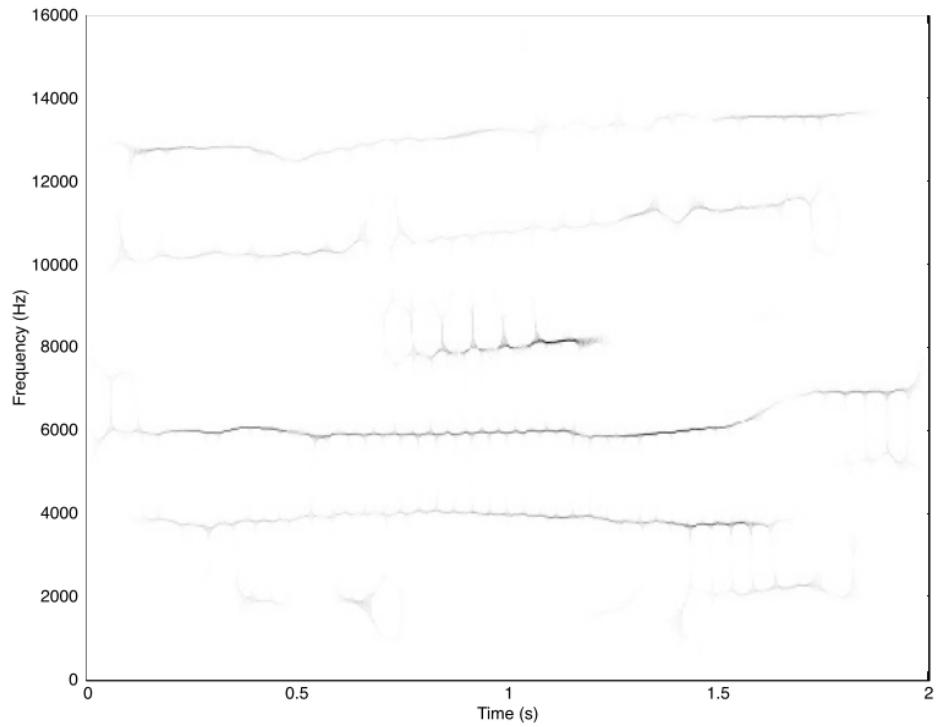


(a) Spectrogram

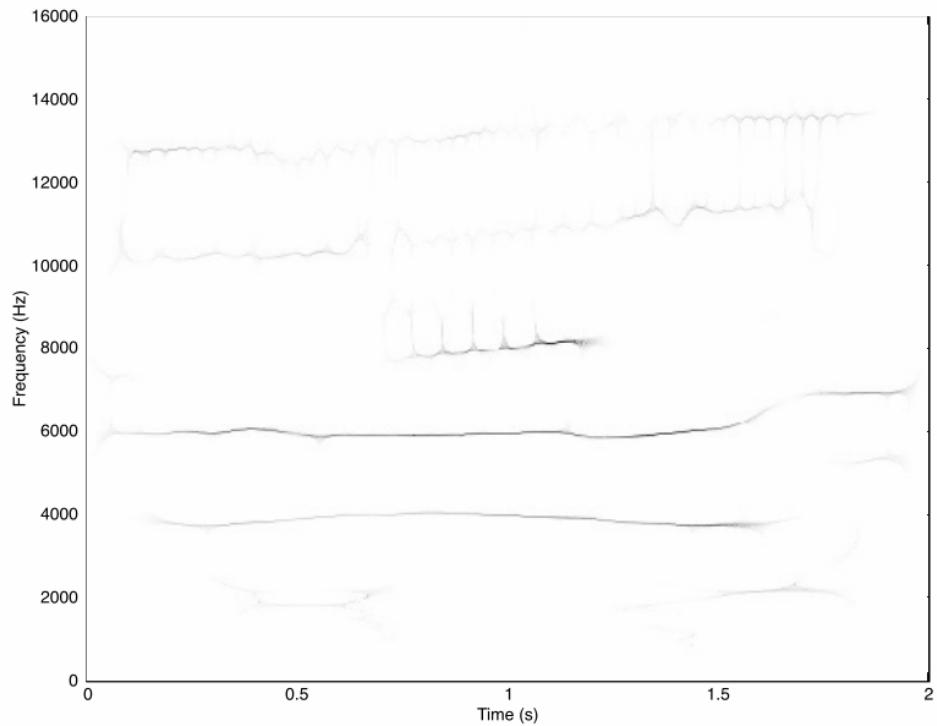


(b) Scalogram

Figure 6.2: Ordinary spectrogram and scalogram of a synthesized signal



(a) Reassigned spectrogram



(b) Reassigned scalogram

Figure 6.3: Reassigned spectrogram and scalogram of a synthesized signal

Chapter 7

Synthesis

The synthesis of a signal is the construction of a signal in the time domain. The synthesis of a signal associated with an atomic decomposition is the construction of this signal through a discrete superposition of *expansion functions*, which are the functions used in the expansion. These expansion functions are, just as the analysing functions used for analysis, elementary functions that are indexed by elements of a discrete index set. For this reason, the synthesis associated with an atomic decomposition can be seen as an *additive synthesis* since it is based on the construction of signals as a discrete sum of elementary functions. The general form of such a synthesis is

$$\alpha \mapsto \sum_{k \in \mathbb{I}} \alpha_k f_k$$

where $\{\alpha_k\}_{k \in \mathbb{I}}$ is a collection of scalars and where $\{f_k\}_{k \in \mathbb{I}}$ is a collection of expansion functions. The scalars associated with the expansion functions can be interpreted as the relative weight of the particular expansion function in the discrete superposition. The synthesis associated with an atomic decomposition constructs therefore a signal from a collection of scalars. This type of synthesis can therefore be seen as a mapping that maps a collection of scalars into a signal. The domain of this mapping is the transform domain and the range of this mapping is the time domain. The scalars in the transform domain will for this reason be called *transform coefficients*. Note that the domain and range of an analysis operator are exactly the reverse of that of a synthesis operator: the domain of an analysis operator is the time domain and its range is the transform domain. The analysis and synthesis operators are therefore adjoint operators.

The signal that is constructed from the transform coefficients is inseparable related to the expansion functions from which it is constructed. In fact, the properties of the expansion functions determine the properties that a synthesized signal can possess. For example, if the desired signal should have certain temporal and spectral properties, then the expansion functions should possess these properties, otherwise it is impossible to construct the desired signal using these expansion functions. The transform coefficients that are associated with the expansion functions determine, as stated previously, the relative weight of the particular expansion functions. For this reason, the constructed signal is completely described by these transform coefficients. The constructed signal in the time domain and the transform coefficients in the transform domain are two different representations of the same signal, just as the transform coefficients obtained through an analysis is just another representation of the analysed signal. Due to this it should be possible to map the signal in the time domain back to the transform coefficients in the transform domain. This adjoint mapping is the analysis operator and was discussed in the previous chapter. However, since both mappings are inseparable related to atomic decompositions, it will also be an important feature in the synthesis of signals.

Transform coefficients The transform coefficients from which a synthesis starts are, just as the transform coefficients obtained through an analysis, in general complex-valued. In the introduction of the previous chapter it was stated that these complex-valued transform coefficients can therefore be represented in several forms. Note that it was also stated that not all of these forms lead to an obvious interpretation. In fact, only the magnitude of the transform coefficients in polar form lead to an obvious interpretation. Unfortunately, in order to perfectly (re)construct a signal from a collection of transform coefficients in polar form, also the arguments or the phases of the coefficients are in general needed. However, the construction of a suitable collection of complex-valued transform coefficients is in general not obvious since the phase values can't be represented and constructed in an obvious form. In the case that the transform coefficients are obtained through an analysis, the coefficients possess already values of the phases and the synthesis operator can be applied on these transform coefficients without any problem. Note that in this case the synthesis is preceded by an analysis, which leads to so-called *analysis-resynthesis* — the topic of the next chapter. Since the second process in analysis-resynthesis is synthesis, methods that synthesize a signal from a collection of complex-valued transform coefficients will be described in this chapter in section §7.1, called *linear synthesis*. Now, if the synthesis mapping is not preceded by an analysis, the complex-valued transform coefficients should be constructed in order to synthesize signals. However, a procedure to construct these transform coefficients is not obvious since there isn't an obvious representation for the values of the phase — values that the transform coefficients should possess. Methods that allow the synthesis of a signal from the magnitudes of the transform coefficients only are therefore valuable. In that case, it would be possible to associate to a collection of expansion functions a collection of magnitudes, and to construct a signal by a linear combination of these expansion functions and corresponding magnitudes. The existence of such methods was, under the assumption that the expansion functions form a frame, proved by Balan, Casazza & Edidin (2006). Note that proving the existence of such methods doesn't denote specific methods explicitly. However, there are several methods that allow the construction of signals from the magnitude of coefficients only. Some of these methods will be treated in section §7.2, called *non-linear synthesis*.¹

7.1 Linear synthesis

The synthesis mapping that maps a collection of complex-valued transform coefficients into a signal is a *linear* mapping from the transform domain to the time domain. This kind of synthesis will therefore be called *linear synthesis*. The signal that is constructed through such a linear synthesis can be two types of signal. The first type is a signal that is constructed through a linear combination of the transform coefficients and a predetermined collection of expansion functions. The second type is a signal whose inner product with a predetermined collection of analysing functions yields again the original transform coefficients. The first type of signal is thus a signal that is constituted by a predetermined collection of expansion functions. The second type of signal is a signal that is constituted by a non-predetermined collection of expansion functions, but from which the transform coefficients can again be derived through an analysis operator associated with the predetermined collection of analysing functions. Note that from both types of signals the original transform coefficients can be derived through an analysis operator, but then the analysing functions associated with these analysis operators are different.

Since both types of signals that can be constructed through a linear synthesis are different, also the procedures to construct both types of signals are different. The first type of signal can be constructed through a direct calculation associated with the predetermined expansion functions. The second type can be constructed through an iterative procedure associated with the predetermined analysing functions.

¹During the final stages of this thesis some research on the inversion of the reassigned transforms was done by Holighaus, Průša & Søndergaard (2015). The inversion of the reassigned transforms allows also the (re)construction of a signal from magnitudes only and could therefore be included in §7.2. Due to the lack of time, these results aren't treated in this thesis, but they seem to have a lot of potential. They are therefore stated as steps for further research in the conclusion of this thesis.

7.1.1 Linear synthesis with predetermined expansion functions

The signal associated with a predetermined collection of expansion functions can, as stated previously, directly be computed. This type of linear synthesis constructs a signal as a discrete superposition of the predetermined expansion functions. The kind of linear synthesis can therefore be seen as an operator \mathcal{R} given by

$$\mathcal{R} : \ell^2(\mathbb{I}) \rightarrow L^2(\mathbb{R}), \quad \mathcal{R}c = \sum_{k \in \mathbb{I}} c_k f_k \quad (7.1)$$

where $\{c_k\}_{k \in \mathbb{I}} \subseteq \ell^2(\mathbb{I})$ is a collection of transform coefficients, and where $\{f_k\}_{k \in \mathbb{I}} \subseteq L^2(\mathbb{R})$ is a collection of expansion functions. If this collection of expansion functions forms a frame for $L^2(\mathbb{R})$, what is assumed, then the operator in (7.1) is the frame synthesis operator associated with this frame. In this case, the synthesis is the construction of a function as a discrete superposition of elements of the frame $\{f_k\}_{k \in \mathbb{I}}$. The transform coefficients from which the signal is constructed are in this case frame coefficients. By applying the frame analysis operator \mathcal{C} , associated with a dual frame of the frame from which the signal was constructed, on the constructed signal, results again in the original transform coefficients, that is,

$$c_k = \left\langle \sum_{k \in \mathbb{I}} c_k f_k, g_k \right\rangle_{k \in \mathbb{I}}$$

where $\{g_k\}_{k \in \mathbb{I}}$ is a dual frame of the frame $\{f_k\}_{k \in \mathbb{I}}$. Thus, after a signal is synthesized from a collection of frame coefficients, the frame coefficients can again be obtained from the synthesized signal.

7.1.2 Linear synthesis with predetermined analysing functions

The linear mapping that constructs a signal, whose frame coefficients are associated with a predetermined collection of analysing functions, is based on an iterative procedure rather than a direct computation — as is the case by the construction of a function as a linear combination of predetermined expansion functions. This iterative procedure consists of estimating a signal from which the transform coefficients can again be derived by the application of the analysis operator, associated with a predetermined collection of analysing functions. This iterative procedure allows the construction of such signal if, and only if, the collection of analysing functions form a frame. Since it is assumed that the collection of analysing functions forms a frame, the iterative procedure is applicable to the analysing functions studied in this chapter. This allows the construction of a signal, whose frame coefficients are associated with a predetermined collection of analysing functions, that correspond to the original transform coefficients. There are several algorithms that can perform the iterative procedure including the, what will be called, the frame synthesis algorithm, the Chebyshev algorithm and the conjugate gradient algorithm. In this thesis only the frame synthesis algorithm will be treated. The reader that is interested in the other algorithms as well is referred to the corresponding sections in Christensen's book (Christensen, 2003)

Frame synthesis algorithm The frame synthesis algorithm performs an iterative procedure that allows the construction of a signal whose frame coefficients, associated with a predetermined collection of analysing functions, are the transform coefficients from which the algorithm was started. This iterative procedure is intimately related to the frame bounds of the collection of analysing functions, and depends only on this collection of analysing functions and the transform coefficients. So, the frame synthesis algorithm estimates a signal — as a discrete superposition of expansion functions — without any a priori knowledge of the expansion functions. The signal is constructed by successive approximations that are completely based on the transform coefficients and the predetermined analysing functions. Applying the frame analysis operator associated with these analysing functions on the constructed signal, results again in the original transform coefficients.

The main principle behind the frame synthesis algorithm is that it approximates a signal based on the difference between the original frame coefficients and the frame coefficients of the approximated signal. The frame coefficients of the approximated signal are obtained by applying the frame analysis operator, associated with the predetermined collection of analysing functions, on a previous approximation of the analysed signal, which is obtained in a previous iteration.

The frame synthesis algorithm is explicitly given in pseudocode as algorithm 2. Here, \mathcal{C} and \mathcal{R} denotes the frame analysis and synthesis operator, respectively. Both operators are associated with the predetermined collection of analysing functions $\{f_k\}_{k \in \mathbb{I}}$.

Algorithm 2 Frame synthesis algorithm

```

1:  $x^{(0)} \leftarrow 0$ 
2: repeat for  $n = 1, 2, \dots$ 
3:  $c^{(n)} \leftarrow \{\langle x, f_k \rangle - \mathcal{C}x^{(n-1)}\}_{k \in \mathbb{I}}$ 
4:  $x^{(n)} \leftarrow x^{(n-1)} + (\frac{2}{A+B})\mathcal{R}c^{(n)}$ 
5: until  $\|x - x^{(n)}\|_2 = 0$ 
```

In words, the estimated signal after the zeroth iteration, $x^{(0)}$, is the constant function $x^{(0)} = 0$. The difference between the input frame coefficients $\{\langle x, f_k \rangle\}_{k \in \mathbb{I}}$ and the frame coefficients of the signal after the zeroth iteration, $\mathcal{C}x^{(0)}$, are just the input frame coefficients $\{\langle x, f_k \rangle\}_{k \in \mathbb{I}}$ since $\mathcal{C}x^{(0)}$ for all $k \in \mathbb{I}$. The estimated signal after the first iteration, $x^{(1)}$, is constructed by applying the frame synthesis operator on the constructed frame coefficients, multiplying this result by the constant $2/(A+B)$, and adding the estimated signal of the previous iteration, which is in this case $x^{(0)}$. By iterating this procedure, the estimated signal after the n 'th iteration, $x^{(n)}$, is constructed by applying the frame synthesis operator on the difference between the input frame coefficients and the frame coefficients of the estimated signal after the previous iteration, $\mathcal{C}x^{(n-1)}$, multiplying this result by the constant $2/(A+B)$, and adding the estimated signal of the previous iteration, $x^{(n-1)}$. By iterating this procedure until a certain stopping criterion is met, the difference between the estimated signal x and the estimated signal after the n 'th iteration, $x^{(n)}$, decreases with each iteration. It can be shown that the difference between x and $x^{(n)}$ is less than or equal to the constant $(\frac{B-A}{B+A})^n \|x\|_2$, where n denotes the number of iterations (Gröchenig, 1993). Note that in algorithm 2, the stopping criterion is a difference of zero between the constructed signal x and the estimated signal after the n 'th iteration, $x^{(n)}$. If this stopping criterion is met, then the constructed signal is the signal whose frame coefficients, with respect to the predetermined collection of analysing functions, are the original transform coefficients.

7.2 Non-linear synthesis

The synthesis mapping that maps a collection of magnitudes into a signal is a *non-linear* mapping from the transform domain to the time domain. This kind of synthesis will therefore be called *non-linear synthesis*. The aim of such a non-linear mapping is the construction of a signal of which the magnitude of its transform coefficients is as close as possible to the collection of magnitudes from which it was constructed. Thus, after analysing the constructed signal should, ideally, a collection of transform coefficients be obtained whose magnitudes are exactly the magnitudes from which the analysed signal was constructed.

The mapping that maps a collection of magnitudes into a signal can be described by the non-linear operator \mathcal{N} given by

$$\mathcal{N} : \ell^2(\mathbb{I}) \rightarrow L^2(\mathbb{R}), \quad \mathcal{N}|\alpha| = \sum_{k \in \mathbb{I}} |\alpha_k| g_k,$$

where $\{\alpha_k\}_{k \in \mathbb{I}} \subseteq \ell^2(\mathbb{I})$ is a collection of transform coefficients and where $\{g_k\}_{k \in \mathbb{I}} \subseteq L^2(\mathbb{R})$ is a collection of expansion functions. If the collection of scalars is a collection of frame coefficients and if the collection of expansion functions forms a frame, which is assumed, then the aim of the non-linear mapping is the construction of a signal of which the absolute value of its frame coefficients is as close as possible to the magnitudes from which the signal was constructed. As was already stated above, the absolute value of the frame coefficients of the signal f should, ideally, be equal to the collection of magnitudes $\{|\alpha_k|\}_{k \in \mathbb{Z}}$ from which it is constructed, that is, f should satisfy

$$\|\mathcal{C}f - \{|\alpha_k|\}\|_2 = \sqrt{\sum_{k \in \mathbb{I}} \left| |\langle f, f_k \rangle| - |\alpha_k| \right|^2} = 0,$$

where \mathcal{C} denotes the frame analysis operator associated with the frame $\{f_k\}_{k \in \mathbb{I}}$. However, not every collection of non-negative scalars is a collection absolute values of frame coefficients since it might be possible that there isn't a function that has these scalars as the absolute value of its frame coefficients. If this is the case, then it is not possible to construct a signal f that has the collection $\{|\alpha_k|\}_{k \in \mathbb{I}}$ as the absolute value of its frame coefficients. In this case, it is only possible to approximate such a signal, and the collection $\{|\alpha_k|\}_{k \in \mathbb{I}}$ will be called *non-representable*. Thus, a non-representable collection of magnitudes doesn't represent an existing signal. On the contrary, if it is possible to construct a signal that has the collection $\{|\alpha_k|\}_{k \in \mathbb{I}}$ as the absolute values of its frame coefficients, then this collection will be called *representable*. Thus, a representable collection of magnitudes corresponds to the magnitude of the frame coefficients of an existing signal. The non-linear mapping that constructs such a representable collection $\{|\alpha_k|\}_{k \in \mathbb{I}}$ is then the mapping given by

$$\mathcal{Q} : L^2(\mathbb{R}) \rightarrow \ell^2(\mathbb{I}), \quad \mathcal{Q}f = \{|\langle f, f_k \rangle|\}_{k \in \mathbb{I}}, \quad (7.2)$$

where $\{f_k\}_{k \in \mathbb{I}}$ is a collection of analysing functions. It can be deduced that if the mapping \mathcal{Q} is injective, then any signal can be constructed from the modulus of its frame coefficients up to a global constant phase shift. In this case, it is thus possible to represent such functions as

$$f \approx \sum_{k \in J} |\langle f, f_k \rangle| g_k, \quad (7.3)$$

where the collection of analysing functions $\{f_k\}_{k \in \mathbb{I}}$ and expansion functions $\{g_k\}_{k \in \mathbb{I}}$ are assumed to be dual frames of each other. Thus, in this case, it is possible to synthesize signals from a collection of magnitudes only.

The reasoning above is related to the reasoning in (Balan et al., 2006). In this paper, Balan, Casazza & Edidin proved that for frames with a sufficient “redundancy”, it is possible to (re)construct a signal up to a global constant phase shift from the modulus of its frame coefficients. There are several explicit reconstruction methods through which such a construction is possible, including the Griffin-Lim and fast Griffin-Lim algorithms. Both algorithms will be treated in the following subsection.

7.2.1 Algorithms for non-linear synthesis

Griffin-Lim algorithm The Griffin-Lim algorithm is an algorithm that was originally introduced by Griffin and Lim to estimate a signal whose magnitude of its short-time Fourier transform is as close as possible to a given collection of magnitudes (Griffin & Lim, 1984). While the algorithm was originally introduced as a method associated with the short-time Fourier transform, it can be generalized and applied to any type of frame (Perraudin et al., 2013). A discussion of the Griffin-Lim algorithm for general frames does, to the knowledge of the author, not exist in the literature yet. The algorithm given below will therefore be new.

The basic principle behind the Griffin-Lim algorithm for general frames is the estimation of a signal of which the magnitude of its frame coefficients is as close as possible to a given collection of magnitudes. The estimation of the signal is based on the minimization of the difference between the magnitude of the frame coefficients, of an estimated signal after a certain number of iterations, and the collection of magnitudes that is used as an input to the algorithm.

The actual Griffin-Lim algorithm is given as pseudocode in algorithm 3. Here $\{|\alpha|\}_{k \in \mathbb{I}}$ is a representable collection of magnitudes, \mathcal{R} is the frame synthesis operator associated with the expansion functions $\{f_k\}_{k \in \mathbb{I}}$ and \mathcal{C} is the frame analysis operator associated with the analysing functions $\{g_k\}_{k \in \mathbb{I}}$. Furthermore, it is assumed that $\{f_k\}_{k \in \mathbb{I}}$ and $\{g_k\}_{k \in \mathbb{I}}$ are dual frames of each other.

Algorithm 3 Griffin-Lim algorithm

```

1:  $c^{(0)} \leftarrow |\alpha|$ 
2: repeat for  $n = 1, 2, \dots$ 
3:  $x^{(n)} \leftarrow \mathcal{R}c^{(n-1)}$ 
4:  $d^{(n)} \leftarrow \mathcal{C}x^{(n)}$ 
5:  $c^{(n)} \leftarrow |\alpha|e^{i \arg d^{(n)}}$ 
6: until  $\|\mathcal{C}x - |\alpha|\|_2 = 0$ 
```

In words, the estimated signal after the first iteration, $x^{(1)}$, is constructed by applying the frame synthesis operator on the collection of magnitudes $\{|\alpha|\}_{k \in \mathbb{I}}$. On this estimated signal, $x^{(1)}$, the frame analysis operator is applied such that a collection of frame coefficients, $d_k^{(1)}$, is obtained. This constructed collection of frame coefficients consists of magnitudes $r_k^{(1)}$ and phases $\varphi_k^{(1)}$. These phases $\varphi_k^{(1)}$ and the magnitudes from which the algorithm was started, $|\alpha_k|$, are then combined into a new sequence of coefficients, $c_k^{(1)} = |\alpha_k|e^{i\varphi_k^{(1)}}$. Applying the frame synthesis operator on this new collection of frame coefficients, $c_k^{(1)}$, yields the estimated signal after the second iteration, $x^{(2)}$. By iterating this procedure until a certain stopping criterion is met, the difference between the magnitude of the frame coefficients of the estimated signal and the magnitudes from which this signal is estimated decreases with each iteration. The associated stopping criterion could be a maximum number of iterations or it could be a difference of zero between the magnitude of the frame coefficients of the estimated signal and the magnitudes from which the signal is estimated. In algorithm 3 there is chosen for this last criterion.

Fast Griffin-Lim algorithm The fast Griffin-Lim algorithm is an algorithm whose procedure is based on the Griffin-Lim algorithm. In fact, the fast Griffin-Lim algorithm is a generalization of the traditional Griffin-Lim algorithm and leads in general to a faster and more accurate convergence than the traditional version of the algorithm (Perraudin et al., 2013). Just as for the traditional Griffin-Lim algorithm, a discussion of the fast Griffin-Lim algorithm for general frames does, to the knowledge of the author, not exist in the literature yet. However, the algorithm will below be given for general frames.

The basic principle behind the fast Griffin-Lim algorithm is equal to the traditional Griffin-Lim algorithm, namely: the estimation of a signal of which the magnitude of its frame coefficients is as close as possible to a given collection of magnitudes. However, the main difference between both algorithms is that after an estimation of a collection of frame coefficients a new collection is constructed that depends on a scalar β .

The actual Griffin-Lim algorithm is given as pseudocode in algorithm 4. Here, as in the traditional version of the Griffin-Lim algorithm, is $\{|\alpha|\}_{k \in \mathbb{I}}$ a representable collection of magnitudes, \mathcal{R} the frame synthesis operator associated with the expansion functions $\{f_k\}_{k \in \mathbb{I}}$ and \mathcal{C} the frame analysis operator associated with the analysing functions $\{g_k\}_{k \in \mathbb{I}}$. Furthermore, it is assumed that $\{f_k\}_{k \in \mathbb{I}}$ and $\{g_k\}_{k \in \mathbb{I}}$ form, again, dual frames of each other.

Algorithm 4 Fast Griffin-Lim algorithm

```

1:  $p^{(0)} \leftarrow |\alpha|$ ,  $c^{(0)} \leftarrow |\alpha|$ 
2: repeat for  $n = 1, 2, \dots$ 
3:  $x^{(n)} \leftarrow \mathcal{R}p^{(n-1)}$ 
4:  $d^{(n)} \leftarrow \mathcal{C}x^{(n)}$ 
5:  $c^{(n)} \leftarrow |\alpha|e^{i \arg d^{(n)}}$ 
6:  $p^{(n)} \leftarrow c^{(n)} + \beta(c^{(n)} - c^{(n-1)})$ 
7: until  $\||\mathcal{C}x| - |\alpha|\|_2 = 0$ 
```

In words, the estimated signal after the first iteration, $x^{(1)}$, is constructed by applying the frame synthesis operator on the collection of magnitudes $|\alpha_k|$. On this estimated signal $x^{(1)}$, the frame analysis operator is applied such that a sequence of frame coefficients c_k is obtained. From the phase of these coefficients, φ_k , and $|\alpha_k|$, new coefficients are defined as $c_k = |\alpha_k|e^{i\varphi_k}$. From these coefficients another sequence p_k is constructed as $p_k = c_k + \beta(c_k - |\alpha_k|)$, where β is a scalar. By applying the frame synthesis operator on the sequence p_k , an estimated signal, $x^{(2)}$, is constructed. By iterating this procedure until a certain stopping criterion is met, the estimated signal will, depending on the representability of $|\alpha_k|$, approximate or converge to a signal whose magnitude of its frame coefficients equals $|\alpha_k|$.

Since the only difference between the fast Griffin-Lim algorithm and the traditional version of the algorithm is the construction of the sequence p_k , it can be deduced that both algorithms are equal when in the construction of this sequence $\beta = 0$. Something else that can be deduced from the fact that the only difference between the fast and traditional version of the algorithm is the construction of the sequence p_k , is that the speed and accuracy of convergence is determined by the scalar β . It can be shown that an optimal value of β near, but not greater than, 1, leads to the best results (Perraudin et al., 2013).

7.3 Examples

In this section several examples of synthesized signals will be given. All these examples consist of the construction of a signal in the time domain after having constructed a collection of transform coefficients in the transform domain.

Example 1 In this example, the linear and non-linear mapping, through which a signal can be constructed from a collection of magnitudes only, will be compared. In figure 7.1, a collection of magnitudes that corresponds to a Gabor frame is shown. These magnitudes are constructed without an analysis. A signal can be constructed from these coefficients through the Griffin-Lim algorithm or the frame synthesis operator. The collection of Gabor coefficients that corresponds to the signal that is constructed through the Griffin-Lim algorithm is shown in figure 7.2a. The collection of Gabor coefficients that corresponds to the signal that is constructed through the frame synthesis operator is shown in figure 7.2b. It can immediately be deduced from these figures that the construction of a signal from a collection of magnitudes through the Griffin-Lim algorithm results in a valuable signal, whereas the construction of a signal from a collection of magnitudes through the frame synthesis operator doesn't. The signal that is constructed from the magnitudes through the Griffin-Lim algorithm doesn't have exactly the same magnitudes as where it was synthesized from. This is caused by the fact that the magnitudes from which the signal was synthesized isn't a representable collection of magnitudes of Gabor coefficients — see §7.2. For this reason, the signal that is constructed from these magnitudes through the Griffin-Lim algorithm is the signal of which the magnitude of its Gabor coefficients is as close as possible to the given collection of magnitudes. Although the magnitude of the Gabor coefficients in figure 7.1 and figure 7.2a aren't exactly equal, the main characteristics are still present in the constructed signal of figure 7.2a. This is definitely not the case by the constructed signal of figure 7.2b. Thus, applying the frame synthesis operator on a collection of magnitudes does in general *not* correspond to a valuable signal.

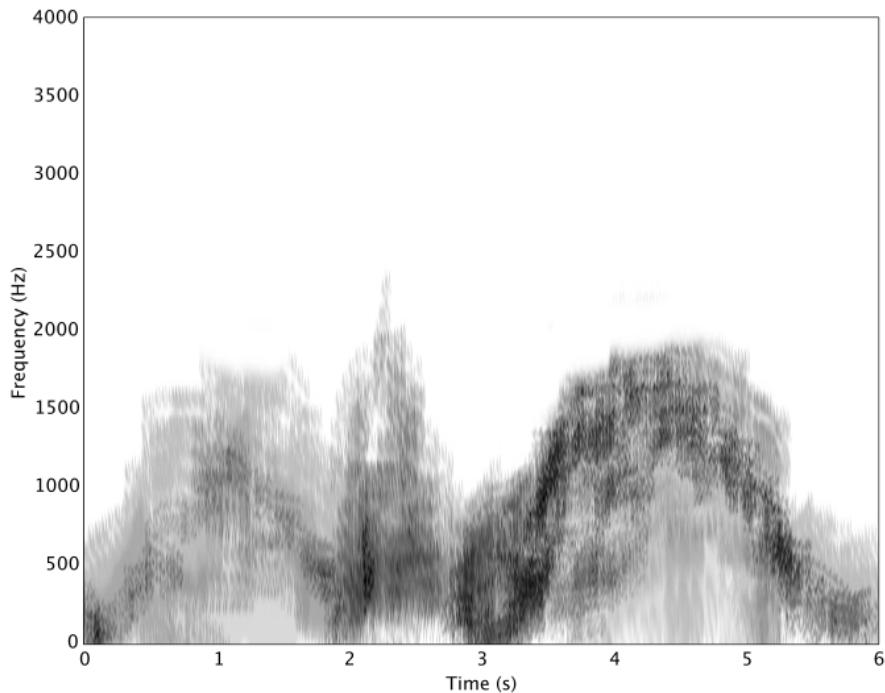
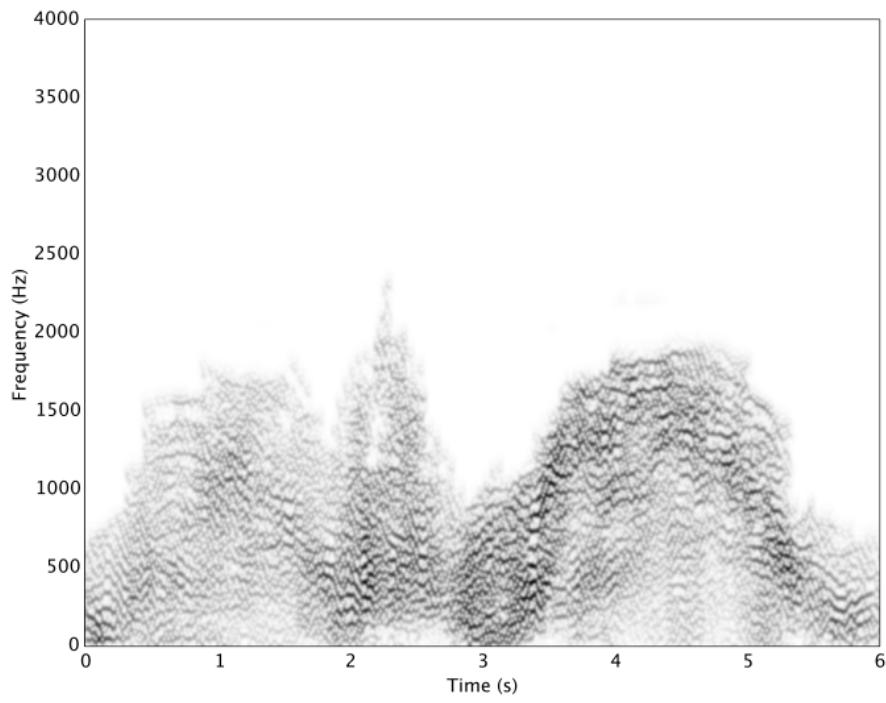
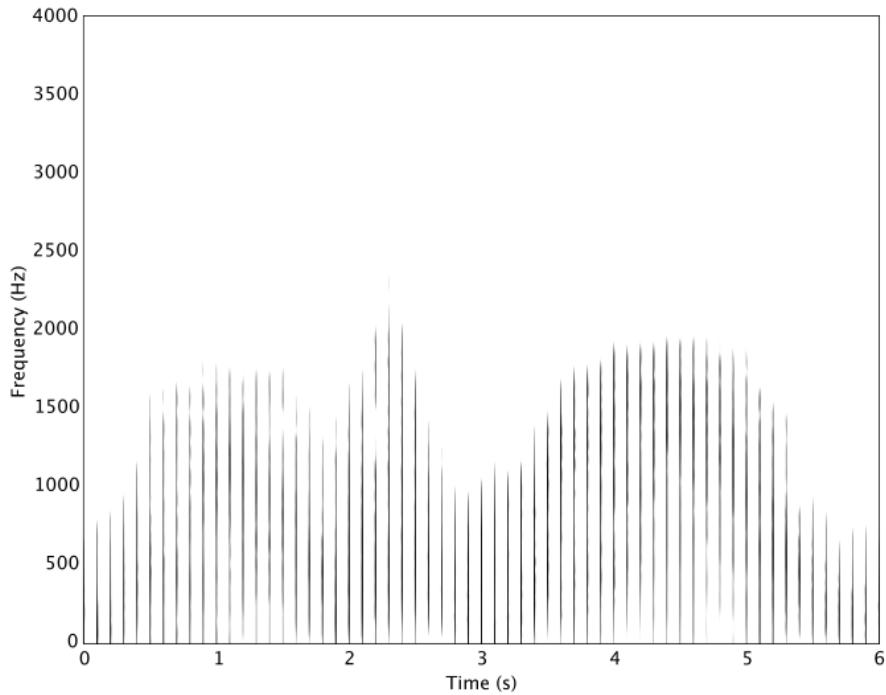


Figure 7.1: Constructed magnitudes



(a) Magnitude of Gabor coefficients after non-linear synthesis



(b) Magnitude of Gabor coefficients after linear synthesis

Figure 7.2: Magnitude of Gabor coefficients

Example 2 In this example, a signal is reconstructed from the magnitudes of the Gabor coefficients shown in figure 7.2a. The main difference between the reconstruction of a signal from the constructed magnitudes in figure 7.1 and the magnitudes of the Gabor coefficients in figure 7.2a is that the Gabor coefficients in figure 7.2a are frame coefficients whereas the constructed magnitudes in figure 7.1 aren't. For this reason, the construction of a signal from the constructed magnitudes was an *approximation*, whereas the construction of a signal from the magnitude of the Gabor coefficients is an *reconstruction*. The magnitude of the Gabor coefficients of the signal that was constructed from the collection of magnitude of Gabor coefficients shown in figure 7.2a is shown in figure 7.3. Theoretically, both collections of magnitudes of Gabor coefficients should be equal, but in practice this is in general not the case due to rounding errors et cetera. However, the difference between an approximation and a reconstruction can clearly be seen from the figures 7.1 and 7.2a respectively 7.2a and 7.3.

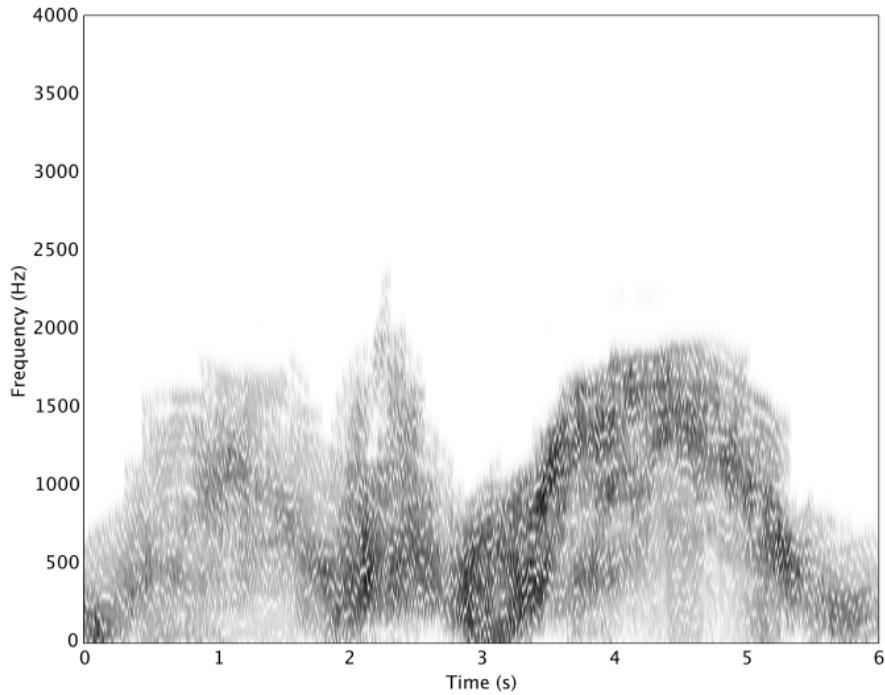


Figure 7.3: Magnitude of Gabor coefficients of reconstructed signal

Chapter 8

Analysis-Resynthesis

The analysis-resynthesis of a signal is the construction of a signal in the time domain after having extracted information from a(nother) signal in the time domain. An analysis-resynthesis is thus a mapping constituted by an analysis and a synthesis mapping. The synthesis mapping is in this case preceded by the analysis mapping. In the context of atomic decompositions is an analysis-resynthesis a mapping that combines the analysis and synthesis associated with atomic decompositions. For this reason, there is with this type of analysis-resynthesis a mapping that maps a signal into a collection of transform coefficients, and a mapping that constructs a signal from a collection of transform coefficients. An analysis-resynthesis maps a signal thus first into the transform domain and after that back into the time domain. An analysis-resynthesis is therefore a mapping from the time domain into the time domain. The general form of such an analysis-resynthesis mapping is

$$f \longmapsto \sum_{k \in \mathbb{I}} \tilde{c}_k f_k,$$

where $\{\tilde{c}_k\}_{k \in \mathbb{I}}$ is a collection of scalars that is based on a collection of transform coefficients $\{\langle f, h_k \rangle\}_{k \in \mathbb{I}}$, which are obtained through an analysis.. Furthermore, f is the signal to be analysed, $\{h_k\}_{k \in \mathbb{I}}$ a collection of analysing functions, and $\{f_k\}_{k \in \mathbb{I}}$ a collection of expansion functions. Note that both the collection of analysing functions and the collection of expansion functions are indexed by elements of the same index set, which is in this case denoted by \mathbb{I} . This is necessary since with each expansion function there should be one corresponding transform coefficient. For analysis-resynthesis in general there are no further restrictions on the type of collections of analysing and expansion functions.

If the outcome of the analysis-resynthesis should be the analysed signal again, then the collections of analysing and expansion functions should form dual frames, and the obtained transform coefficients shouldn't be modified. However, reconstructing an analysed signal through an analysis-resynthesis is in general *not* the goal of an analysis-resynthesis, but the fact that it guarantees the possibility for perfect reconstruction is fundamental for the goal of an analysis-resynthesis, which is the construction of new signals based on information obtained from analysed signals. These new signals are in general modified versions of the analysed signals, and in order to get only the desired modifications, it is essential that the analysed signal could also have been perfectly reconstructed. For example, if it is not possible to perfectly reconstruct a signal, that is, if there are undesired properties in the reconstructed version of the analysed signal, then these undesired properties will also be present in the intentionally modified version of the analysed signal. Thus, the fact that perfect reconstruction is possible with an analysis-resynthesis allows the possibility to construct modified versions of an analysed signal with only the desired modifications, which is, again, the goal of an analysis-resynthesis. For this reason, the types of analysis and synthesis mappings that lead to a perfect reconstruction of a signal are important in an analysis-resynthesis. The different types of analysis and synthesis mappings were discussed in chapter 6 and 7, respectively. However, since the combination of these mappings lead to atomic decompositions and allow analysis-resynthesis, their properties will also be an important feature in this chapter.

Transform coefficients The transform coefficients associated with an analysis-resynthesis are obtained through an analysis mapping. For this reason, these transform coefficients are in general complex-valued. These transform coefficients can therefore be represented in several forms, but in general does only the magnitude of these coefficients lead to an obvious representation — see the introduction of chapter 6. However, the argument of the transform coefficients are, as in the case of a synthesis, in general of fundamental importance for the construction of a signal. For this reason, both the magnitude and the argument of the transform coefficients are important for an analysis-resynthesis. In order to modify a signal decently through an analysis-resynthesis, both the magnitude and argument of the transform coefficients should be modified.

8.1 Linear analysis-resynthesis

The analysis-resynthesis mapping is constituted by two mappings, namely an analysis and a synthesis mapping. The analysis mapping maps a signal from the time domain into a collection of inner products, the transform coefficients, in the transform domain. The synthesis mapping maps a collection of transform coefficients in the transform domain into a signal in the time domain. If both the analysis and synthesis mapping are linear mappings, then the analysis-resynthesis mapping, which is constituted by these mappings, will be called a *linear analysis-resynthesis*.

8.1.1 Combination of linear analysis and synthesis mappings

The linear analysis-resynthesis mapping consists by definition of a linear analysis and a linear synthesis mapping. Recall that the linear analysis mapping is a mapping of the form

$$L^2(\mathbb{R}) \rightarrow \ell^2(\mathbb{I}), \quad f \mapsto \{\langle f, h_k \rangle\}_{k \in \mathbb{I}}, \quad (8.1)$$

where f is the signal to be analysed and where $\{h_k\}_{k \in \mathbb{I}}$ is a collection of analysing functions. The linear synthesis mapping is the adjoint operator of the linear analysis operator and is given by

$$\ell^2(\mathbb{I}) \rightarrow L^2(\mathbb{R}), \quad \{c_k\}_{k \in \mathbb{I}} \mapsto \sum_{k \in \mathbb{I}} c_k f_k, \quad (8.2)$$

where $\{c_k\}_{k \in \mathbb{I}}$ is a collection of transform coefficients and where $\{f_k\}_{k \in \mathbb{I}}$ is a collection of expansion functions. The main characteristic of a linear analysis-resynthesis is now that the transform coefficients of the mapping given in (8.2) are derived from coefficients obtained through a linear analysis mapping, as given in (8.1). The word *derived* is used here to indicate that the transform coefficients in the synthesis mapping doesn't have to be original transform coefficients that are obtained through an analysis mapping, but can also be modified versions of such original transform coefficients.

The linear analysis-resynthesis mapping is thus, by definition, the concatenation of the mappings given in (8.1) and (8.2), but the actual collections of analysing and expansion functions aren't determined by the definition of the linear analysis-resynthesis mapping. Consequently, the actual procedures that represent the mappings given in (8.1) and (8.2) aren't determined by the definition of the linear analysis-resynthesis mapping. For example, the linear analysis mapping as given in (8.1) can be both types of analysis mappings that are described in §6.1, namely the frame analysis operator and the frame analysis algorithm. Similarly, the synthesis mapping as given in (8.2) can be both types of synthesis mappings that are described in §7.1, namely the frame synthesis operator and the frame synthesis algorithm. Note, however, that in order to obtain a proper linear analysis-resynthesis mapping, the linear analysis and synthesis mappings should be combined decently and can't be combined arbitrarily. The following enumeration recalls some of the main characteristics of the linear analysis and synthesis mappings, and states how these mappings should be combined in order to get a proper linear analysis-resynthesis mapping:

- (i) Frame analysis operator: The frame analysis operator constructs a collection of frame coefficients that correspond to a predetermined collection of analysing functions — see §6.1.1. In order to perfectly reconstruct an analysed signal from these frame coefficients, a synthesis mapping that reconstructs a signal from a collection of frame coefficients, associated with a predetermined collection of analysing functions, should be used. There are two of such synthesis mappings, namely the frame synthesis operator and the frame synthesis algorithm. In the case of the frame synthesis operator, the collection of expansion functions should form a dual frame of the predetermined collection of analysing functions that was used in order to obtain the frame coefficients — see §7.1.1. In the case of the frame synthesis algorithm, the mapping is entirely based on the predetermined collection of analysing functions and its frame bounds — see §7.1.2.
- (ii) Frame analysis algorithm: The frame analysis algorithm constructs a collection of frame coefficients that corresponds to a predetermined collection of expansion functions — see §6.1.2. In order to perfectly reconstruct an analysed signal from these frame coefficients, a synthesis mapping associated with the predetermined collection of expansion functions should be used. There is only one such synthesis mapping, namely the frame synthesis operator associated with the predetermined collection of expansion functions. It is in general not possible to use the frame synthesis algorithm in combination with the frame analysis algorithm since the frame analysis algorithm corresponds with a predetermined collection of expansion functions, whereas the frame synthesis algorithm corresponds with a predetermined collection of analysing functions.
- (iii) Frame synthesis operator: The frame synthesis operator constructs a function that corresponds with a predetermined collection of expansion functions — see §7.1.1. Therefore both the frame analysis operator and the frame analysis algorithm can be used in a linear analysis-resynthesis — see points (i) and (ii) for the reasoning behind these combinations.
- (iv) Frame synthesis algorithm: The frame synthesis algorithm constructs a function that corresponds with a predetermined collection of analysing functions — see §7.1.1. Therefore only the frame analysis operator can be used in a linear analysis-resynthesis — see points (i) and (ii) for the reasoning behind this combination.

8.1.2 Modification of transform coefficients

In the introduction of this chapter, it was stated that it is in general not the goal of an analysis-resynthesis to perfectly reconstruct an analysed signal. On the contrary, the main goal of an analysis-resynthesis is in general to construct new signals based on information obtained from analysed signals. An appropriate method to construct such new signals is to modify the transform coefficients of an analysed signal and map the modified transform coefficients back into a signal in the time domain. The actual modification of the transform coefficients happens trivially always after an analysis mapping. However, the actual determination of the type of modification can happen both before or after the analysis. In the first case, when the type of modification is determined *a priori*, the modification cannot be based on information obtained through the analysis of the analysed signal. In the second case, when the type of modification is determined *a posteriori*, the modification could be based on information obtained through the analysis of the analysed signal. Examples of both types of modifications are given in the following section.

8.2 Examples

In this section several examples of types of modifications of transform coefficients will be given. All these examples consist of a representation of an analysed signal in terms of transform coefficients, a representation of a modified version of these transform coefficients and a representation of transform coefficients of the signal that is constructed through an analysis-resynthesis.

Example 1 In this example, a pink noise signal will be analysed, modified by multiplication with a function in the transform domain, and a new signal will be synthesized from these modified transform coefficients. In figure 8.1 the magnitude of a collection of wavelet coefficients associated with a pink noise signal is shown. These wavelet coefficients are the same as the wavelet coefficients in figure 6.1b, and therefore is for a discussion of these wavelet coefficients referred to §6.2. In figure 8.2, a function that is constructed on purpose is shown. This function is a function in the transform whose value is 1 in the regions where its colour is black, and whose value is 0 where its colour in figure 8.2 is white. The purpose of this constructed function is to multiply it point-wise with the wavelet coefficients obtained through the analysis of the pink noise signal. The product of the wavelet coefficients and the constructed function is then a collection of wavelet coefficients whose values are unchanged in the region where the construction function has a value of 1 and whose values are 0 in the region where the construction function has a value of 0. The product of the wavelet coefficients and the symbol is shown in figure 8.3. The result of this analysis-resynthesis is that the wavelet coefficients of the analysed signal are restricted to a certain region in the time-frequency plane. The wavelet coefficients that are outside of this region are set to 0.

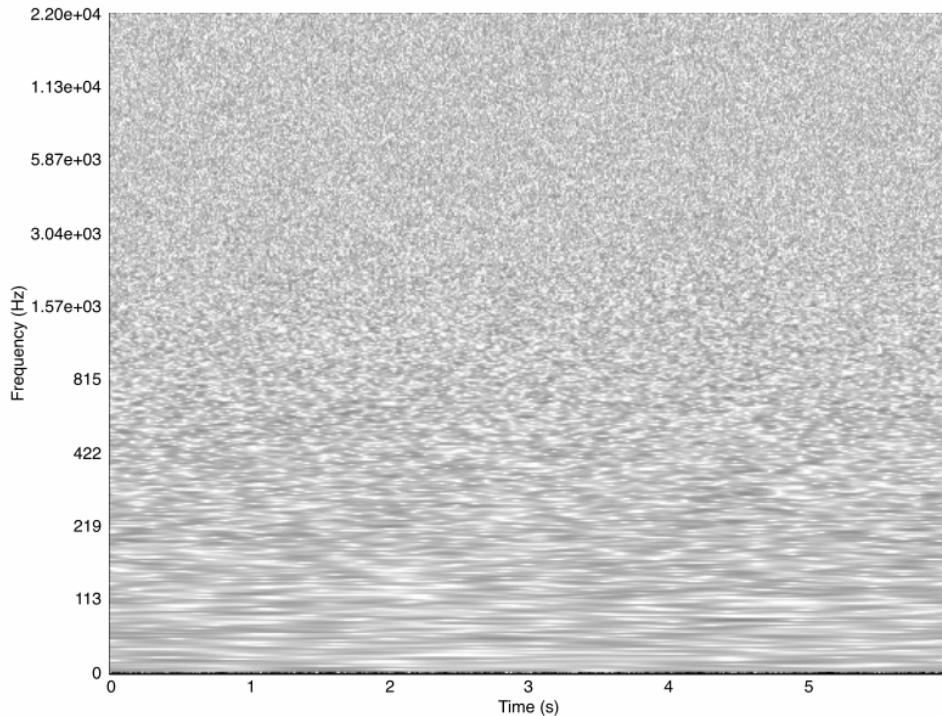


Figure 8.1: Wavelet coefficients of pink noise

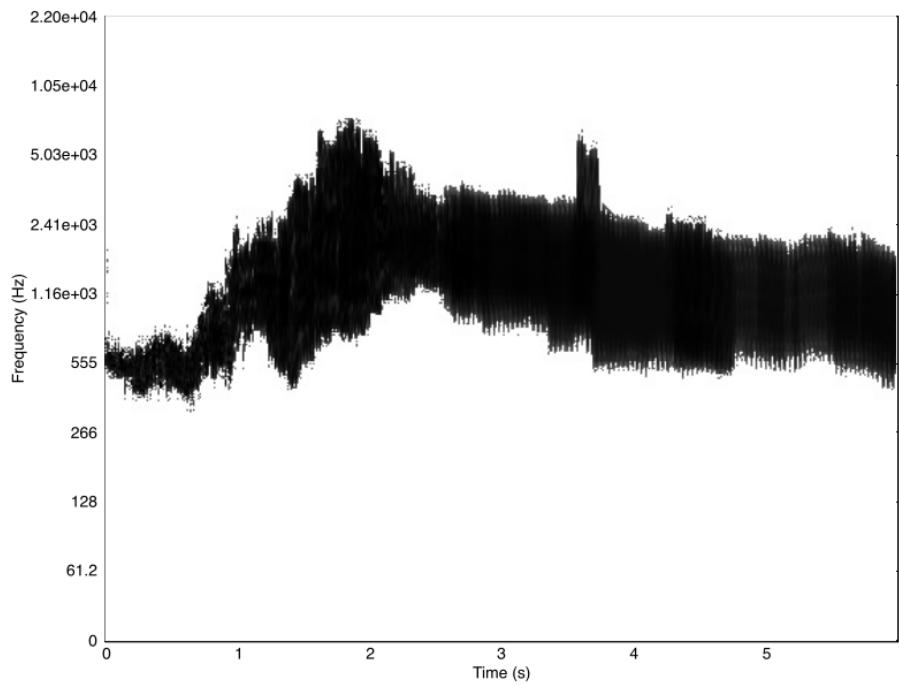


Figure 8.2: Constructed function for modification

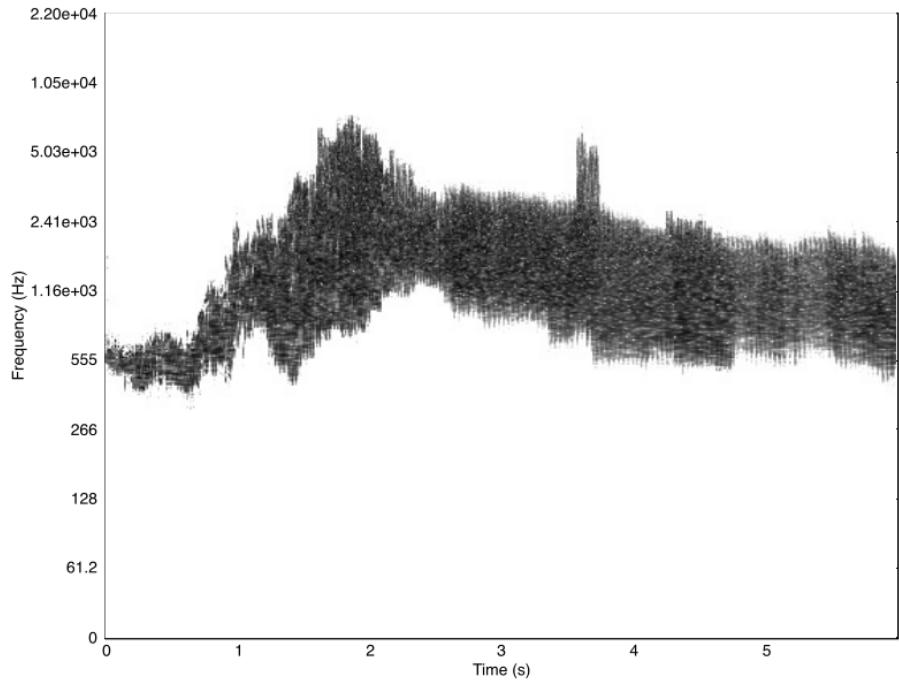


Figure 8.3: Modified wavelet coefficients

Example 2 In this example, a pink noise signal will be analysed, modified in the transform domain, and a new signal will be synthesized from these modified transform coefficients. In figure 8.4, the magnitude of a collection of Gabor coefficients associated with a pink noise signal is shown. These Gabor coefficients are the same as the Gabor coefficients shown in figure 6.1a, and therefore is for a discussion of these Gabor coefficients referred to §6.2. The modification of this collection of Gabor coefficients is based on information extracted from these Gabor coefficients. The actual modification is that Gabor coefficients with a value below a certain threshold are set to 0. This modification results in a subset of the complete collection of Gabor coefficients, namely the subset whose elements are Gabor coefficients with a value above a certain threshold. Since the energy of a pink noise signal is primarily concentrated at lower frequencies, the constructed subset of Gabor coefficients consists primarily of Gabor coefficients that correspond to a low frequency. An example of such a subset of Gabor coefficients is shown in figure 8.5. The threshold that was used with this modification was constant for all positions on the time-frequency plane, and therefore all Gabor coefficients whose value was below this threshold were set to 0. However, it is possible to change the threshold over time. In this case, the Gabor coefficients whose value are below a certain threshold in a certain time interval are set to 0. An example of such a modification is given in figure 8.6.

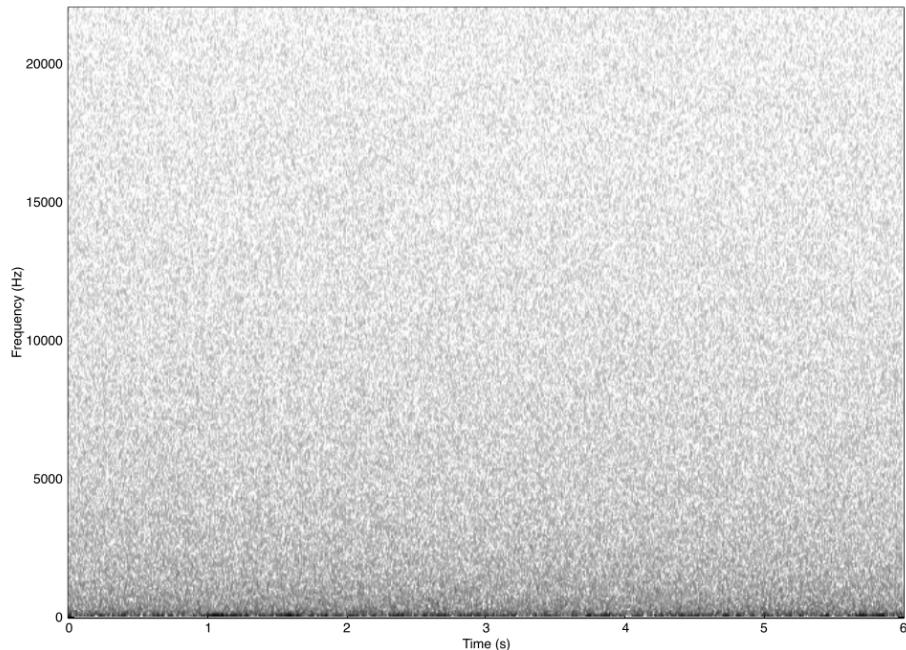


Figure 8.4: Magnitude of Gabor coefficients of a pink noise signal

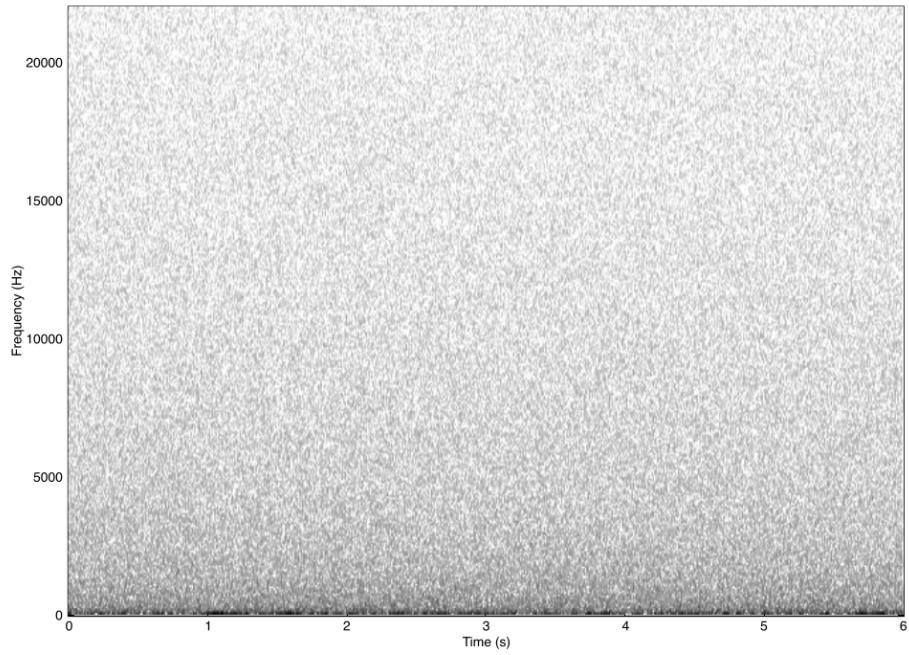


Figure 8.5: Magnitude of modified Gabor coefficients 1

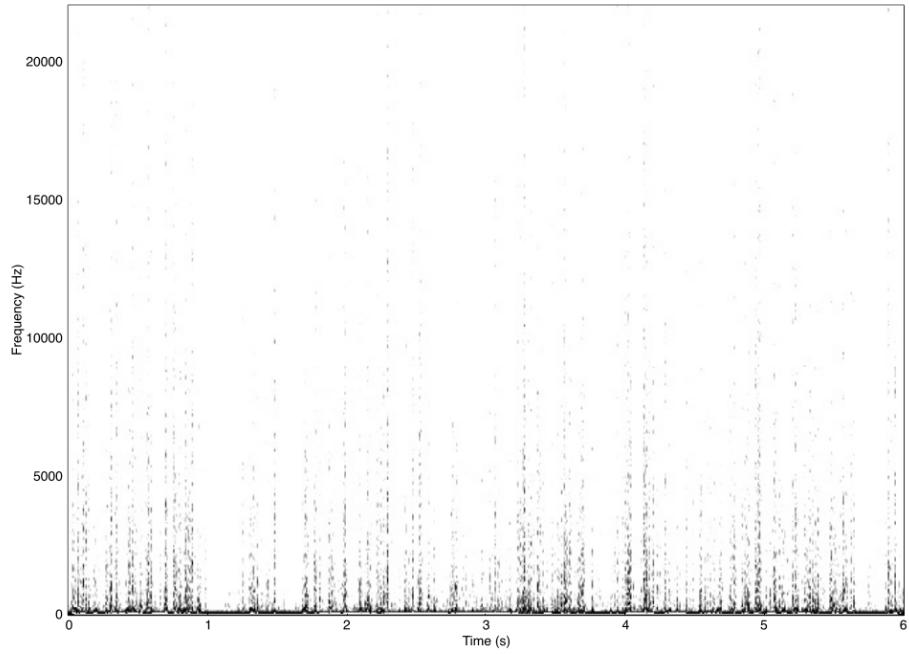


Figure 8.6: Magnitude of modified Gabor coefficients 2

Epilogue

In this epilogue, the methods for the analysis and synthesis of acoustical signals that were treated in the main part of this thesis will be compared to their origins. These origins are the theories of Dennis Gabor and Iannis Xenakis, which were introduced in the prologue of this thesis. This epilogue functions therefore as a comparison of the methods treated in the main matter of the thesis and the theories of Gabor and Xenakis.

Dennis Gabor proposed that any arbitrary signal can be represented as a sum of elementary functions that have time and frequency as independent variables. These elementary functions can be interpreted as translated and modulated versions of a single Gaussian window function. The collection of elementary functions that Gabor proposes can therefore be interpreted as a specific case of a Gabor system. If this Gabor system forms a frame, then Gabor's proposition regarding the representation of a signal as a sum of these elementary functions should be true. However, it can be shown that the specific Gabor system that Gabor proposed doesn't provide a valuable time-frequency representation of a signal. In order to get a valuable time-frequency representation of the signal, the elementary functions proposed by Gabor should be interpreted more carefully (Janssen, 1981). The technical details regarding this "more careful interpretation" will not be discussed here, but the main point is that the type of elementary functions that Gabor proposed can, under certain conditions, indeed lead to valuable time-frequency representations. In this case, such a collection of elementary functions forms an appropriate Gabor frame. Furthermore, as long as the assumption of an appropriate Gabor frame is satisfied, all types of window functions can be used as an atom. Thus, the Gaussian window function proposed by Gabor isn't the only window function that can be used. This has as advantage that for a certain application the most appropriate window function can be used.

Gabor's proposition can thus also be realized by translating and modulating a single window function other than a Gaussian. In fact, this notion can even be extended further. In the context of non-stationary Gabor frames it was namely shown that one can construct a frame by translating and modulating a collection of window functions. Thus, instead of using only one single window function for the generation of a collection of elementary functions, a collection of several window functions can be used. In §5.4.1 it was shown that when only one single window function was used for the generation of the collection of elementary functions, then this corresponds to a standard Gabor system. Gabor's original proposition can therefore be seen as a specific case of the general theory associated with non-stationary Gabor systems. This fact isn't only of interest for the sake of generalization, but is also of interest for certain applications. For example, Gabor's proposition leads to a time-frequency plane in which all "logons"—to use Gabor's term— are of an equal shape, but the shapes of the logons that can be obtained with a non-stationary Gabor system may vary over time or frequency. This property is of interest for the study of signals with special characteristics since the shapes of the logons can then be adapted to these characteristics.

Just as Gabor's suggestion regarding an elementary function for the decomposition of arbitrary signals can be extended, also Xenakis's suggestion of such an elementary function can be extended. Xenakis proposed, just as Gabor, that any signal can be represented as a sum of elementary functions that have time and frequency as independent variables. The elementary functions that Xenakis proposed for this purpose were Gabor's elementary functions, and translated and modulated versions of a rectangular window function. However, these elementary signals can, according to the same reasoning used above, be extended to translated and modulated versions of

any single window function. Thus, as long as the elementary functions form an appropriate Gabor frame, then they can be used to synthesize signals. In fact, it would even be possible to extend the elementary signals proposed by Xenakis to a non-stationary Gabor system. However, this will also drastically change Xenakis' concept of a screen. For this reason, the concept of a screen will first be compared with the methods that were treated in this thesis.

A screen is an “intensity-frequency plane” with amplitude and frequency as independent axes. A screen could therefore easily be compared with an amplitude spectrum, that is, the magnitude of a Fourier transform. In fact, a screen is virtually equal to an amplitude spectrum. The only difference between a screen and an amplitude spectrum is that the amplitude axis of a screen is independent of the frequency axis, whereas the amplitude axis of an amplitude spectrum is a function of frequency, and thus depends on the frequency axis. The independence of the amplitude axis in a screen allows that elementary functions with the same frequency can have different values as amplitude. However, the result of several elementary functions having the same frequency, but different amplitudes, is just a summation in amplitude. There will not be any audible difference. These elementary functions can therefore be interpreted as one elementary function with the sum of the individual amplitudes as amplitude. The concept of a screen can therefore be reduced to an amplitude spectrum. The concept of a book, that is, a sequence of screens, can now be interpreted as a sequence of amplitude spectra and is therefore virtually equal to the frame synthesis operator associated with a Gabor frame.

In the case that a screen isn't reduced to an amplitude spectrum, then a screen will lead to a non-unique representation of a signal. The non-uniqueness of such a representations has as a result that it isn't possible to use the concept of a screen for the analysis of signals. As a result, analysis-resynthesis is also not possible with screens. Note that Xenakis developed the concept of a screen purely for the synthesis of signals, thus the fact that it isn't possible to analyse signals with screens doesn't detract the purpose for which screens were originally developed. However, there are even some characteristics of the original concept of screens due to which the synthesis of signals might become problematic. Remember that within the concept of a screen the elementary signals are determined by a value of amplitude, time and frequency. The synthesis of a signal through screens is therefore a synthesis from magnitudes only. There isn't any information on the phase of an elementary signal within the concept of a screen. In §7.2, it was shown that synthesis from magnitudes only might be problematic when there isn't an appropriate algorithm used. Thus, even when the concept of a screen is reduced to an amplitude spectrum, it is necessary to use an appropriate algorithm for the synthesis. Therefore, in order to synthesize signals in the manner proposed by Xenakis, it is necessary to reduce the concept of a screen to an amplitude spectrum and to use one of the algorithms described in §7.2. In this case, the method for the synthesis of signals proposed by Xenakis equals the synthesis of a signal from a spectrogram.

Now, after having compared the original concept of a screen with some of the methods described in this thesis, the concept of a screen will be extended by relating it to the theory of non-stationary Gabor systems. In the prologue it was stated that a screen can be considered as a temporal unit since it represents a constant time duration. In fact, when screens are interpreted as amplitude spectra, this constant time duration is caused by the use of one single window function, which has exactly this time duration. If instead of one single window function, multiple window functions are used, then this allows that the time duration that a screen represents can be variable over time. Thus, in the case that Xenakis' concepts are related to non-stationary Gabor systems, then the duration that a screen possesses can differ over time.

Conclusion

In conclusion of this thesis the main parts and results of the main matter will be summarized, and a next step for future studies will be stated.

In this thesis, it was attempted to give a systematic exposition of methods for atomic decompositions, and their possibilities for the analysis and synthesis of acoustical signals. In order to do so, both the theory associated with specific methods for atomic decompositions and the general theory of frames were treated. These specific methods were first introduced as independent theories, but were later interpreted within the general framework of frame theory. The treatment of the specific methods for atomic decompositions as independent theories clearly showed the motivation behind the existence of these methods. Aside that, these treatments lead to several useful interpretations of the same method. On the other hand, the interpretation of the specific methods for atomic decompositions as special cases of frame theory clearly showed that the principles behind all these different methods are in fact the same. For this reason, a lot of comparable methods for analysis and synthesis can be treated in the same manner. For example, all methods for the analysis of signals that decompose a signal into elementary functions can be treated as the analysis associated with atomic decompositions. Similarly, all methods for the synthesis of signals that construct a signal as a discrete superposition of elementary functions can be treated as the synthesis associated with atomic decompositions. Thus, all these different methods can be treated and used in the same manner, namely as a concrete example of a part of an atomic decomposition. Since several methods for analysis and synthesis can be considered as parts of an atomic decomposition, it is also possible to connect them to another part of an atomic decomposition. So, it is possible to connect to a method that is in general exclusively used for analysis a synthesis counterpart. An example of such an analysis method is the constant-Q transform, which can easily be inverted when interpreted within frame theory. In the same manner, it is possible to connect to a method that is in general exclusively used for synthesis an analysis counterpart. Examples of such synthesis methods are types of granular synthesis, whose main concepts can also be used for the analysis of signals, when interpreted within frame theory.

Frame theory forms a firm foundation for several methods for the analysis and synthesis of acoustical signals. However, representations or models that are based on frame theory are sometimes more interesting than frame theory on itself. For example, the reassigned versions of methods for analysis are more interesting for the analysis of acoustical signals than the ordinary versions of a linear analysis are. Another example, a non-linear synthesis, which allows the synthesis from the magnitude of frame coefficients only, is more interesting than an ordinary linear synthesis. Note that both of these examples can't exist without the firm foundation of frame theory. However, they clearly show that frame theory on itself is in general not sufficient enough to provide the most interesting methods for the analysis and synthesis of acoustical signals.

To finalize this conclusion a possible next step for future studies will be stated. In the opinion of the author, the reassigned methods and the methods for the synthesis from magnitudes only are the most interesting methods, of the methods that are treated in this thesis, for the analysis respectively the synthesis of acoustical signals. For this reason, the synthesis from reassigned methods would be of special interest since this would combine the properties of reassignment and the synthesis from magnitudes only. Although the synthesis from reassigned methods is in general considered as impossible, there seem to be some possibilities. For this reason, this would be an interesting next step for future studies.

References

- Auger, F. & Flandrin, P. (1995). Improving the readability of time-frequency and time-scale representations by the reassignment method. *Signal Processing, IEEE Transactions on*, 43(5), 1069–1089.
- Balan, R., Casazza, P., & Edidin, D. (2006). On signal reconstruction without phase. *Applied and Computational Harmonic Analysis*, 20(3), 345–356.
- Balazs, P., Dörfler, M., Jaillet, F., Holighaus, N., & Velasco, G. (2011). Theory, implementation and applications of nonstationary gabor frames. *Journal of computational and applied mathematics*, 236(6), 1481–1496.
- Brown, J. C. (1991). Calculation of a constant Q spectral transform. *The Journal of the Acoustical Society of America*, 89(1), 425–434.
- Brown, J. C. & Puckette, M. S. (1992). An efficient algorithm for the calculation of a constant Q transform. *The Journal of the Acoustical Society of America*, 92(5), 2698–2701.
- Chassande-Mottin, E. & Flandrin, P. (2001). Reassigned scalograms and singularities. In *European Congress of Mathematics*, (pp. 583–590). Springer.
- Christensen, O. (2003). *An Introduction to Frames and Riesz Bases*. Basel: Birkhäuser.
- Cohen, L. (1995). *Time-Frequency Analysis*. New Jersey: Prentice Hall PTR.
- Daubechies, I. (1992). *Ten Lectures on Wavelets*. Pennsylvania: SIAM.
- Daubechies, I., Grossmann, A., & Meyer, Y. (1986). Painless nonorthogonal expansions. *Journal of Mathematical Physics*, 27(5), 1271–1283.
- De Poli, G., Piccialli, A., & Roads, C. (1991). *Representations of Music Signals*. Cambridge, MA: MIT Press.
- Duffin, R. J. & Schaeffer, A. C. (1952). A class of nonharmonic fourier series. *Transactions of the American Mathematical Society*, 72, 341–366.
- Feichtinger, H. & Gröchenig, K. (1992). Non-orthogonal wavelet and gabor expansions, and their group representations. In G. Beylkin, R. Coifman, & Daubechies (Eds.), *Wavelets and their applications* (pp. 353–376). Boston, MA: Jones and Bartlett.
- Flanagan, J. L. & Golden, R. (1966). Phase vocoder. *Bell System Technical Journal*, 45(9), 1493–1509.
- Flandrin, P. (1998). *Time-frequency/Time-scale analysis*. San Diego: Academic Press.
- Folland, G. B. & Sitaram, A. (1997). The uncertainty principle: a mathematical survey. *Journal of Fourier Analysis and Application*, 3(3), 207–238.
- Gabor, D. (1946). Theory of communication. part 1: The analysis of information. *Journal of the Institution of Electrical Engineers-Part III: Radio and Communication Engineering*, 93(26), 429–441.

- Gabor, D. (1947). Acoustical quanta and the theory of hearing. *Nature*, 159(4044), 591–594.
- Gabor, D. (1952). Lectures on communication theory. Technical report, Massachusetts Institute of Technology, Research Laboratory of Electronics.
- Goodwin, M. & Vetterli, M. (1997). Atomic decompositions of audio signals. In *Applications of Signal Processing to Audio and Acoustics, 1997 IEEE ASSP Workshop on*, (pp. 4–7).
- Griffin, D. & Lim, J. S. (1984). Signal estimation from modified short-time Fourier transform. *Acoustics, Speech and Signal Processing, IEEE Transactions on*, 32(2), 236–243.
- Gröchenig, K. (1993). Acceleration of the frame algorithm. *Signal Processing, IEEE Transactions on*, 41(12), 3331–3340.
- Gröchenig, K. (2001). *Foundations of time-frequency analysis*. Basel: Birkhäuser.
- Gröchenig, K. (2003). Uncertainty principles for time-frequency representations. In *Advances in Gabor Analysis* (pp. 11–30). Basel: Birkhäuser.
- Guillemain, P. & Kronland-Martinet, R. (1996). Characterization of acoustic signals through continuous linear time-frequency representations. *Proceedings of the IEEE*, 84(4), 561–585.
- Hogan, J. & Lakey, J. (2007). *Time-Frequency and Time-Scale Methods: Adaptive Decompositions, Uncertainty Principles, and Sampling*. Basel: Birkhäuser.
- Holighaus, N., Průša, Z., & Søndergaard, P. L. (submitted, 2015). New ideas in reassignment: General time-frequency filter banks, sampling and processing. *IEEE Signal Processing Letters*.
- Jaillet, F. (2005). *Représentation et traitement temps-fréquence des signaux audionumériques pour des applications de design sonore*. PhD thesis, Université de la Méditerranée - Aix-Marseille II.
- Janssen, A. (1981). Gabor representation of generalized functions. *Journal of Mathematical Analysis and Applications*, 83(2), 377–394.
- Kronland-Martinet, R., Morlet, J., & Grossmann, A. (1987). Analysis of sound patterns through wavelet transforms. *International Journal of Pattern Recognition and Artificial Intelligence*, 1(02), 273–302.
- Meyer-Eppler, W. (1959). *Grundlagen und Anwendungen der Informationstheorie*. Berlin: Springer-Verlag.
- Perraudin, N., Balazs, P., & Soendergaard, P. (2013). A fast griffin-lim algorithm. In *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2013 IEEE Workshop on*, (pp. 1–4).
- Průša, Z., Søndergaard, P. L., Holighaus, N., Wiesmeyr, C., & Balazs, P. (2014). The large time-frequency analysis toolbox 2.0. In M. Aramaki, O. Derrien, R. Kronland-Martinet, & S. Ystad (Eds.), *Sound, Music, and Motion*. Berlin: Springer International Publishing.
- Roads, C. (2004). *Microsound*. Cambridge, MA: MIT Press.
- Roads, C., Pope, S., Piccialli, A., & De Poli, G. (1997). *Musical Signal Processing*. Lisse: Swets & Zeitlinger.
- Søndergaard, P. L., Torrésani, B., & Balazs, P. (2012). The linear time frequency analysis toolbox. *International Journal of Wavelets, Multiresolution Analysis and Information Processing*, 10(4).
- Sturm, B. L., Roads, C., McLeran, A., & Shynk, J. J. (2009). Analysis, visualization, and transformation of audio signals using dictionary-based methods. *Journal of New Music Research*, 38(4), 325–341.
- Tempelaars, S. (1996). *Signal Processing, Speech and Music*. Lisse: Swets & Zeitlinger.

- Van Velthoven, J. (2015). *Atomic decompositions of square-integrable functions*. <http://ltfat.sourceforge.net/notes/ltfatnote031.pdf>.
- Van Velthoven, J. & Søndergaard, P. L. (2014). *The Large Time-Frequency Analysis Toolbox Tutorial*. <http://ltfat.sourceforge.net/notes/ltfatnote004.pdf>.
- Xenakis, I. (1960). Elements of stochastic music. *Gravesaner Blätter*, 18, 84–105.
- Xenakis, I. (1963). *Musique formelles*. Paris: Editions Richard-Masse.
- Xenakis, I. (1992). *Formalized Music: Thought and Mathematics in Composition* (Revised ed.). New York: Pendragon Press.

Appendix A

Atomic decompositions of square-integrable functions

A.1 Introduction

A decomposition of a function is a representation of a function in terms of elementary functions. The representation of a function in terms of elementary functions is in general called an *expansion* of a function. A discrete expansion of an arbitrary function f is a representation of f in terms of a *series expansion* of the form

$$f = \sum_{n \in J} c_n f_n,$$

where J denotes a discrete, countable index set. The functions f_n , $n \in J$, are in this case the elementary functions and are called the *expansion functions*. These expansion functions are in general derived by elementary operations on a single, fixed function called the *generator* or *atom*. A collection of expansion functions $\{f_n\}_{n \in J}$ that is used to represent a function as a convergent series is called a *system*. The scalars c_n , $n \in J$, that are associated with the expansion functions are in general called the *coefficients*. The collection of coefficients $\{c_n\}_{n \in J}$ form a representation of the expanded function in a domain that is in general called the *transform domain* associated with the expansion functions. The process of mapping an arbitrary function into the transform domain associated with the expansion functions, is in general called *analysis*. The mapping that maps a function into a collection of coefficients in the transform domain is called the *coefficient mapping*. The process that maps a collection of coefficients from the transform domain into an arbitrary function is in general called *synthesis*. The mapping that constructs a function from a collection of coefficients in the transform domain is called the *reconstruction mapping*. Both the analysis and synthesis of functions in terms of a linear combination of expansion functions belong to the general framework of an atomic decomposition.

A.2 Atomic decompositions

The concept of representing an arbitrary function as a series expansion with respect to a system of expansion functions can be generalized to classes of functions. In this case can *each* element of the class of functions be represented as a series expansion with respect to a system of expansion functions. An example of a general method for such decompositions is an *atomic decomposition*. An atomic decomposition can be defined for general Banach spaces and includes therefore the space of square-integrable functions, the space that is considered in this paper, as a special case.

Definition 1 (Atomic decomposition). Let \mathcal{B} be a Banach space and let \mathcal{B}_d be an associated Banach space of scalar-valued sequences indexed by elements of the countable index set $J \subseteq \mathbb{Z}$. Let $\{f_n\}_{n \in J} \subset \mathcal{B}^*$ and $\{g_n\}_{n \in J} \subset \mathcal{B}$. Then the pair of systems $(\{f_n\}_{n \in J}, \{g_n\}_{n \in J})$ is called an *atomic decomposition* of \mathcal{B} with respect to \mathcal{B}_d if it satisfies the following conditions:

- (i) $\{\langle f, f_n \rangle\} \in \mathcal{B}_d, \quad \forall f \in \mathcal{B},$
- (ii) $\exists 0 < A, B < \infty, \forall f \in \mathcal{B}, A\|f\|_{\mathcal{B}} \leq \|\langle f, f_n \rangle\|_{\mathcal{B}^d} \leq B\|f\|_{\mathcal{B}},$
- (iii) $f = \sum_{n \in J} \langle f, f_n \rangle g_n \quad \forall f \in \mathcal{B}.$

The constants A and B that appear in (ii) called *atomic bounds* of the atomic decomposition $(\{f_n\}_{n \in J}, \{g_n\}_{n \in J})$.

An atomic decomposition allows that any function $f \in \mathcal{B}$ can be represented as

$$f = \sum_{n \in J} \langle f, f_n \rangle g_n. \quad (\text{A.1})$$

This formula is in general called the *reproducing formula* for $f \in \mathcal{B}$. If the series of the reproducing formula associated with an atomic decomposition converges unconditionally, then the atomic decomposition is called *unconditional*. The collection of scalars $\{\langle f, f_n \rangle\}_{n \in J}$ in the reproducing formula are obtained through the coefficient mapping defined by

$$f \mapsto \{\langle f, f_n \rangle\}. \quad (\text{A.2})$$

The definition of this coefficient mapping is allowed by conditions (i) and (ii) of definition 1. If the Banach spaces \mathcal{B} and \mathcal{B}_d of definition 1 are respectively the separable Hilbert space \mathcal{H} and $\ell^2(J)$, then conditions (i) and (ii) of definition 1 imply the existence of a system $\{g_n\}_{n \in J} \subset \mathcal{H}$ that forms together with the system $\{f_n\}_{n \in J} \subset \mathcal{H}$ an atomic decomposition of \mathcal{H} with respect to $\ell^2(J)$. Thus, if the norms $\|f\|_{\mathcal{H}}$ and $\|\langle f, f_n \rangle\|_{\ell^2}$ are equivalent, then the reproducing formula in (A.1) exists for all $f \in \mathcal{H}$. The system $\{f_n\}_{n \in J} \subset \mathcal{H}$ is in this case called a *discrete frame* for \mathcal{H} .

A.2.1 Frame theory

A system $\{f_n\}_{n \in J} \subset \mathcal{H}$ is a frame for the Hilbert space \mathcal{H} if it satisfies the so-called *frame condition*.

Definition 2 (Frame condition). A system $\{f_n\}_{n \in J} \subseteq \mathcal{H}$ is called a *discrete frame* for a separable Hilbert space \mathcal{H} if there exist two constants $A, B > 0$ such that

$$A\|f\|_{\mathcal{H}}^2 \leq \sum_{n \in J} |\langle f, f_n \rangle|^2 \leq B\|f\|_{\mathcal{H}}^2, \quad \forall f \in \mathcal{H}. \quad (\text{A.3})$$

The constants A and B are called the lower respectively the upper frame bound of the frame $\{f_n\}_{n \in J}$.

If a system $\{f_n\}_{n \in J}$ is a frame for \mathcal{H} , then $\overline{\text{span}}\{f_n\}_{n \in J} = \mathcal{H}$. This shows that if a system is a frame for a separable Hilbert space, then it is *complete* in this Hilbert space. The elements $f_n, n \in J$, of a frame $\{f_n\}_{n \in J}$ are called the *frame elements* of $\{f_n\}_{n \in J}$. If $\{f_n\}_{n \in J}$ is a frame, but ceases to be a frame after the removal of any of its frame elements, then $\{f_n\}_{n \in J}$ is called an *exact frame*. If the frame $\{f_n\}_{n \in J}$ is still a frame after the removal of any of its frame elements, then it is called an *overcomplete frame*.

The frame bounds of a frame are inseparable related to the possibility of representing a $f \in \mathcal{H}$ in terms of the frame elements as in the reproducing formula (A.1). However, the frame bounds of a frame are in general not unique. An example of two important frame bounds are the so-called *optimal frame bounds*.

Definition 3. The largest lower frame bound and the smallest upper frame bound are called the *optimal frame bounds*. The optimal lower frame bound A respectively the optimal upper frame bound B are defined as

$$\begin{aligned} A &:= \inf_{f \in \mathcal{H}} \left(\frac{\sum_{n \in J} |\langle f, f_n \rangle|^2}{\|f\|_{\mathcal{H}}^2} \right), \\ B &:= \sup_{f \in \mathcal{H}} \left(\frac{\sum_{n \in J} |\langle f, f_n \rangle|^2}{\|f\|_{\mathcal{H}}^2} \right). \end{aligned}$$

If the frame bounds of a frame are equal, then the frame is called *tight*. For a tight frame $\{f_n\}_{n \in J}$ whose frame bounds are equal to 1, the frame condition becomes Parseval's equality for orthonormal bases,

$$\sum_{n \in J} |\langle f, f_n \rangle|^2 = \|f\|_{\mathcal{H}}^2, \quad \forall f \in \mathcal{H}. \quad (\text{A.4})$$

From this it can be deduced that a orthonormal basis is a tight frame with frame bound 1. It can further be deduced that if a system $\{f_n\}_{n \in J}$ only satisfies the upper bound in the frame condition, then the frame condition becomes *Bessel's inequality*,

$$\sum_{n \in J} |\langle f, f_n \rangle|^2 \leq B \|f\|_{\mathcal{H}}^2, \quad \forall f \in \mathcal{H}. \quad (\text{A.5})$$

In this case is the system $\{f_n\}_{n \in J}$ called a *Bessel sequence* with Bessel bound B .

Operators associated with frames

There are a few important operators associated with frames, including the coefficient operator, the representation operator and the frame operator.

Definition 4. Let $\{f_n\}_{n \in J} \subseteq \mathcal{H}$ be a frame for a Hilbert space \mathcal{H} .

(i) The *coefficient operator* \mathcal{C} associated with $\{f_n\}_{n \in J}$ is the mapping defined by

$$\mathcal{C} : \mathcal{H} \rightarrow \ell^2(J), \quad f \mapsto \mathcal{C}f = \{\langle f, f_n \rangle\}_{n \in J}.$$

(ii) The *representation operator* \mathcal{R} associated with $\{f_n\}_{n \in J}$ is the mapping defined by

$$\mathcal{R} : \ell^2(I) \rightarrow \mathcal{H}, \quad f \mapsto \mathcal{R}f = \sum_{n \in J} c_n f_n.$$

(iii) The *frame operator* \mathcal{S} associated with $\{f_n\}_{n \in J}$ is the mapping defined by

$$\mathcal{S} : \mathcal{H} \rightarrow \mathcal{H}, \quad f \mapsto \mathcal{S}f = \sum_{n \in J} \langle f, f_n \rangle f_n.$$

Several important properties of the operators defined above are given in the following proposition.

Proposition 1. Let $\{f_n\}_{n \in J} \subseteq \mathcal{H}$ be a frame for a Hilbert space \mathcal{H} .

(i) The coefficient operator \mathcal{C} and the representation operator \mathcal{R} are both bounded operators.

(ii) The representation operator \mathcal{R} is the adjoint operator of the coefficient operator \mathcal{C} .

(iii) The frame operator \mathcal{S} is bounded, self-adjoint, positive and invertible.

Proof. (i) Since the frame $\{f_n\}_{n \in J}$ associated with the coefficient operator \mathcal{C} is a Bessel sequence,

$$\|\mathcal{C}f\|_{\ell^2} = \left(\sum_{n \in J} |\langle f, f_n \rangle|^2 \right)^{1/2} \leq B \|f\|_{\mathcal{H}},$$

it can immediately be deduced that \mathcal{C} is bounded. The representation operator \mathcal{R} is bounded since it is the adjoint of \mathcal{C} (see (ii)).

(ii) Let $(c_n)_{n \in J} \subseteq \ell^2(J)$. Then

$$\langle c, \mathcal{C}f \rangle = \sum_{n \in J} c_n \overline{\langle f, f_n \rangle} = \left\langle \sum_{n \in J} c_n f_n, f \right\rangle = \langle \mathcal{R}c, f \rangle.$$

(iii) Since $\mathcal{S} = \mathcal{R}\mathcal{C}$, \mathcal{S} is the composition of two bounded operators and is itself also bounded since

$$\|\mathcal{S}f\| = \|\mathcal{R}\mathcal{C}f\| \leq \|\mathcal{R}\| \|\mathcal{C}f\| \leq \|\mathcal{R}\| \|\mathcal{C}\| \|f\|, \quad \forall f \in \mathcal{H}.$$

That \mathcal{S} is self-adjoint follows from the direct computation

$$\langle \mathcal{S}f, f \rangle = \sum_{n \in J} |\langle f, f_n \rangle|^2 = \langle f, \mathcal{S}f \rangle.$$

That \mathcal{S} is positive follows directly from the frame condition rewritten in the form

$$A\|f\|_{\mathcal{H}}^2 \leq \langle \mathcal{S}f, f \rangle \leq B\|f\|_{\mathcal{H}}^2$$

with $A, B > 0$. Since $A > 0$ and $\mathcal{S}f = 0 \implies f = 0$, \mathcal{S} is invertible. \square

A.2.2 Frame decompositions

The invertibility of the frame operator \mathcal{S} leads to the construction of the system $\{\mathcal{S}^{-1}f_n\}_{n \in J}$. In the following lemma it is proved that this system constitutes a frame, called the *canonical dual* of the frame $\{f_n\}_{n \in J} \subseteq \mathcal{H}$.

Lemma 1. *If $\{f_n\}_{n \in J} \subseteq \mathcal{H}$ is a frame for \mathcal{H} with frame bounds A, B and if \mathcal{S} is the frame operator associated with $\{f_n\}_{n \in J}$, then is $\{\mathcal{S}^{-1}f_n\}_{n \in J}$ a frame with frame bounds B^{-1}, A^{-1} .*

Proof. Since \mathcal{S} is a positive operator it holds that

$$A\|f\|_{\mathcal{H}}^2 \leq \langle \mathcal{S}f, f \rangle \leq B\|f\|_{\mathcal{H}}^2 \iff A\|f\|_{\mathcal{H}} \leq \|\mathcal{S}f\| \leq B\|f\|_{\mathcal{H}}.$$

Let now $f = \mathcal{S}^{-1}g$ where $g \in \mathcal{H}$. Then

$$\begin{aligned} A\|f\|_{\mathcal{H}} \leq \|\mathcal{S}f\| \leq B\|f\|_{\mathcal{H}} &\iff B^{-1}\|\mathcal{S}f\| \leq \|f\| \leq A^{-1}\|\mathcal{S}f\| \\ &\iff B^{-1}\|g\| \leq \|\mathcal{S}^{-1}g\| \leq A^{-1}\|g\| \\ &\iff B^{-1}\|g\|^2 \leq \langle \mathcal{S}^{-1}g, g \rangle \leq A^{-1}\|g\|^2 \\ &\iff B^{-1}\|g\|^2 \leq \sum_{n \in J} |\langle g, \mathcal{S}^{-1}g \rangle|^2 \leq A^{-1}\|g\|^2 \end{aligned}$$

\square

Theorem 1 (Frame decomposition). *If a system $\{f_n\}_{n \in J}$ forms a frame for a Hilbert space \mathcal{H} and \mathcal{S} is the frame operator associated with this frame, then any $f \in \mathcal{H}$ can be represented as the series*

$$f = \sum_{n \in J} \langle f, \mathcal{S}^{-1}f_n \rangle f_n \tag{A.6}$$

$$f = \sum_{n \in J} \langle f, f_n \rangle \mathcal{S}^{-1}f_n \tag{A.7}$$

where both series converge unconditionally in the norm of \mathcal{H} .

Proof. Since the frame operator \mathcal{S} associated with a frame $\{f_n\}_{n \in J}$ is invertible and self-adjoint, it follows from direct computations that

$$f = \mathcal{S}\mathcal{S}^{-1}f = \sum_{n \in J} \langle \mathcal{S}^{-1}f, f_n \rangle f_n = \sum_{n \in J} \langle f, \mathcal{S}^{-1}f_n \rangle f_n$$

$$f = \mathcal{S}^{-1}\mathcal{S}f = \mathcal{S}^{-1} \sum_{n \in J} \langle f, f_n \rangle f_n = \sum_{n \in J} \langle f, f_n \rangle \mathcal{S}^{-1}f_n$$

From the fact that both $\{f_n\}_{n \in J}$ and $\{\mathcal{S}^{-1}f_n\}_{n \in J}$ are Bessel sequences and both $\{\langle \mathcal{S}^{-1}f, f_n \rangle\}_{n \in J}$ and $\{\langle f, f_n \rangle\}_{n \in J}$ are in $\ell^2(J)$, it can be deduced that the series (A.6) respectively (A.7) converge unconditionally. For the unconditional convergence of Bessel sequences, see [8]. \square

The frame decomposition allows an atomic decomposition of \mathcal{H} with respect to $\ell^2(J)$. Furthermore, since the frame decomposition implies unconditional convergence of the series, the atomic decomposition that is connected to the frame decomposition is an unconditional atomic decomposition. The atomic decompositions are in this case the pairs of systems $(\{f_n\}_{n \in J}, \{\mathcal{S}^{-1}f_n\}_{n \in J})$ and $(\{\mathcal{S}^{-1}f_n\}_{n \in J}, \{f_n\}_{n \in J})$ which consist of a frame and its canonical dual.

Remark 1. Two frames that satisfy the reproducing formula in (A.1) are in general called *dual frames*.

Theorem 2. Let $\{f_n\}_{n \in J}$ and $\{g_n\}_{n \in J}$ be Bessel sequences in \mathcal{H} . If $\{f_n\}_{n \in J}$ and $\{g_n\}_{n \in J}$ satisfy the equality

$$\langle f, g \rangle = \sum_{n \in J} \langle f, f_n \rangle \langle g_n, g \rangle, \quad \forall f, g \in \mathcal{H}. \quad (\text{A.8})$$

Then $\{f_n\}_{n \in J}$ and $\{g_n\}_{n \in J}$ are dual frames for \mathcal{H} .

Proof. That a Bessel sequences $\{f_n\}_{n \in J}$ that satisfies (A.8) is a frame follows from the computation

$$\|f\|_{\mathcal{H}}^2 = \langle f, f \rangle = \sum_{n \in J} \langle f, f_n \rangle \langle g_n, f \rangle \leq \sqrt{\sum_{n \in J} |\langle f, f_n \rangle|^2} \sqrt{\sum_{n \in J} |\langle f, g_n \rangle|^2} \leq \sqrt{\sum_{n \in J} |\langle f, f_n \rangle|^2} \sqrt{B} \|f\|_{\mathcal{H}},$$

which is equivalent to

$$B^{-1} \|f\|_{\mathcal{H}}^2 \leq \sum_{n \in J} |\langle f, f_n \rangle|^2.$$

In the same way it can be shown that the Bessel sequence $\{g_n\}_{n \in J}$ satisfies the lower frame bound B^{-1} through which it becomes a frame. That the frames $\{f_n\}_{n \in J}$ and $\{g_n\}_{n \in J}$ are dual frames follows from

$$0 = \langle f, g \rangle - \sum_{n \in J} \langle f, f_n \rangle \langle g_n, g \rangle = \left\langle f - \sum_{n \in J} \langle f, f_n \rangle g_n, g \right\rangle, \quad \forall g \in \mathcal{H},$$

which shows that $f = \sum_{n \in J} \langle f, f_n \rangle g_n$. \square

A dual frame of a frame that isn't the canonical dual is called an *alternative dual* of the frame. The existence of alternative dual frames is related to the completeness of the frame.

Theorem 3. Let $\{f_n\}_{n \in J}$ be an overcomplete frame. Then there exists a frame $\{g_n\}_{n \in J} \neq \{\mathcal{S}^{-1}f_n\}_{n \in J}$ that satisfy

$$f = \sum_{n \in J} \langle f, g_n \rangle f_n, \quad \forall f \in \mathcal{H}.$$

Proof. See [2]. \square

A.3 Fourier expansion

The theory of a Fourier expansion is based on the notion that any periodic function can be decomposed into a superposition of sinusoids. A function f on \mathbb{R} is called a \mathbb{Z} -periodic or 1-periodic function if

$$f(t) = f(t + p), \quad \forall p \in \mathbb{Z}. \quad (\text{A.9})$$

These functions can be represented as functions on the closed interval $[0, 1]$ or as functions on the torus $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and are therefore elements of $L^2[0, 1]$ or $L^2(\mathbb{T})$. An example of an element of $L^2(\mathbb{T})$ is the complex exponential $e^{i2\pi kt}$ with $k \in \mathbb{Z}$ and $t \in \mathbb{T}$. A collection of such complex exponentials given by

$$\{e_k\}_{k \in \mathbb{Z}} = \{e^{i2\pi kt}\}_{k \in \mathbb{Z}}, \quad t \in \mathbb{T} \quad (\text{A.10})$$

forms a Parseval frame for $L^2(\mathbb{T})$ [8]

A.3.1 Operators associated with Fourier expansions

The coefficient operator associated with the frame $\{e_k\}_{k \in \mathbb{Z}}$ is given by

$$\mathcal{F} : L^2(\mathbb{T}) \rightarrow \ell^2(\mathbb{Z}), \quad \mathcal{F}f = \{\langle f, e_k \rangle\}_{k \in \mathbb{Z}} \quad (\text{A.11})$$

The coefficient operator \mathcal{F} maps a function f from the so-called *time domain* into the *frequency domain*. The collection of scalars $\{\langle f, e_k \rangle\}_{k \in \mathbb{Z}}$ in the frequency domain are called the *Fourier coefficients* of f . The adjoint of the coefficient operator \mathcal{F} , \mathcal{F}^* , is given by

$$\mathcal{F}^* : \ell^2(\mathbb{Z}) \rightarrow L^2(\mathbb{T}), \quad \mathcal{F}^*c = \sum_{k \in \mathbb{Z}} c_k e_k \quad (\text{A.12})$$

and is called the representation operator associated with the frame $\{e_k\}_{k \in \mathbb{Z}}$. This representation operator \mathcal{F}^* maps a collection of scalars $\{c_k\}_{k \in \mathbb{Z}} \subset \ell^2(\mathbb{Z})$ into a function $f \in L^2(\mathbb{T})$. The concatenation of \mathcal{F}^* and \mathcal{F} , that is, $\mathcal{F}^* \mathcal{F}$, is the frame operator associated with the frame $\{e_k\}_{k \in \mathbb{Z}}$ and is given by

$$\mathcal{F}^* \mathcal{F}f = \sum_{k \in \mathbb{Z}} \langle f, e_k \rangle e_k \quad (\text{A.13})$$

The series in (A.13) is in general called a *Fourier series*.

Theorem 4. *Let $\{e_k\}_{k \in \mathbb{Z}} = \{e^{i2\pi kt}\}_{k \in \mathbb{Z}}$ with $t \in \mathbb{T}$, then any $f \in L^2(\mathbb{T})$ can be represented as*

$$f = \sum_{k \in \mathbb{Z}} \langle f, e_k \rangle e_k$$

where the Fourier series converges unconditionally in the L^2 -norm.

Proof. Since $\{e_k\}_{k \in \mathbb{Z}}$ is a Bessel sequence and since $\{\langle f, e_k \rangle\}_{k \in \mathbb{Z}} \subset \ell^2(\mathbb{Z})$, the Fourier series defined through the frame operator associated with $\{e_k\}_{k \in \mathbb{Z}}$ converges unconditionally. \square

A.4 Gabor expansion

Through the frame $\{e^{i2\pi kt}\}_{k \in \mathbb{Z}}$ any $f \in L^2(\mathbb{T})$ could be decomposed into a linear combination of complex exponentials with associated coefficients. A sequence of such frames can describe any square-integrable function on the interval $(-\infty, \infty)$, that is, any $f \in L^2(\mathbb{R})$.

A.4.1 Gabor frames

A Gabor expansion of a function $f \in L^2(\mathbb{R})$ into a sequence of linear combinations of complex exponentials with associated coefficients is based on the translation and modulation of a *generator*. Under certain conditions any $g \in L^2(\mathbb{R})$ could be such a generator, but a generator is admissible if it belongs to the so-called *Segal algebra* or *Feichtinger algebra* $S_0(\mathbb{R})$.

Definition 5 (Feichtinger algebra). The function space $S_0(\mathbb{R})$ defined as

$$S_0(\mathbb{R}) = \left\{ g \in L^2(\mathbb{R}) \mid \iint_{\mathbb{R}} \left| \int_{\mathbb{R}} g(t) e^{-\pi(t-\tau)^2} e^{-i2\pi\xi t} dt \right| d\tau d\xi < \infty \right\} \quad (\text{A.14})$$

is called the *Feichtinger algebra* $S_0(\mathbb{R})$.

Remark 2. The Feichtinger algebra $S_0(\mathbb{R})$ is a Banach space with respect to the norm

$$\|g\|_{S_0} = \iint_{\mathbb{R}} \left| \int_{\mathbb{R}} g(t) e^{-\pi(t-\tau)^2} e^{-i2\pi\xi t} dt \right| d\tau d\xi, \quad (\text{A.15})$$

and is a dense subspace in $L^2(\mathbb{R})$ [6].

Translated and modulated versions of a generator $g \in S_0(\mathbb{R})$ are obtained through the translation and modulation operator on $L^2(\mathbb{R})$.

Definition 6. Let $g \in L^2(\mathbb{R})$.

(i) The translation operator \mathcal{T}_τ , that translates a function by $\tau \in \mathbb{R}$, is given by

$$\mathcal{T}_\tau : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}), \quad (\mathcal{T}_\tau g)(t) = g(t - \tau), \quad \tau \in \mathbb{R}.$$

(ii) The modulation operator \mathcal{E}_ξ , that translates a function by $\xi \in \mathbb{R}$, is given by

$$\mathcal{E}_\xi : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}), \quad (\mathcal{E}_\xi g)(t) = g(t) e^{i2\pi\xi t}, \quad \xi \in \mathbb{R}.$$

A fundamental property of the translation and modulation operators, their so-called *commutation relation*, is proved in the next lemma.

Lemma 2. For all $\tau, \xi \in \mathbb{R}$ the following hold:

$$(\mathcal{T}_\tau \mathcal{E}_\xi g)(t) = e^{-i2\pi\xi t} (\mathcal{E}_\xi \mathcal{T}_\tau g)(t). \quad (\text{A.16})$$

Proof. The relation follows from the direct computation,

$$\begin{aligned} (\mathcal{T}_\tau \mathcal{E}_\xi g)(t) &= \mathcal{E}_\xi g(t - \tau) \\ &= e^{i2\pi\xi(t-\tau)} g(t - \tau) \\ &= e^{-i2\pi\xi\tau} e^{i2\pi\xi t} g(t - \tau) \\ &= e^{-i2\pi\xi\tau} \mathcal{E}_\xi \mathcal{T}_\tau g(t) \end{aligned}$$

□

Next, a version of a translated and modulated generator $g \in L^2(\mathbb{R})$ is given by

$$(\mathcal{E}_\xi \mathcal{T}_\tau g)(t) = g(t - \tau) e^{i2\pi\xi t}. \quad (\text{A.17})$$

A natural discretization of such a translated and modulated generator is $\tau = n\alpha$, $\xi = m\beta$ where $\alpha, \beta > 0$ are fixed and $n, m \in \mathbb{Z}$. Such a subsampled version of a translated and modulated generator is called a *Gabor atom*.

Definition 7. Let $g \in L^2(\mathbb{R})$ and $\alpha, \beta \in \mathbb{R}^+$. The collection of functions given by

$$\mathcal{G}(g, \alpha, \beta) = \{g_{m,n}\}_{m,n \in \mathbb{Z}}, \quad g_{m,n} := \mathcal{E}_{m\beta} \mathcal{T}_{n\alpha} g,$$

is called a *Gabor system*.

Definition 8. A Gabor system $\mathcal{G}(g, \alpha, \beta)$ that satisfies the frame condition (A.3) is called a *Gabor frame*.

Conditions and properties for Gabor frames

There are several sufficient conditions for a Gabor system to form a Gabor frame for $L^2(\mathbb{R})$. One of these sufficient conditions is stated in [1].

Theorem 5. *Let $g \in L^2(\mathbb{R})$ and $\alpha, \beta \in \mathbb{R}^+$. Suppose that*

$$A := \inf_{t \in [0, \alpha]} \left(\sum_{n \in \mathbb{Z}} |g(t - n\alpha)|^2 - \sum_{n \neq 0} \left| \sum_{n \in \mathbb{Z}} g(t - n\alpha) \overline{g(t - n\alpha - \frac{k}{\beta})} \right| \right) > 0,$$

$$B := \sup_{t \in [0, \alpha]} \sum_{n \in \mathbb{Z}} \left| \sum_{n \in \mathbb{Z}} g(t - n\alpha) \overline{g(t - n\alpha - \frac{k}{\beta})} \right| < \infty.$$

Then $\{\mathcal{E}_{mb}\mathcal{T}_{n\alpha}g\}_{m,n \in \mathbb{Z}}$ forms a frame for $L^2(\mathbb{R})$ with frame bounds $\frac{A}{\beta}$ and $\frac{B}{\beta}$.

Proof. See [1]. □

The completeness of a Gabor system $\{\mathcal{E}_{mb}\mathcal{T}_{n\alpha}g\}_{m,n \in \mathbb{Z}}$ with $\alpha, \beta \in \mathbb{R}^+$ is related to the density of the lattice $\Lambda = \alpha\mathbb{Z} \times \beta\mathbb{Z}$ and is therefore determined by the product of the parameters α, β .

Theorem 6. *Let $\mathcal{G}(g, \alpha, \beta)$ be a Gabor system. Then the following hold:*

- (i) *If $\alpha\beta \leq 1$, then $\mathcal{G}(g, \alpha, \beta)$ is overcomplete in $L^2(\mathbb{R})$,*
- (ii) *If $\alpha\beta = 1$, then $\mathcal{G}(g, \alpha, \beta)$ is complete in $L^2(\mathbb{R})$,*
- (iii) *If $\alpha\beta > 1$, then $\mathcal{G}(g, \alpha, \beta)$ is incomplete in $L^2(\mathbb{R})$.*

Proof. See [7]. □

A.4.2 Operators associated with Gabor frames

In this section the Gabor system $\mathcal{G}(g, \alpha, \beta)$ is assumed to form a Gabor frame for $L^2(\mathbb{R})$.

The coefficient operator \mathcal{C} associated with $\mathcal{G}(g, \alpha, \beta)$ is explicitly given by

$$\mathcal{C} : L^2(\mathbb{R}) \rightarrow \ell^2(\mathbb{Z}^2), \quad \mathcal{C}f = \{\langle f, \mathcal{E}_{mb}\mathcal{T}_{n\alpha}g \rangle\}_{m,n \in \mathbb{Z}}. \quad (\text{A.18})$$

This coefficient operator maps a function f into a collection of scalars in the so-called *time-frequency domain*. The scalars $\{\langle f, \mathcal{E}_{mb}\mathcal{T}_{n\alpha}g \rangle\}_{m,n \in \mathbb{Z}}$ are called the *Gabor coefficients* associated with $\mathcal{G}(g, \alpha, \beta)$ and are points on the lattice $\Lambda := \alpha\mathbb{Z} \times \beta\mathbb{Z}$.

The adjoint of the coefficient operator \mathcal{C} associated with $\mathcal{G}(g, \alpha, \beta)$ is the reconstruction operator \mathcal{R} associated $\mathcal{G}(g, \alpha, \beta)$. This reconstruction operator is given by

$$\mathcal{R} : \ell^2(\mathbb{Z}^2) \rightarrow L^2(\mathbb{R}), \quad \mathcal{R} = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{m,n} \mathcal{E}_{mb}\mathcal{T}_{n\alpha}g. \quad (\text{A.19})$$

This reconstruction operator expands a function in terms of a collection of scalars in the time-frequency domain. The concatenation of \mathcal{C} and \mathcal{R} , that is, $\mathcal{R}^*\mathcal{C}$, is the frame operator \mathcal{S} associated with the $\mathcal{G}(g, \alpha, \beta)$ and is called the *Gabor frame operator*. The Gabor frame operator is given by

$$\mathcal{S} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}), \quad \mathcal{S}f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, \mathcal{E}_{mb}\mathcal{T}_{n\alpha}g \rangle \mathcal{E}_{mb}\mathcal{T}_{n\alpha}g \quad (\text{A.20})$$

An important property of the Gabor frame operator is its commutation with the operators \mathcal{E}_{mb} and $\mathcal{T}_{n\alpha}$.

Lemma 3. *Let $\mathcal{G}(g, \alpha, \beta)$ be a Gabor frame and let \mathcal{S} be the associated frame operator, then*

(i)

$$\mathcal{S}\mathcal{E}_{mb}\mathcal{T}_{n\alpha} = \mathcal{E}_{mb}\mathcal{T}_{n\alpha}\mathcal{S}, \quad \forall m, n \in \mathbb{Z}.$$

(ii)

$$\mathcal{S}^{-1}\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha} = \mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}\mathcal{S}^{-1}, \quad \forall m, n \in \mathbb{Z}.$$

Proof. (i) From a direct computation, using the commutator relation expressed in (A.16), it can be deduced that for any $f \in L^2(\mathbb{R})$,

$$\begin{aligned} \mathcal{S}\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}f &= \\ &= \sum_{m' \in \mathbb{Z}} \sum_{n' \in \mathbb{Z}} \langle \mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}f, \mathcal{E}_{m'\beta}\mathcal{T}_{n'\alpha}g \rangle \mathcal{E}_{m'\beta}\mathcal{T}_{n'\alpha}g \\ &= \sum_{m' \in \mathbb{Z}} \sum_{n' \in \mathbb{Z}} \langle f, \mathcal{T}_{-n\alpha}\mathcal{E}_{(m'-m)\beta}\mathcal{T}_{n'\alpha}g \rangle \mathcal{E}_{m'\beta}\mathcal{T}_{n'\alpha}g \\ &= \sum_{m' \in \mathbb{Z}} \sum_{n' \in \mathbb{Z}} \langle f, e^{2\pi in\alpha(m'-m)\beta} \mathcal{E}_{(m'-m)\beta}\mathcal{T}_{(n'-n)\alpha}g \rangle \mathcal{E}_{m'\beta}\mathcal{T}_{n'\alpha}g \\ &= \sum_{m' \in \mathbb{Z}} \sum_{n' \in \mathbb{Z}} e^{2\pi in\alpha m'\beta} \langle f, \mathcal{E}_{m'\beta}\mathcal{T}_{n'\alpha}g \rangle \mathcal{E}_{(m'-m)\beta}\mathcal{T}_{(n'-n)\alpha}g, \quad m' + m := m', \quad n' + n := n' \\ &= \sum_{m' \in \mathbb{Z}} \sum_{n' \in \mathbb{Z}} e^{2\pi in\alpha m'\beta} \langle f, \mathcal{E}_{m'\beta}\mathcal{T}_{n'\alpha}g \rangle \mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}\mathcal{E}_{m'\beta}\mathcal{T}_{n'\alpha}g \\ &= \mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}\mathcal{S} \end{aligned}$$

(ii) Since \mathcal{S} is a homeomorphism, it can be replaced by \mathcal{S}^{-1} in (i). \square

Due to the commutativity of the inverse of the Gabor frame operator and the translation and modulation operator, the canonical dual of a Gabor frame $\{\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}g\}_{m,n \in \mathbb{Z}}$ is explicitly given by

$$\mathcal{S}^{-1}\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}g = \mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}\mathcal{S}^{-1}g, \quad (\text{A.21})$$

where $\gamma = \mathcal{S}^{-1}g$ is called the *canonical dual generator*. This shows that the canonical dual of a Gabor frame is again a Gabor system, namely the Gabor frame $\{\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}\gamma\}_{m,n \in \mathbb{Z}}$.

Theorem 7. *If a Gabor system $\{\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}g\}_{m,n \in \mathbb{Z}}$ forms a frame for $L^2(\mathbb{R})$, then there exists a Gabor frame for $L^2(\mathbb{R})$ given by $\{\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}\gamma\}_{m,n \in \mathbb{Z}}$ such that any $f \in L^2(\mathbb{R})$ can be represented as the series*

$$f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, \mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}g \rangle \mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}\gamma \quad (\text{A.22})$$

$$f = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, \mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}\gamma \rangle \mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}g \quad (\text{A.23})$$

where both series converges unconditionally in the L^2 -norm.

Proof. Since, by assumption, the system $\{\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}g\}_{m,n \in \mathbb{Z}}$ forms a frame for $L^2(\mathbb{R})$ its canonical dual is given by $\{\mathcal{S}^{-1}\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}g\}_{m,n \in \mathbb{Z}}$. From lemma 3 it can be deduced that this canonical dual frame has the form $\{\mathcal{E}_{m\beta}\mathcal{T}_{n\alpha}\gamma\}_{m,n \in \mathbb{Z}}$ where $\gamma = \mathcal{S}^{-1}g$. Since both (A.22) and (A.23) are frame decompositions their unconditional convergence follows from theorem 1. \square

This theorem provides two unconditional atomic decompositions of $L^2(\mathbb{R})$ with respect to $\ell^2(\mathbb{Z}^2)$, namely the ones given that consist of the Gabor frames $\mathcal{G}(g, \alpha, \beta)$ and $\mathcal{G}(\mathcal{S}^{-1}g, \alpha, \beta)$. In general, any two Gabor systems $\mathcal{G}(g, \alpha, \beta)$ and $\mathcal{G}(h, \alpha, \beta)$ form an atomic decomposition if they are dual frames of each other. In this case are the generators $g, h \in L^2(\mathbb{R})$ called *dual generators*, and is it possible to represent any $f \in L^2(\mathbb{R})$ as the series (A.22) and (A.23). Two Gabor frames are dual frames if they satisfy the general general duality condition for frames, but another condition for two Gabor systems to be dual frames is the so-called *Wexler-Raz duality condition*.

Theorem 8 (Wexler-Raz). *Let $\mathcal{G}(g, \alpha, \beta)$ and $\mathcal{G}(h, \alpha, \beta)$ be two Gabor frames. The systems are dual frames if and only if*

$$\frac{\langle h, \mathcal{T}_{\frac{n}{\alpha}} \mathcal{E}_{\frac{m}{\beta}} g \rangle}{\alpha\beta} = \delta_{n,0}\delta_{m,0}, \quad m, n \in \mathbb{Z}. \quad (\text{A.24})$$

Proof. See [2]. □

Any two Gabor frames that satisfy (A.24) form an atomic decomposition of $L^2(\mathbb{R})$ with respect to $\ell^2(\mathbb{Z}^2)$.

A.5 Wavelet expansion

The theory of wavelets is based on the notion that a function can be constructed from translated and dilated versions of a generator. Such a generator ψ is in general a square-integrable function and is called *admissible* if it satisfies the *Calderón admissible condition*.

Definition 9. Let $\psi \in L^2(\mathbb{R})$. The admissible constant associated with ψ is defined as

$$C_\psi := \int_{\mathbb{R} \setminus \{0\}} \frac{|\hat{\psi}(\xi)|^2}{|\xi|} d\xi, \quad (\text{A.25})$$

where $\hat{\psi}(\xi)$ denotes the Fourier transform of $\psi(t)$. If $C_\psi \in \mathbb{R}^+$, then ψ is called *admissible*.

Remark 3. Another form to express the admissibility of $\psi \in L^2(\mathbb{R})$ is that it should satisfy

$$\frac{\hat{\psi}(\xi)}{\sqrt{\xi}} \in L^2(\mathbb{R}). \quad (\text{A.26})$$

Translated and dilated versions of a generator $\psi \in L^2(\mathbb{R})$ are obtained by applying the translation operator and dilation operator on ψ .

Definition 10. Let $\psi \in L^2(\mathbb{R})$. The dilation operator \mathcal{D}_v , that dilates a function by $v \in \mathbb{R} \setminus \{0\}$, is given by

$$\mathcal{D}_v : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}), \quad (\mathcal{D}_v \psi)(t) = \frac{1}{\sqrt{|v|}} \psi\left(\frac{t}{v}\right), \quad v \in \mathbb{R} \setminus \{0\}.$$

The commutation relation of the translation and dilation operator is given in the next lemma.

Lemma 4. *For all $\tau \in \mathbb{R}$ and $v \in \mathbb{R} \setminus \{0\}$ the following hold:*

$$(\mathcal{T}_\tau \mathcal{D}_v \psi)(t) = (\mathcal{D}_v \mathcal{T}_{\tau/v} \psi)(t).$$

Proof. The relation follows from the direct computation,

$$\begin{aligned} (\mathcal{T}_\tau \mathcal{D}_v \psi)(t) &= \mathcal{T}_\tau \left(\frac{1}{\sqrt{|v|}} \psi\left(\frac{t}{v}\right) \right), \\ &= \frac{1}{\sqrt{|v|}} \psi\left(\frac{t}{v} - \frac{\tau}{v}\right), \\ &= (\mathcal{D}_v \mathcal{T}_{\tau/v} \psi)(t). \end{aligned}$$

□

A version of a translated and dilated generator ψ is

$$(\mathcal{T}_\tau \mathcal{D}_v \psi)(t) = \frac{1}{\sqrt{|v|}} \psi \left(\frac{t - \tau}{v} \right). \quad (\text{A.27})$$

A discretized version of such a translated and dilated generator can be obtained by setting $v = v^j$, $\tau = n\alpha v^j$ with $j, n \in \mathbb{Z}$ and $v > 1, \alpha > 0$. Such a discretized version of a translated and dilated wavelet is called a *wavelet atom*

Definition 11. Let $\psi \in L^2(\mathbb{R})$ and $v > 1, \alpha > 0$. The collection of functions given by

$$\mathcal{W}(\psi, v, \alpha) = \{\psi_{j,n}\}_{j,n \in \mathbb{Z}}, \quad \psi_{j,n} := \mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \psi.$$

Definition 12. A wavelet system $\mathcal{W}(\psi, v, \alpha)$ that satisfies the frame condition (A.3) is called a *wavelet frame*.

There are several sufficient conditions for the system $\{\mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \psi\}_{j,n \in \mathbb{Z}}$ to form a wavelet frame for $L^2(\mathbb{R})$. One of these sufficient conditions is stated in [1].

Theorem 9. Let $\psi \in L^2(\mathbb{R})$ and $\alpha > 0, \beta > 1$. Suppose that

$$A := \inf_{|\xi| \in [1, v]} \left(\sum_{j \in \mathbb{Z}} |\hat{\psi}(v^j \xi)|^2 - \sum_{n \neq 0} \sum_{j \in \mathbb{Z}} |\hat{\psi}(v^j \xi) \hat{\psi}(v^j \xi + \frac{n}{\alpha})| \right) > 0,$$

$$B := \sup_{|\xi| \in [1, v]} \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} |\hat{\psi}(v^j \xi) \hat{\psi}(v^j \xi + \frac{n}{\alpha})| < \infty.$$

Then $\{\mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \psi\}_{j,n \in \mathbb{Z}}$ forms a frame for $L^2(\mathbb{R})$ with frame bounds $\frac{A}{\alpha}$ and $\frac{B}{\alpha}$.

Proof. See [2]. □

A.5.1 Operators associated with wavelet frames

In this section the wavelet system $\mathcal{W}(\psi, v, \alpha)$ is assumed to form a wavelet frame for $L^2(\mathbb{R})$.

The coefficient operator \mathcal{C} associated with $\mathcal{W}(\psi, v, \alpha)$ is given by

$$\mathcal{W} : L^2(\mathbb{R}) \rightarrow \ell^2(\mathbb{Z}^2), \quad \mathcal{W}f = \{\langle f, \mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \psi \rangle\}_{j,n \in \mathbb{Z}} \quad (\text{A.28})$$

This coefficient operator maps a function f into a collection of scalars $\{\langle f, \mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \psi \rangle\}_{j,n \in \mathbb{Z}}$ in the so-called *time-scale plane*. This collection of scalars are called the wavelet coefficients associated with $\mathcal{W}(\psi, v, \alpha)$ and are points on the lattice $\Gamma := \alpha v^j \mathbb{Z} \times v^j \mathbb{Z}$.

The adjoint of the coefficient operator \mathcal{C} associated with $\mathcal{W}(\psi, v, \alpha)$ is the reconstruction operator \mathcal{R} associated with $\mathcal{W}(\psi, v, \alpha)$. This reconstruction operator is given by

$$\mathcal{R} : \ell^2(\mathbb{Z}^2) \rightarrow L^2(\mathbb{R}), \quad \mathcal{R}c = \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} c_{j,n} \mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \psi. \quad (\text{A.29})$$

This synthesis operator constructs a function f from a collection of scalars in the time-scale plane. The concatenation of \mathcal{C} and \mathcal{R} , that is, \mathcal{RC} , is the frame operator \mathcal{S} associated with $\mathcal{W}(\psi, v, \alpha)$. This frame operator is explicitly given by

$$\mathcal{S} : L^2(\mathbb{R}) \rightarrow L^2(\mathbb{R}), \quad \mathcal{S}f = \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, \mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \psi \rangle \mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \psi. \quad (\text{A.30})$$

In general, the frame operator associated with a wavelet frame, and its inverse, doesn't commute with the translation operator $\mathcal{T}_{n\alpha v^j}$. For this reason the next inequality holds in general,

$$\mathcal{S}^{-1} \mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \psi \neq \mathcal{T}_{n\alpha v^j} \mathcal{D}_{v^j} \mathcal{S}^{-1} \psi. \quad (\text{A.31})$$

This inequality shows that the canonical dual of a wavelet frame doesn't need to be a wavelet frame itself.

As opposed to the translation operator $\mathcal{T}_{n\alpha v^j}$, the frame operator associated with a wavelet frame does commute with the dilation operator \mathcal{D}_{v^j} .

Lemma 5. Let \mathcal{D}_{v^j} be the dilation operator and \mathcal{S} be the frame operator associated with a wavelet frame, then

(i)

$$\mathcal{S}\mathcal{D}_{v^j} = \mathcal{D}_{v^j}\mathcal{S}, \quad \forall j \in \mathbb{Z}.$$

(ii)

$$\mathcal{S}^{-1}\mathcal{D}_{v^j} = \mathcal{D}_{v^j}\mathcal{S}^{-1}, \quad \forall j \in \mathbb{Z}.$$

Proof. (i) First, note that for a wavelet frame for $L^2(\mathbb{R})$ it holds that

$$\mathcal{T}_{n\alpha v^j}\mathcal{D}_{v^j}\psi = \mathcal{D}_{v^j}\mathcal{T}_{n\alpha}\psi.$$

Next, it can be deduced from a direct computation that for all $f \in L^2(\mathbb{R})$,

$$\begin{aligned} \mathcal{S}\mathcal{D}_{v^j}f &= \sum_{j' \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle \mathcal{D}_{v^j}f, \mathcal{D}_{v^{j'}}\mathcal{T}_{n\alpha}\psi \rangle \mathcal{D}_{v^{j'}}\mathcal{T}_{n\alpha}\psi, \\ &= \sum_{j' \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, \mathcal{D}_{v^{j'-j}}\mathcal{T}_{n\alpha}\psi \rangle \mathcal{D}_{v^{j'}}\mathcal{T}_{n\alpha}\psi, \\ &= \sum_{j' \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, \mathcal{D}_{v^{j'}}\mathcal{T}_{n\alpha}\psi \rangle \mathcal{D}_{v^{j+j'}}\mathcal{T}_{n\alpha}\psi, \\ &= \mathcal{D}_{v^j}\mathcal{S}f. \end{aligned}$$

Here the second equality followed from the alternative form of the wavelet coefficients, namely

$$\begin{aligned} \langle f, \mathcal{D}_{v^j}\mathcal{T}_{n\alpha}\psi \rangle &= \frac{1}{v^{j/2}} \int_{\mathbb{R}} f(t)\overline{\psi(\frac{t}{v^j} - n\alpha)} dt \\ &= v^{j/2} \int_{\mathbb{R}} f(v^j t)\overline{\psi(t - n\alpha)} dt \\ &= \langle \mathcal{D}_{v^{-j}}f, \mathcal{T}_{n\alpha}\psi \rangle. \end{aligned}$$

(ii) Since \mathcal{S} is a homeomorphism, it can be replaced by \mathcal{S}^{-1} in (i). □

Theorem 10. If a wavelet system $\{\mathcal{T}_{n\alpha v^j}\mathcal{D}_{v^j}\psi\}_{j,n \in \mathbb{Z}}$ forms a frame for $L^2(\mathbb{R})$, then there exists a frame $\{\mathcal{D}_{v^j}\mathcal{S}^{-1}\mathcal{T}_{n\alpha}\psi\}_{j,n \in \mathbb{Z}}$ such that any $f \in L^2(\mathbb{R})$ can be represented as the series

$$f = \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, \mathcal{T}_{n\alpha v^j}\mathcal{D}_{v^j}\psi \rangle \mathcal{D}_{v^j}\mathcal{S}^{-1}\mathcal{T}_{n\alpha}\psi \tag{A.32}$$

$$f = \sum_{j \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} \langle f, \mathcal{D}_{v^j}\mathcal{S}^{-1}\mathcal{T}_{n\alpha}\psi \rangle \mathcal{T}_{n\alpha v^j}\mathcal{D}_{v^j}\psi \tag{A.33}$$

where both series converges unconditionally in the L^2 -norm.

Proof. Since, by assumption, the system $\{\mathcal{T}_{n\alpha v^j}\mathcal{D}_{v^j}\psi\}_{j,n \in \mathbb{Z}}$ forms a frame for $L^2(\mathbb{R})$ its canonical dual is given by $\{\mathcal{S}^{-1}\mathcal{T}_{n\alpha v^j}\mathcal{D}_{v^j}\psi\}_{j,n \in \mathbb{Z}}$. Since $\{\mathcal{T}_{n\alpha v^j}\mathcal{D}_{v^j}\psi\}_{j,n \in \mathbb{Z}}$ is equal to $\{\mathcal{D}_{v^j}\mathcal{T}_{n\alpha}\psi\}_{j,n \in \mathbb{Z}}$ it follows from lemma 5 that the canonical dual has the form $\{\mathcal{D}_{v^j}\mathcal{S}^{-1}\mathcal{T}_{n\alpha}\psi\}_{j,n \in \mathbb{Z}}$. Since both (A.32) and (A.33) are frame decompositions their unconditional convergence follows from theorem 1. □

This theorem provides two unconditional atomic decompositions of $L^2(\mathbb{R})$ with respect to $\ell^2(\mathbb{Z}^2)$, namely the ones given by a wavelet frame $\mathcal{W}(\psi, v, \alpha)$ and its canonical dual. Note that the canonical dual frame is not generator from one single generator ψ , but from a collection of generators $\tilde{\psi}_{0,n} = \mathcal{S}^{-1}\mathcal{T}_{n\alpha}\psi$. In general, any dual frame of $\mathcal{W}(\psi, v, \alpha)$ forms together with this wavelet frame an atomic decomposition.

References of Appendix A

- [1] Casazza, P.G. and Christensen, O. Weyl-Heisenberg frames for subspaces of $L^2(\mathbb{R})$. *Proceedings of the American Mathematical Society*, 129, 145-154, 2001.
- [2] Christensen, O. (2003). *An Introduction to Frames and Riesz Bases*. Basel: Birkhauser.
- [3] Daubechies, I. *Ten Lectures on Wavelets*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 1992.
- [4] Daubechies, I. and Han, B. The Canonical Dual Frame of a Wavelet Frame. *Applied and Computational Harmonic analysis*, 12, 269-285, 2002.
- [5] Feichtinger, H. and Gröchenig, K. Banach spaces related to integrable group representations and their atomic decompositions, I. *Journal of functional analysis*, 86, 307-340, 1989.
- [6] Feichtinger, H. and Zimmerman, G. A Banach space of test functions for Gabor analysis. In H.Feichtinger and T.Strohmer (eds.), *Gabor analysis and algorithms - Theory and applications*, pp. 123-170. Birkhauser, 1998.
- [7] Gröchenig, K. *Foundations of Time-Frequency Analysis*. Birkhauser, 2001.
- [8] Heil, C. *A Basis Theory Primer*. Birkhauser, 2010.

Appendix B

Contents of accompanying CD

The accompanying CD contains two subfolders with different types of files. The subfolders of which the CD consists are:

Audio This subfolder consists of a collection of sound examples that support the conclusions made in this thesis. This collection is accompanied by a document that describes each sound example.

Text This subfolder consists of a digital version of the thesis.